The London School of Economics and Political Science

The Bayesian and the Realist: Friends or Foes?

Eleftherios Farmakis

A thesis submitted to the Department of Philosophy, Logic and Scientific Method of the London School of Economics for the degree of Doctor of Philosophy, London, November 2007

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

The main purpose of my thesis is to bring together two seemingly unrelated topics in the philosophy of science and extract the philosophical consequences of this exercise. The first topic is Bayesianism - a well-developed, and popular, probabilistic theory of confirmation. The second topic is Scientific Realism - the thesis that we have good reason to believe that our best scientific theories are (approximately) true. It seems natural to assume that a sophisticated probabilistic theory of confirmation is the most appropriate framework for the treatment of the issue of scientific realism. Despite this intuition, however, the bulk of the literature is conspicuous for its failure to apply the Bayesian apparatus when discussing scientific realism. Furthermore, on the rare occasions that this has been attempted, its outcomes have been strikingly negative. In my thesis I systematise and critically examine the segmented literature in order to investigate whether, and how, Bayesianism and scientific realism can be reconciled. I argue for the following claims: 1) that those realists who claim that Bayesians lack a proper notion of 'theory acceptance' have misunderstood the nature of Bayesianism as a reductive account of 'theory acceptance'; 2) that it is possible to reconstruct most of the significant alternative positions involved in the realism debate using this new account of 'theory acceptance'; 3) that Bayesianism is best seen as a general framework within which the standard informal arguments for and against realism become transparent, thus greatly clarifying the force of the realist argument; 4) that a Bayesian reconstruction does *not* commit one to any particular position as ultimately the right one, and, 5) that this result does *not* amount to succumbing to relativism. I conclude that the attempt to apply Bayesianism to the realism issue

enjoys a considerable amount of success, though not enough to resolve the dispute definitively.

Acknowledgements

My greatest debt goes to my supervisors, John Worrall and Stephan Hartmann. Working under John's supervision has been a unique intellectual experience. I have benefited immensely by John's deep and wide-ranging philosophical insight as well as his thorough and attentive criticism of my views and ideas. I am also grateful to him for invariably taking the time to read and comment meticulously on every single draft of each chapter, despite his overly busy schedule. I have no doubts that this thesis is now much more complete than it would have been without his help and guidance. Stephan has also been an invaluable source of support and guidance. His remarkable ability to come up with ideas for new research was crucial for stimulating my interest in the topic that I discuss in this thesis. His enthusiasm for philosophy has also been inspirational and helped me overcome various difficult periods in the process. I should also thank all my teachers in the Dept. of Philosophy, Logic and Scientific Method at the London School of Economics, and especially Colin Howson for enlightening me on Bayesianism on various occasions. And, of course, all my friends and fellow PhD students for making the Dept. of Philosophy a fun place to work, and especially Matteo Morganti, Mauro Rossi and Stan Larski for countless hours of discussion in philosophy and other more important things. Finally, special mention should be made to my family for all their unconditional support, financial and other, without which this thesis would be an impossible task.

Contents

Introduction

This thesis concerns scientific realism and Bayesian Confirmation Theory. *Scientific Realism* is a celebrated and also intuitively plausible doxastic attitude towards the theories of modern science. Realists claim that there are good reasons to think of these theories as (approximately) true descriptions of the physical world. Their belief is standardly supported by the No-Miracles argument, which suggests that the empirical success of modern science is so extraordinary that it would be a miracle if our well-confirmed, empirically successful theories failed to be at least approximately true descriptions of the world. *Bayesian Confirmation Theory* (or, simply, *Bayesianism*), on the other hand, is a quantitative analytical approach to the issue of theory-confirmation, which refers to the way that the evidence bears on our assessment of the merits of theoretical hypotheses. Bayesianism has received increasing attention recently and by now it is safe to assert that it at least counts among the dominant approaches to confirmation.

It might seem natural to suppose that Bayesianism offers a rigorous framework within which the question of realism can be investigated anew and that, as a result, advances in Bayesian Confirmation Theory are a welcome extension to the analytical resources we possess in debating the issue of realism. After all, the very basis of the No-Miracles argument is the extent to which our best theories are well-*confirmed* on the grounds of their empirical success. What else could be better suited for the analysis of the argument and the debate that stems from it than a theoretical approach whose very subject matter is the concept of 'confirmation'?

Quite paradoxically, this presumption is denied by many realists and by many Bayesians, each for their own reasons. Realists argue that Bayesians lack the conceptual resources to construe the debate properly, and this is why they think Bayesianism is insufficient as a framework within which the realism debate ought to be recast. While Bayesians utilize the conceptual resources offered by Bayesian Confirmation Theory only to reach the other extreme and suggest that both the realist argument and the ensuing debate are fundamentally flawed, marred by elementary logical errors and bound to lead us astray.

This situation struck me as really paradoxical, no less because both realism and Bayesianism have, independently of each other, received considerable attention and have been investigated with great care and subtlety. The foundational merits of Bayesianism are significant, resting as they do on seemingly compelling arguments. Intuitively, a rigorous analysis of partial belief ought to shed light on the ampliative inferences we perform to the (approximate) truth of our theories. Similar thoughts, however, apply to the realism debate. Many arguments and counter-arguments have been put forward during the last 30 years or so and many alternatives to realism have been entertained. The level of analysis has reached impressive depths, which it sounds just too far-fetched to think of as straightforwardly logically flawed and misguided.

In researching the topic, the belief that a Bayesian rendition of the realism debate will yield important intellectual benefits became stronger. Hence, I here defend the compatibility of the Bayesian mode of analysis with the aims and objectives of the epistemic dimension of the debate regarding scientific realism. In the first half of the thesis, I reconstruct and neutralize the various incompatibilist positions that have been put forth so far and only then (in the second half) present my positive proposal and analyze its contribution to the debate. In final analysis, I hold that employing Bayesian Confirmation Theory in the discussion of realism yields important benefits. As so often happens in analytical philosophy, Bayesianism primarily contributes clarity and rigor. Occasionally sloppy and obscure language is substituted by the more precise probabilistic idiom, thus greatly clarifying the concepts and arguments employed in defence of the various positions. Moreover, though primarily a neutral framework of analysis capable of accommodating all versions of the realist argument, Bayesianism also helps to bring out the realist claim in its sharpest form, thus offering further service to the realists.

At the same time, however, looking into the way that Bayesianism bears on the question of realism also reveals its limitations. I argue that these stem from the recognition of the fact that, for all its rigor and clarity, Bayesianism *cannot* offer a definitive resolution of the question of realism, at least any more definitive than the informal treatments already at hand allow for. I do not think that this qualification renders the application of Bayesian techniques to the question of realism redundant or pointless. On the contrary, it helps bring out in another way one of the salient features of all genuine philosophical problems, namely their open-ended, seemingly eternal character. Being a reflection of this inherent characteristic of all philosophical problems, the inability of Bayesianism to settle the question of realism once and for all becomes more of an underlining of an obvious limitation on philosophical analysis in general than a powerful argument against Bayesianism in particular.

So far I have only hinted at the contributions of Bayesianism to the treatment of the realism issue. Yet, I think there are gains to be found when looking at the converse direction too. More specifically, the application of the Bayesian apparatus on a problem as vexing as realism leads to a better appreciation of the character of Bayesianism and particularly its most widely accepted and, at the same time, most controversial version, i.e. Subjective Bayesianism. Chapter 5 contains, among other things, an analysis of the relationship between Subjective Bayesianism and 'scientific rationality', which I think is much more faithful to the fundamentals of the subjectivist position than the 'demonized' received view one frequently encounters in the literature. Overall then, there is much to be gained from applying the Bayesian apparatus to the question of realism, both with regards to our understanding of the realist argument as well as our appreciation of the character of Subjective Bayesianism.

In more detail, the structure and content of the thesis is as follows. In chapter 1 I introduce some of the fundamental ideas underlying Bayesian Confirmation Theory. The presentation is not exhaustive; it is rather tailored to what follows in subsequent chapters and the ensuing discussion. Chapter 1 is sub-divided into two. In the first part particular emphasis is given to betting behaviour and the Dutch-book arguments that aim to establish the connection between probabilities and rational degrees of belief, and also to the related question of the role of Bayesian Conditionalisation within Bayesian Confirmation Theory. In the second part, the focus shifts to the character of the prior probabilities and the various attempts to find rational constraints moving beyond the axioms of probability. Once more, the presentation is provisional and commits the analysis that follows to no particular interpretation; it merely sets the stage for the main discussion in later chapters.

The treatment of the dialectic between realism and Bayesianism begins in Chapter 2. There I discuss and disarm the first main challenge to the project of bringing these two 'isms' in contact, namely the claim that Bayesianism is ill-equipped to enlighten the realism debate because it lacks an appropriately strong notion of 'theory acceptance'. Realists, it is often argued, *accept* the theories of modern science on the basis of an argument for their approximate truth. Bayesians, however, cannot use the concept of 'theory acceptance' meaningfully, except in the extreme circumstance when a hypothesis is assigned probability 1 (and this is inapplicable to scientific theories). As against this view, I argue that Bayesianism already serves as a *reductive theory of 'theory acceptance'*, which allows us to replace the somewhat primitive idiom of 'theory acceptance' with the more advanced and precise language of probabilities. In addition to this, I show that the Bayesian conception of probabilities as degrees of belief best conveys the essentials of the epistemic dimension of the realism debate and, finally, go on to reconstruct the various alternative positions according to my proposed reductive analysis. The upshot of this chapter is that Bayesians possess sufficient conceptual resources to make sense of the realist thesis and contribute creatively to the debate regarding its validity.

If Bayesianism is an adequate analytical tool, however, might it not be the case that it helps us reveal major flaws in the informal statement of the most prominent realist argument? In Chapter 3 I explore the various forms of this accusation. I start by explicating what the No-Miracles argument for realism amounts to, noting the two markedly different ways in which it has been standardly understood; namely as a *plausibility consideration* and as an *Inference to the Best Explanation*. Although my sympathies lie with the former, at this stage I do not take sides. My primary interest is to investigate two important challenges to the argument, which purport to show that Bayesianism decisively defeats some or all versions of it. The first is Bas van Fraassen's (1989) claim that a Bayesian rendition of the No-Miracles argument *qua* Inference to the Best Explanation shows it to be *incoherent*; while the second is due to Colin Howson (2000) and suggests that the No-Miracles argument, in any of its forms, is a straightforward probabilistic fallacy. My criticism of van Fraassen's incoherence charge makes heavy use of the discussion regarding Bayesian Conditionalisation and its status as a principle of rationality found in Chapter 1. I argue that van Fraassen assigns the wrong role to Bayesian Conditionalisation and conclude that his argument does not go through even if we grant all his premises. Howson's argument, on the contrary, is valid. Rather than showing that the No-Miracles argument is fundamentally flawed, however, it merely points to the right way in which the argument should be probabilistically reconstructed. Hence, instead of offering an argument for incompatibilism, Howson ends up merely (but correctly) stressing the significance of prior probabilities in any Bayesian rendition of the No-Miracles argument.

A careful assessment of Howson's conclusions ultimately reveals how the No-Miracles argument can be faithfully captured in Bayesian terms. In Chapter 4 I take up the positive task of arguing for the compatibility between Bayesian Confirmation Theory and the realist claim. Taking my cue from Peter Lipton's (2004) reconciliatory approach between the Bayesians and the explanationists, I show how Bayesianism serves as a neutral framework within which *both* versions of the No-Miracles argument can be probabilistically reconstructed. I think that such a reconstruction conduces to clarity and reveals all the hidden presuppositions that the No-Miracles claim contains. At the same time, it brings to the surface from a different angle the notorious problem of the *subjectivity of the priors*. Perhaps unsurprisingly, *Subjective Bayesianism* is deemed the most adequate interpretational stance for the prior probabilities involved in our reconstruction. This means, however, that the strong claims to objectivity that the explanationist version of the No-Miracles argument makes are untenable. 'Inference to the Best Explanation' cannot provide the realist with the powerful argument he hopes to offer. Nonetheless, Subjective Bayesianism fares much better in conveying the essence of the No-Miracles claim construed as a plausibility argument. This I take to be direct evidence that, in the end, the No-Miracles consideration is just another plausibility claim. Consequently, Bayesianism also helps us discern the true character and reach of the realist argument, thus adjudicating a long-standing disagreement within the realist camp.

Even if Bayesianism can help in this way, however, can the same be done with respect to the dispute between realism and the various competing positions, like constructive empiricism or epistemic structural realism? Here the situation is more complicated. *Prima facie* it might seem that the allusion to what is plausible and what is not precludes any kind of objective treatment of this question. In the second part of chapter 4 I discuss the extent to which Bayesianism helps us decide between the various alternative options. Unsurprisingly, a Bayesian rendition of the situation will *not* settle the issue definitively. It helps illuminate, nonetheless, the interplay between the various considerations that can be adduced in favour of or against realism and the way each might turn the balance one way or the other. Prominent among the anti-realist arguments is the 'pessimistic induction'. It, too, has been accused of probabilistic invalidity by Peter Lewis (2001). Chapter 4 also includes a discussion of Lewis' claim and concludes that his assessment fails to do justice to the import of the argument. Properly understood, a probabilistic reading of the 'pessimistic induction' brings out with remarkable precision the ways that the historical record *can* influence our belief or disbelief in realism. Ultimately, from a Bayesian viewpoint no definitive winner will be found. What will be found is an *explanation* of why this is so. Chapter 4 concludes by explaining why the inability of our Bayesian reconstruction to settle the debate is merely a reflection of the nature of the philosophical problem of realism *per se* rather than a deficiency of the Bayesian framework of analysis.

Chapter 5 focuses on the general epistemological presuppositions behind the two versions of the realist argument. In very general terms, its main aim is to substantiate what was implicitly assumed in chapter 4, i.e. a) that Subjective Bayesianism *is* in a position to capture the normative force of the No-Miracles argument consistently with its own principles; and b) that the normative force of the plausibility version of the argument, sanctioned by our probabilistic reconstruction, is in fact *significant* despite its appeal to intuition. The need for the defense of these two claims arises out of the following, commonly held, views: I) that a Subjective Bayesian reconstruction of the No-Miracles argument by definition undermines its strength, due to the (alleged) relationship between Subjective Bayesianism and relativism; and II) that plausibility considerations alluding to intuition, irrespective of whether they take a probabilistic expression, are utterly unconvincing, since 'intuitiveness' is a subjective notion.

I tackle (I) and (II) in order: first, I deny that a sensible construal of the nature of the Subjective Bayesian interpretation of probabilities brings with it any consequences with respect to the issue of relativism. The problem of relativism is independent of one's stance towards the probability calculus and has to be countered on general epistemological grounds. Appreciating this point, however, means that one has to delve into the basic epistemological presuppositions of the No-Miracles claim, in its various forms, in order to locate the way in which it aspires to defeat the relativist challenge. My second task, then, is to examine various alternative ways this has been attempted and evaluate their merits. By defending a foundationalist answer to the problem of relativism based on the notion of 'inductive intuition', I directly deny that plausibility considerations are powerless. Quite to the contrary, on mature reflection they are our only hope in the battle against the relativist.

The last important issue which is brought out by the investigation of the epistemological fundamentals of the realist argument is the way in which the various proposed answers to the relativist challenge have an impact on our ways of interpreting partial belief and probability. Even if it is the case, as I argue it is, that Subjective Bayesianism is neutral with respect to the problem of relativism, an adequate answer to this problem is expected to influence our ways of interpreting prior probabilities. Consequently, my last concern is to show how and why Subjective Bayesianism best conveys the essentials of the 'intuitive' foundationalist answer I embrace.

In sum, this thesis contains an argument that the Bayesian and the realist *should* be friends. They should be friends because they both possess adequate conceptual means to tackle effectively the question of realism; because bringing

Bayesianism to bear on the realist's worries clarifies the logical structure of the realist argument and its potential rejoinders; because Bayesianism helps settling a long-standing debate within the realist camp as to the proper form of the No-Miracles argument; because the inability of a Bayesian reconstruction to settle the debate definitively reveals the distinctively philosophical nature of the problem of realism; and, finally, because the attention drawn towards the fundamental epistemological assumptions employed in the debate further clarifies not only the epistemological character of the realist case but also the nature of the most controversial variant of Bayesian Confirmation Theory, i.e. Subjective Bayesianism.

To my mind, all the above are compelling reasons why one *should want* the realism debate to be (also) cast in Bayesian terms. This conclusion, however, is *not* to be interpreted as yet another attempt to reduce philosophical discourse to a mere exercise in formal reasoning. All the central arguments contained in this essay are philosophical rather than technical. For better or for worse, Bayesianism (as well as any other formal tool of analysis) is intended only to facilitate, not substitute for philosophical thinking.

Chapter 1

Bayesian Confirmation Theory

In spite of the widespread usage of the term 'Bayesianism' in contemporary philosophy of science, it is best thought of as a cover-term. There are elements that all Bayesians more or less agree on but the term 'Bayesianism' conceals several different interpretations of the probability calculus. In this chapter I shall first highlight the main points of agreement between Bayesians of all persuasions and then analyze the essentials of the diverging interpretations.

1.1 Elements of Bayesianism

1.1.1 Probabilities as Degrees of Belief

The first major common thesis that all Bayesians share is an interpretation of the probability calculus as representing *degrees of belief* rather than objective features of the world. The point of view we are concerned with here is *ontological* in nature and reflects an attitude towards the nature of the calculus. Interpretations that take probabilities to be degrees of belief can be classified under the label of *epistemic interpretations*, in contradistinction to *physical* ones that take probabilities to be features of the physical world. Hence, Bayesians accept as legitimate an epistemic account of probabilities. They need not, however, be committed to the stronger view that an epistemic account is the *only* legitimate account of the calculus. Pluralist approaches to probabilities have been adopted from as early as Poincaré's *Science* and *Hypothesis* (1902, ch. XI) and Ramsey's classic (1926) paper "Truth and Probability" and, as we shall see shortly, are still quite popular[1]. Hence, if one is a Bayesian, one adopts *at least* an epistemic interpretation of the probability calculus.

1.1.2 Justification via Betting Behaviour

A second thesis that most Bayesians adopt involves reference to betting behaviour in order to justify the connection between probabilities and degrees of belief. The paradigm of such justification takes the form of the celebrated Dutch-book argument, originally presented by Ramsey (1926) and De Finetti (1937). This argument shows that if one's degrees of belief fail to comply with the formal axioms of the probability calculus, then one must be susceptible to accepting a bet which guarantees a loss whatever happens (Dutch-book); and, conversely, that if one is susceptible to a Dutch-book, then one's degrees of belief fail to comply with the axioms of the probability calculus. Failure of your degrees of belief to instantiate the axioms of the calculus is equivalent to employing degrees of belief that are *incoherent*, in the sense that there is a system of bets which you regard as

[1] Over the years pluralist approaches to probability have been adopted by *inter alia* Carnap (1950), Howson and Urbach (1993) and Gillies (2000).

'fair' but on which you are bound to lose come what may (that is, whatever way the world turns out to be).

Despite its apparent simplicity, however, a lot of ink has been spilled over the exact character of the Dutch-book argument and the kind of justification it offers for equating degrees of belief with probabilities. This has resulted in much confusion.

Reference to betting behaviour in justifying the connection between degrees of belief and probabilities goes back to Ramsey's intuition that "the old-established way of measuring a person's belief is to propose a bet, and see what are the lowest odds which he will accept" (Ramsey 1926, 172). The betting situation is set up as follows: person A is the bookmaker and person B the agent, whose degrees of belief on the truth-value of a proposition E we want to elicit. B agrees to state a number q, called his betting quotient on E, and only after this is done the bookmaker chooses the stake S, which can be either positive or negative. The agent is to pay the bookmaker qS in exchange for S, if E comes out true and nothing if it comes out false. Hence, in case E comes out true the agent receives $S\text{-}qS = S(1\text{-}q)$ and he loses qS in case E turns out to be false. It is assumed that, under the above conditions, q represents B's degree of belief in the truth of E. It can be proved that B's degrees of belief are coherent in the sense of avoiding a Dutch-book if and only if they satisfy the formal axioms of probability. This is the *Ramsey-De Finetti theorem.*

The Ramsey-De Finetti theorem applies in a straightforward way relative to the following axioms of the probability calculus:

1. $0 \leq \Pr(E) \leq 1$, and $\Pr(T) = 1$, where T is a necessary truth.

2. If two propositions E_1 and E_2 are inconsistent, $\Pr(E_1 \vee E_2) = \Pr(E_1) + Pr(E_2)$. *Finite additivity* then follows, since for any finite number *n* of mutually inconsistent propositions $E_1 \ldots E_n$, $\Pr(E_1 \vee \ldots \vee E_n) = \sum_{i=1}^{n} \Pr(E_i)$.
3. For any two propositions *E* and *F*, assuming $\Pr(F) \neq 0$, $\Pr(E/F) = \Pr(E \wedge F)/\Pr(F)$.

Consider, for example, the simple case of axiom 1. Assume that one fails to assign probability 1 to a tautology, setting instead q = Pr(*T*) < 1. All the bookie has to do then is set a negative stake. Then, the agent is guaranteed to lose $|S|(1\text{-}q)$, since a tautology is always true. The rest of the axioms are justified in a similar manner[2].

Disagreement has been quite widespread regarding *countable additivity*, however, in which the domain of the propositions is (countably) infinite. More formally, the following principle had for long been deemed either illegitimate or at least problematic for Bayesians:

2'. For a countably infinite number of inconsistent propositions $E_1 \ldots E_n \ldots$,

$$\Pr(E_1 \vee \ldots \vee E_n \vee \ldots) = \sum_{i=1}^{\infty} \Pr(E_i).$$

The reason is that countable additivity does not allow for a uniform distribution over the countable set of propositions. To see this, consider the hypothetical case that we do assign a uniform distribution over the uncountable domain. Then, by axiom (1), $\Pr(E_i) \geq 0$. If it is greater than zero, then the sum will be infinite, violating the other half of axiom (1), which demands that a probability assignment

[2] See Gillies (2000, 59-64) for a complete demonstration of the Ramsey-De Finetti theorem relative to all axioms.

cannot be greater then one. If it is zero, then $\sum_{i=1}^{\infty} \Pr(E_i) = 0$, violating axiom (1) in that a necessary truth has probability 1. It is assumed here that our domain is denumerably infinite, so that denumerably infinitely many inconsistent propositions exhaust this domain and, hence, $\sum_{i=1}^{\infty} \Pr(E_i) = 1$.

De Finetti, very much in the spirit of his subjectivism, famously rejected countable additivity on the grounds that there is no good *a priori* reason why one should not bet according to a uniform distribution. On the other hand such a decision has many undesirable consequences for the actual practice of mathematics, since countable additivity has for long been a very useful probabilistic assumption. Hence, until recently, many authors had either used it or even stipulated it as a strengthening of axiom (2) on the basis of its instrumental value despite the absence of a foundational justification[3]. Recently, however, Williamson (1999) has given a Dutch-book argument for countable additivity to the effect that whoever violates the principle is irrational precisely in the sense that a bet ensuring loss can be arranged against him. Williamson's argument invokes just one additional constraint - that only a finite amount of money can change hands after the outcome of the bet is determined. Hence, betting according to a uniform distribution is in fact excluded by the same considerations of coherence that necessitate adherence to the other axioms of probability.

These important results have been taken to provide a complete justification for equating rational degrees of belief with probabilities. Furthermore, the appeal to coherence led Ramsey (and his followers) to assert that "the theory of probability is in fact a generalization of formal [i.e. deductive] logic" (Ramsey 1926, 82).

[3] See Williamson (1999) and references therein.

According to this view, what probability theory tells us is what degrees of belief we *should* employ about certain propositions *given* our simultaneous degrees of belief on other propositions. Such remarks, though, sound at least over-ambitious if one of the most common objections to Dutch-book justifications is correct - namely that such arguments offer only a *pragmatic* justification of the calculus (cf. Kennedy and Chihara 1979; Joyce 1998, 584-586).

The motivation for this complaint is simply the very down-to-earth betting set-up as well as the explicit invocation of utilities or monetary units in the statement of the Dutch-book argument. These have been taken to suggest that the success of the argument offers only pragmatic support to the conclusion without any apparent relevance for epistemology. James Joyce has stated this complaint as follows:

> "A more significant problem has to do with the *pragmatic* character of the Dutch-book argument. There is a distinction to be drawn between *prudential* reasons for believing, which have to do with the ways in which holding certain opinions can affect one's happiness, and epistemic reasons for believing, which concern the accuracy of the opinions as representations of the world's state. Since the Dutch-book argument provides only a prudential rationale for conforming one's partial beliefs to the laws of probability, it is an open question whether it holds any interest for epistemology" (1998, 584-585)[4].

The main problem the Dutch-book argument faces, it is claimed, is that it focuses too much on the *prudential* reasons for conforming one's degrees of belief to the laws of probability. Those reasons, however, relate to one's happiness rather than to

[4] Emphasis in the original.

any epistemically significant concept, and this why their relevance or usefulness to epistemology is questioned. What is worse, the *descriptive inaccuracy* of the Dutch-book assumptions might also lead one to doubt that even the pragmatic reading of the argument is of any value at all. It surely is a truism that factual evidence about the behaviour of individuals reveals that the Dutch-book assumptions are high-order idealizations rather than direct realistic descriptions. But then, one might argue, the Dutch-book argument offers not only a pragmatic justification of the axioms but also a *poor* one for that matter, since its results can not be regarded as representative even of what is involved in actual cases of prudence.

This line of reasoning has given rise to many alternative justification procedures, which are purported to avoid the perceived shortcomings of the Dutch-book argument. Some of them define (appropriately measured) 'epistemically significant concepts', such as *accuracy* (cf. Rosenkrantz 1981; Joyce 1998), as the goal of rational agents in their inquiries; and argue that a violation of the axioms of probability undermines that goal in the sense that a system of beliefs, which satisfies the axioms, is invariably more accurate. Others stipulate plausible desiderata and then show that any real-valued function, which satisfies them, can be transformed into a probability function (cf. Cox 1961, Good 1950 and Lucas 1970). Still others justify the axioms on the grounds that violating them precludes the possibility of one's degrees of beliefs being accurate estimates of frequencies, since frequencies themselves do obey the axioms of probability (cf. van Fraassen 1983). Nonetheless, most Bayesians do subscribe to the Dutch-book argument and, I think, with good reason.

This is that the charge against the Dutch-book argument that it offers merely a pragmatic justification of the axioms is unfair. Joyce, to be sure, has drawn our attention to a very important distinction, namely that between offering *pragmatic* versus offering *epistemic* reasons for belief. It may be the case that acting on the basis of certain beliefs can be shown to serve the satisfaction of our behavioral, psychological or financial aims in particular situations. If so, then we certainly have pragmatic reasons to co-ordinate our actions with these beliefs whenever we pursue such aims in those situations. This possibility, however, need not be relevant also to epistemic justification. This is because achieving our behavioral aims may very well depend on factors other than the strictly epistemic merits of our beliefs. It appears, then, that epistemic justification can be obtained only through arguments pertaining to show that our beliefs allow us to satisfy our strictly *epistemic* aims (e.g. truth, consistency etc.), irrespective of their usefulness as a guide for action in particular practical situations. Joyce correctly points out in this vein that "there does seem to be a clear difference between appraising a system of beliefs in terms of the behavior it generates or in terms of its agreement with the facts" (1998, 585). But, surely, if there is such a difference, then it is not enough merely to cite practical advantages or disadvantages (financial or other) in order to reach normative epistemic conclusions (cf. Christensen 1996, 451-452). What is required is an argument focusing on the exclusively epistemic merits of our beliefs.

Affirming Joyce's distinction, however, does not imply that the Dutch-book argument offers only a pragmatic justification for the axioms of probability. In fact, as Howson notes, "to see [the] axioms as mere assurances of financial safety is to miss their real significance" (2000, 126). Equivalently, to see the Dutch-book argument as a pragmatic consideration aiming at prudent housekeeping is also misguided. The Dutch-book argument is not a pragmatic consideration referring to real life. Rather it is an idealization, an abstract model, which, though inspired by real-life situations such as betting, serves a purely epistemic aim. This is to demonstrate how violating the probability axioms results in one's degrees of belief being *inconsistent* in precisely Ramsey's sense: they violate genuine laws of logic, the logic of partial belief.

Christensen (1991, 238-239) sums up the line of reasoning underlying this position as follows:

1. Consistency in one's degrees of belief is a cognitive desideratum.
2. Consistency in a person's degrees of belief implies that no Dutch-books can be made against that person.
3. However, Dutch-book arguments show that violating the axioms of probability gives rise to such books.
4. Hence, the axioms of probability are criteria of consistency.

Premise (1) is eminently plausible for anyone who takes deductive logic seriously, regardless of whether one purports to understand the probability axioms in a like manner, i.e. as constraints of a logical nature, or not. Premise (2) should also be uncontroversial. Given a few innocuous assumptions about self-interested behaviour, it is quite sensible to assume that no one who respects the minimal constraints of logic would enter voluntarily an arrangement entailing a loss come what may.

Very much like deductive inconsistency, it is overwhelmingly probable, although not guaranteed *in actual life* due to the imperfection of human logical faculties, that probabilistic inconsistency will also have practical consequences. This is not the central issue though. The central issue has to do with whether a normative claim can be established regarding the constraints which the axioms of probability impose. The Dutch-book argument, disentangled from considerations referring to every-day betting scenarios, does succeed in revealing the normative force of these constraints.

Talking in this way, however, also shows that, despite having normative force, Dutch-book arguments are *not* fundamental in the attempt to characterize precisely the nature of the constraints imposed by the axioms of probability. In Christensen's words, "potential vulnerability to [the] particular kind of monetary loss [Dutch-book arguments revolve around] serves [only] as a vivid *symptom* of a real problem" (1991, 238)[5]. The task of providing a more precise and rigorous account of what the 'real problem' of inconsistency consists in has been recently taken up by Howson (2000). Using deductive logic as his model, Howson lays out 3 individually necessary and collectively sufficient conditions for a discipline to qualify as 'logic'. These are: (a) that the discipline involves statements and relations between them, (b) that it legitimizes some form of non-domain-specific-reasoning, and (c) that it is about consistency, in the sense of having a soundness and completeness theorem (2000, 127).

In the case of deductive logic, a soundness theorem states that every sentence that can be derived using the rules of proof is a deductive consequence of the axioms (so that in particular where these are the axioms of pure logic, every derivable sentence is a logical truth), while a completeness theorem states the converse (that is, that every deductive consequence of the axioms is derivable using the system's rules of proof). Since we are now interested in *probabilistic*, rather than deductive, consistency, a soundness theorem will state that any consistent

[5] Emphasis in the original.

assignment of fair betting quotients will satisfy the axioms of probability, while a completeness theorem, like before, the converse.

Howson goes on to show that probability theory satisfies all three conditions and, hence, does qualify as logic, the logic of partial belief (ibid. 127-132). In effect, the constraints of the axioms of probability (including countable additivity) can plausibly be seen as *logical constraints* directly analogous to those of deductive logic, exactly as Ramsey envisaged in the beginning of the 20th century[6].

One can sum up the results of the discussion so far as follows: the Dutch-book argument, despite appearances, does have normative force in that it reveals inconsistencies between our various degrees of belief, if these are not probabilities. The nature of these inconsistencies can be more precisely characterised using the resources of formal logic. This, however, need have no implications regarding the kind of justification Dutch-book arguments offer. It follows, then, that worries arising from its alleged pragmatic dimension vanish, since there is no really pragmatic dimension, or any such element involved, in this particular use of the Dutch-book argument. At the same time, the above picture helps us explain why it is no surprise that alternative justifications successfully demonstrate the superiority of a system of beliefs respecting the axioms compared to ones that don't. One can hardly expect an inconsistent system in the logical sense to do better than a consistent one relative to whatever systematic epistemic standards[7]. Hence, while

[6] De Finetti seems to have a similar interpretation to Ramsey's in mind. His expositions, however, are less clear because of his ultra-empiricism and operationalist spirit. For these aspects in De Finetti's work, see Galavotti (1989).

[7] Non Dutch-book justifications yield results which hold *invariably* for all situations and, hence, are not vulnerable to the objection that an inconsistent agent might do very well in a one-off situation, since a consistent one would do even better. To be sure, these justifications are not without

Earman's conclusion that "collectively [all the above methods of justification] provide powerful persuasion for conforming degrees of belief to the probability calculus" (1992, 46) seems reasonable, one may arguably claim that the betting scenario suffices on its own to correlate degrees of belief with probabilities, especially in the face of its overwhelming simplicity when compared to other methods of justification.

1.1.3 The Principle of Conditionalisation

A third thesis most Bayesians share refers to the way beliefs are updated in the face of evidence. The model used to describe the process is called *Bayesian Conditionalisation*, which takes two forms, a strict form and a more sophisticated one due to Richard Jeffrey. *Strict Conditionalisation* (SC) dictates that if one's probability in evidence *E* is 1 and if *E* is the strongest such proposition, then the new probability we assign to some one hypothesis *H* after we have learned evidence *E* equals the old probability of *H* conditional on the evidence in question. Formally: $\Pr_{new}(H) = \Pr_{old}(H/E)$.

SC assumes that the evidence is known with certainty *presently*[8]. It could be the case, however, that even in present time we are only partially confident regarding evidence *E*, perhaps because some experiment or observation has merely altered the degree of our partial belief in *E* rather than made it certain. Despite the lack of certainty about *E*, changes in its probability are expected to have some

---

problems. Most of them, though, revolve around the plausibility of their underlying assumptions rather than the universality of their results (cf. Earman 1992, 44-46).

[8] If $\Pr_{old}(E) = 1$, then no belief-change would occur at all, since in that case $\Pr_{new}(H) = \Pr_{old}(H/E) = \Pr_{old}(H)$.

effect on the probabilities of some of our hypotheses. Under those circumstances, belief-change accords with *Jeffrey Conditionalisation* (JC), originally proposed by Jeffrey (1965): $\Pr_{new}(H) = \sum_i \Pr_{old}(H/E_i) \times \Pr_{new}(E_i)$, for all members of a partition $E_i, \ldots, E_n$. $\Pr_{new}(E_i)$ reflects our *present* (new) uncertainty over the evidence, which is different from $\Pr_{old}(E_i)$ and arises through an exogenous, i.e. non-conditionalisation, change of belief stemming from, in fact, observation or experiment (at least in standard cases). Clearly, SC is only a special case of JC for $\Pr_{new}(E_i) = 1$, for some one *i*.

Despite this, virtually all recent applications of Bayesianism in philosophy of science make use just of SC on the basis of simplicity of presentation. Mathematically, $\Pr_{old}(H/E)$ is standardly computed using Bayes' Theorem, which in its simplest form says that:

$$\Pr(H/E) = \frac{\Pr(H) \times \Pr(E/H)}{\Pr(E)}, \qquad (1)$$

where Pr(*H/E*) is the posterior probability of the hypothesis, Pr(*H*) the prior probability of the hypothesis, Pr(*E/H*) the likelihood of the evidence and Pr(*E*) the prior probability (or expectedness) of the evidence.

In the literature there have been several attempts to offer Dutch-book arguments for both SC and JC and, hence, render them an indispensable part of the Bayesian apparatus, on an equal footing with the standard probability axioms. The underlying idea is to show that if one violates the principle of conditionalisation in either of its forms, a set of *sequential* bets can be arranged, which, although fair for our agent when taken at any given time-instant, collectively entail an inevitable loss over time. Some such result or other is often taken to establish the thesis that SC and JC enjoy the same status as the formal axioms of probability[9]. In what follows I discuss only the case of SC.

The betting set-up for a diachronic Dutch-book argument for SC is precisely the same as with the synchronic case, with the agent stating his fair betting quotients (presumably as and when these arise at various temporal stages) and then the bookmaker arranging the stakes[10]. Take the simplest case, i.e. a two-stage betting situation. Suppose that one's new degree of belief in a hypothesis H is $\mathrm{Pr}_{new}(H) = 0.5$, while by conditionalisation on his old beliefs he would have reached the conclusion that it was $\mathrm{Pr}_{old}(H/E) = 0.7$. Suppose, also, that he gives the evidence E $\mathrm{Pr}_{old}(E) = 0.6$. The following betting strategy is open to the bookmaker: (a) he offers now (i.e. at 'old' time) a *conditional* bet on H given E, whereupon the bet pays, say, £10 if both E and H are true at the price of £7, and the bet is off in case E is false, b) he also offers a bet on E alone, paying £2 at the price of £1.20. Both these bets, taken individually, are fair for the agent, who accepts them.

If E is false, then the bookmaker wins a net £1.20 pounds from bet (b), since bet (a) is off. In case E turns out to be true, the bookmaker offers to buy a third bet from the agent, this time based on his new degree of belief on H: (c) the bookmaker offers to buy a bet paying £10 in case H is true at the price of £5. Once again, this bet is fair taken by itself. Since it applies *only if* E turns out to be true, the net gain

[9] Van Fraassen, being a liberal Bayesian, would deny this implication for he holds that Bayesian Conditionalisation imposes itself as a rationality requirement *only if* we decide to update belief following a rule. The latter condition, though, is not necessarily true. See van Fraassen (1989) p. 169 and Douven (1999) p. S425.

[10] The discussion of the diachronic Dutch book argument here is a more concrete version of the one found in Howson and Urbach (1993, 100-105). The original exposition can be found in Teller (1973, 1976).

for the bookmaker in this case will be: 1) in case H is true, (£7 - £10) + (£1.20 – £2) + (£10 - £5) = -£3 - £0.8 + £5 = £1.20.

2) in case H is false, £7 + (£1.20 – £2) - £5 = £1.20.

Hence, in any case, regardless of the truth values of E and H the bettor is assured of losing a fixed amount of money or, in other words, a Dutch-book[11].

What does this mean for conditionalisation though? The proponents of conditionalisation have taken this argument to show that the principle of conditionalisation is a requirement of rationality. This, however, is not at all obvious. One worry relates to the fact that here we are talking about a *combination* of bets, not just a single one. For the argument to go through one needs the extra assumption that "the value of the sequence [of bets] to him [i.e. the agent] is the sum of the separate values he sets" (Schick 1986, 117), sometimes referred to as the 'package principle', which, it is argued, need not be true[12]. This objection need not be that troubling, however. All we have to do is recall our interpretation of the betting scenarios as merely abstract models, which help us determine the conditions that give rise to incoherent states of belief, and thus help us end up with conclusions bearing normative status. To the extent that this is our understanding of the betting situation, monetary value-related worries immediately disappear as missing the point of the argument.

[11] A similar argument has also been presented in favour of JC. See Skyrms (1987) and Howson and Urbach (1993, 109-110).

[12] In fact this objection is also intended against the justification of the synchronic axiom of additivity, stating that $\Pr(A \vee B) = \Pr(A) + \Pr(B)$, if $\models \neg(A \wedge B)$, since it also rests on a 'package' of bets, albeit simultaneous and not diachronic (Schick 1986, 112-116).

A more fundamental problem for the diachronic Dutch-book argument is whether it is of the appropriate kind to deliver its conclusion. For the argument to show inconsistency in the logical sense, it has to be assumed that it is a Dutch-book argument of the same kind as the one pertaining to the synchronic axioms of probability. Appearances to the contrary, not only is this assumption not self-evident but it is clearly false. To see this, we have to look closer at the general structure of the argument.

The dynamic Dutch-book argument shows that there is an inconsistency between a) an initial probability assignment Pr(H), b) the principle of conditionalisation as a means for updating belief, and c) a new assignment not in accordance with this principle. The proponents of conditionalisation elect to put the blame either on (a) or (c). Hence, they argue, either we have been incoherent by violating conditionalisation and reaching (c) on the supposition that our initial assignment (a) is unproblematic; or, in case, we know that (c) will be our new degree of belief at some future time, we have been incoherent due to our initial assignment (a). The second case is summarised in van Fraassen's (1984) Principle of Reflection:

$$\mathrm{Pr}_t\,(H/\mathrm{Pr}_{t+1}(H) = \mathrm{r}) = \mathrm{r}.$$

In either case, conditionalisation remains sacrosanct.

There are convincing counter-examples, though, to this line of argument, the most entertaining of which is due to Christensen (1991). Suppose someone swallows a pill of LSQ, a psychedelic drug which makes you believe strongly within one hour of taking it that you can fly (F). Conditionalising on the information that you will be so disposed to believe, yields $\mathrm{Pr}_{t+1}(F) = \mathrm{Pr}_t(F/\mathrm{Pr}_{t+1}(F) = 0.99) = 0$. At the same time, $\mathrm{Pr}_{t+1}(F/\mathrm{Pr}_{t+1}(F) = 0.99) = 0.99$ as a matter of tautology. Hence, the clash between (b) and (c) cannot be resolved by blaming the assignment in (c). What about (a), i.e. that $\Pr(F) = 0$? It seems obvious to me that Christensen is perfectly right when concluding that, if this person is asked about his degree of belief *now* that he can fly, "the answer mandated by Reflection (.99, of course!) is ridiculous" (ibid. 234). This is a clear case, however, in which both ways that the conditionaliser uses to resolve conflict fail, which in turn means that doubt must be cast on the principle of conditionalisation itself as a principle of rationality.

In order to doubt this, though, one must find a way to come around the obstacle presented by the fact that what "justifies" conditionalisation is a *Dutch-book argument*. To avoid inconsistency, then, one has to show that this Dutch-book argument is not of the same kind as the ones used to justify the synchronic axioms. This is precisely the answer sketched by Christensen (1991) and elaborated in more detail by Howson and Urbach (1993) and Howson (2000). It revolves around the idea that probability theory is a kind of logic dealing with the consistency of beliefs. Taking into account that consistency, as in deductive logic, is first and foremost a '*synchronic*' notion, Howson and Urbach (1993, 102-103) argue that, although the series of bets in the diachronic version of the argument *does* result in inevitable loss, this does *not* suffice to demonstrate inconsistency, for it violates the strictly synchronic character of the latter. In the same sense that I can consistently believe *p* today and *not-p* tomorrow, I can also violate today's assignment via conditionalisation tomorrow without being inconsistent.

To my mind this is clearly the correct answer to the problem of conditionalisation: the rule cannot be considered a universally true one without qualifications. Furthermore, it gives new insight to the Dutch-book argument for the standard case: for it to have any normative bearing regarding the restrictions on one's degrees of belief, it has to be *synchronic*, i.e. all the bets must be bought *at one and the same time*. Mere assurance of net loss is not enough. In order to differentiate between these cases, arguments like the diachronic Dutch-book are usually called 'Dutch-strategies' (a term due to van Fraassen (1984)) rather than Dutch-books proper.

The above arguments do not intend, of course, to establish that Bayesian conditionalisation, in any of its two forms, is generally invalid; they clearly suggest, however, that it is valid *only conditionally*. The condition under which conditionalisation is valid, then, is that upon learning evidence E and $\Pr_{new}(E) = 1$, one's conditional probabilities relative to E remain unchanged. In other words, if, upon learning E with certainty, $\Pr_{new}(H/E) = \Pr_{old}(H/E)$, then conditionalisation becomes a valid rule of inference. This is because if $\Pr_{new}(E) = 1$, then $\Pr_{new}(H/E) = \Pr_{new}(H)$. Hence, if $\Pr_{new}(H/E) = \Pr_{old}(H/E)$, conditionalisation also holds, i.e. $\Pr_{new}(H) = \Pr_{old}(H/E)$. Surely though, as Howson and Urbach (1993, 103) remark, imposing this restriction makes the diachronic Dutch-strategy entirely unnecessary, since conditionalisation is now by definition the correct updating rule. Consequently, there seems to be no place for the diachronic Dutch-strategy in the realm of Dutch-books proper[13].

Where does all this leave the discussion regarding conditionalisation? Despite their abstract and near-Cartesian[14] character, one consequence of these

[13] *Mutatis mutandis* the same conclusion can be drawn for JC (cf. Howson and Urbach 1993, 109-110).

[14] I say *near*-Cartesian because it seems obvious that the drug counter-example is not as remote from common sense a circumstance as the original Cartesian demon-hypothesis is.

arguments is that conditionalisation is *not* a 'logical' principle in the way that the rest of the axioms of probability arguably are, since there exist cases, as we saw, where violating it is the reasonable thing to do. That these counter-examples make use of unlikely situations does not, I think, impair their force, since they are employed only for the purpose of demonstrating the failure of Dutch-strategies to count as Dutch-books proper, rather than showing that such counter-examples are the norm in actual life.

Clearly, the same holds for the condition that one holds his conditional probabilities relative to the evidence unchanged. If that condition were analytic, then so would conditionalisation be. Howson and Urbach remarked at one point that "this is a condition [they] can imagine satisfied by [the] ideal scientific reasoner" (1993, 113) Bayesianism refers to. I don't think, however, that this is tantamount to admitting that conditionalisation has got any kind of special status similar to the one of the probability axioms. It sounds more like the *pragmatic* point that this way of modelling confirmation happens, as a matter of fact and not logic, to convey our intuitions about how belief changes over time upon learning new evidence. This means, though, that there is a residual tension between Howson and Urbach's allusion to an 'ideal reasoner' and their overall interpretation of the probability calculus. As we have already mentioned (and shall see in more detail shortly), Howson and Urbach adopt a logical variety of Subjective Bayesianism, according to which the only justifiable constraints probability theory imposes are of a logical nature. Now, if alluding to an 'ideal agent' fails to establish a logical fact, how can Howson and Urbach still maintain conditionalisation as a fundamental principle of Bayesian reasoning?

To the best of my knowledge, the proper answer to this problem has been provided only recently by Colin Howson (unpublished), in a way which greatly clarifies (and corrects) the account encountered in Howson and Urbach (1993). Howson now abandons all reference to an 'ideal reasoner', opting instead to drop the principle of conditionalisation altogether: Bayesian Confirmation Theory does *not* need conditionalisation because confirmation has nothing to do with *temporal* updating of belief.

Recall the statement of conditionalisation: $\Pr_{new}(H) = \Pr_{old}(H/E)$. In words, my 'new' probability, say in time $t + 1$, after I have learned the evidence E with certainty equals my 'old' conditional probability, say at time t, before E was known with certainty. Howson's proposal denies that confirmation has anything to do with time. It should not enter our considerations, in other words, whether in actual life I can only visualise having new degrees of belief at a later time relative to my prior beliefs. On the contrary, confirmation has to do with the objective relations between theory and evidence, *regardless of when* the evidence becomes known and the exact time that *we* hold the relevant beliefs. Put differently, confirmation only deals with how the evidence *forces* us to change our prior beliefs, were we to take it into account. So, 'prior probability' is synonymous with 'probability irrespective of the evidence at time t' and 'posterior probability' with 'probability in the light of the evidence at time t also'. On this account, though, the usual synchronic conditional probability relations will do for the purposes of confirmation theory and no need for conditionalisation emerges. In Howson's own words, "where conditionalisation is valid, *its validity is guaranteed by the axioms of probability themselves*" (unpublished, 17)[15].

[15] Emphasis in the original.

Howson's decision to abandon the principle of conditionalisation, then, essentially makes explicit all the consequences of the position adopted in Howson and Urbach (1993). The principle of conditionalisation is not a logical principle. Consequently, it is of no avail to anyone who admits only of logical constraints on probabilistic assignments. Of course, if one is more lax with respect to the nature of the constraints deemed permissible, one might still accept the principle of conditionalisation on other than logical grounds. For present purposes, however, it is crucial to keep in mind the *non*-logical character of conditionalisation. As we shall see in subsequent chapters, the mistaken view that conditionalisation is a logical principle has been a central tenet in recent Bayesian reconstructions of the debate regarding scientific realism.

1.1.4 A Quantitative Account of Confirmation

Finally, all Bayesians agree that confirmation should be construed as quantitative rather than qualitative in character and that the relation of theories to the evidence ought to be represented by means of the probability calculus. There has been ample disagreement in the literature, however, regarding the exact form this representation should take. So-called *strict Bayesians* (cf. Maher 1986, Howson and Urbach 1993) assert that point-valued degrees of belief are a legitimate way to represent one's state of knowledge, while *weak Bayesians* (cf. Good 1962, Levi 1974, Walley 1991) hold that interval-valued degrees of belief are to be preferred on both descriptive and normative grounds.

In the present context I shall assume, although not discuss extensively, that point-valued degrees of belief are unproblematic, or at any rate, not more problematic than their interval-valued counterparts, and, hence, that strict

Bayesianism goes untroubled. This is not merely an arbitrary assumption. Both Maher (1986, 452) and Howson and Urbach (1993, 88) rightly remark that if strict Bayesianism fails, then the intervals of the weak Bayesians lack determinate bounds, transforming weak Bayesianism into an equally unrealistic theory. Surely, rational agents assume that there *are* determinate bounds in their interval-valued degrees of belief as much as they fail to single out precisely one particular probability function.

On the normative side, Maher (1986, 453-455) offers convincing reasons why weak Bayesianism cannot be taken as normatively superior to the strict variant. Briefly put, his argument takes the form of a dilemma: either interval-valued degrees of belief hamper the whole of Bayesian decision-theory by failing to account for the notion of expected utility, or, by modifying weak Bayesianism in an attempt to rectify this problem, one falls back at best to strict Bayesianism and at worst to principles of dubious standing, like the *maximin* principle. It follows, then, that one does not incur any significant loss by modelling scientific confirmation in terms of point-valued, instead of interval-valued, degrees of belief.

This is not to say, of course, that, after all, people actually entertain point-valued degrees of belief in real-life situations. Still, idealising to this effect has proved an immensely useful practice, due to the accuracy and clarity of the results produced. Arguing in favour of this practice, Howson and Urbach draw a parallel with the case of physical magnitudes, "where people are quite happy to invoke real-number values, even though, strictly speaking, lengths, volumes, masses, densities, and the like do not take exact values" (1993, 88). This habit is vindicated by the fact that "as an idealisation which gives sufficiently accurate results within the ranges of imprecision within which we work, it is indispensable" (ibid. 89). The same happens with Bayesian confirmation theory: as a rational reconstruction of people's attitudes towards theories, its contribution in terms of rigour, clarity and precision outweighs the costs deriving from its descriptive inaccuracy.

Of the various measures of confirmation that have been proposed so far, the most popular one takes the degree of confirmation c of a hypothesis H by evidence E to be given by the difference between the posterior and the prior probability of H, i.e.

$$c = \Pr(H/E) - \Pr(H). \qquad (2)$$

Comparisons of this quantitative account of confirmation with competing proposals, like Hempel's instance confirmation and the hypothetico-deductive method as well as critical evaluations of its success in resolving traditional problems in the philosophy of science can be found in Earman (1992, chs. 3 and 4) and Howson and Urbach (1993, ch. 7).

1.2 Interpretations of Epistemic Probability

Perhaps the most significant way in which Bayesians diverge is over the answer they offer to the *epistemological question* what constraints, if any, should be applied to prior probabilities. The attempts to discover sound constraints for these prior values represent a spectrum of positions, varying from pure personalism to the logical interpretation, and correspond to different ideas regarding the amount of objectivity that can be achieved within the context of an analysis of partial belief. In what follows I provide a summary of these views.

1.2.1 Subjective Bayesianism

At one end of the spectrum lie *Subjective Bayesians* (or pure personalists), like De Finetti and Jeffrey, who claim that *any* prior probability assignment is admissible as rational as long as it satisfies the constraints imposed by the synchronic Dutch Book argument, namely the axioms of probability[16]. Given the minimal and relatively uncontroversial character of these constraints, this version of Bayesianism is taken by most people to be the only clear and unequivocal one[17]. In fact, its greatest virtue is precisely that it seems not to move beyond those constraints, which reflect requirements of logical consistency, and whose justification is as unequivocal as possible.

Consequently, the view that probability is 'a new kind of logic', adumbrated by Ramsey (1926), sketched by Howson and Urbach (1993) and finally systematically outlined by Howson (2000), is not only a rigorous explication of the fundamental logical properties of the probability axioms, but also explains why the pure subjectivist strand of Bayesianism has proved so appealing. The explanation, in straight parallel with deductive logic, consists in the fact that since this new logic deals with consistency questions only, it cannot possibly be concerned with substantive, and hence extra-logical, questions, like which specific value our prior probabilities ought to take. This will inevitably be "*a probability assignment that is not deducible from the probability axioms*" (Howson 2000, 134)[18].

---

[16] Obviously those convinced by the diachronic Dutch-book arguments would add Bayesian Conditionalisation as a requirement of rationality (with the exception of van Fraassen, as we noted in footnote 9). The synchronic constraints though are universally accepted while, as we saw, the principle of conditionalisation is, to the say the least, dubious.

[17] Worrall (2000, 131-132) characterises it as the 'pure' version of Bayesianism.

[18] Emphasis in the original.

It is interesting to note at this point that, although Howson's analysis provides the most rigorous framework for the defence of the subjective variant of Bayesianism, Howson himself is a subjectivist *only* for those domains where there is no frequency information available. In case we can obtain estimates of frequencies, Howson does impose constraints on our degrees of belief, as we shall see shortly in more detail.

There is, however, an obvious problem with this interpretation of the calculus, namely that it (seemingly) bars any attempt to build an objective account of ampliative, non-deductive inference. If *any* prior assignment is admissible as rational, then there seems little hope for a definitive solution of many important problems in the methodology of science, despite the presence of strong intuitive presumptions that only one answer is 'rational'. Put differently, Bayesian personalism seems to be equating rational constraints with logical constraints; all the rest are relegated to the domain of plausibility considerations, justified at best only on pragmatic grounds with no real epistemological significance. Intuitively, though, the probability of the sun rising tomorrow just cannot be (in an intuitive, but still very strong, sense of 'cannot') less than something like 0.9999999999.

The usual subjectivist answer to this problem revolves around the celebrated idea of the 'washing-out of priors'. This is another cover-term referring to a class of mathematical results, which establish that under certain conditions the posterior probabilities of different agents on a proposition given the evidence will converge in the limit, no matter what their prior assignments are. In their simplest form, such theorems assume that agents are equally 'dogmatic' in that they assign zero probability to the same elements of the probability space (cf. Earman 1992, 142)[19]. In this way, it is argued, the problem of subjectivity of the priors is not as troubling as it might look *prima facie* since, in the long run, differing initial opinions will match.

It is not at all obvious, however, that these results suffice to solve the problem. First of all, the rate of convergence of opinion is not specified. This means that the limit could very well be very far, which would render the results useless for all practical purposes. Secondly, the conditions required for these theorems to hold are *not* on a par with the axioms of probability. Hence, one presumably would have to invoke plausibility considerations for their justification, which the subjectivist so much dislikes. Finally, Earman has noted that "for some aspects of the objectivity problem not only is the long run irrelevant, so is the short run" (ibid. 149). More specifically, a plausible account of confirmation ought to explain widely shared judgements as to whether a particular piece of evidence supports one theory better than another. If, however, wide disagreement is permitted over the prior probability assignment those theories merit, then (for some measures of confirmation at least) the answer we shall get will not be unequivocal. It should be obvious that merger-of-opinion results, either applying in the long or in the short run, are just useless for the purpose of explaining the impact a particular bit of evidence has on two or more competing theories[20].

---

[19] Earman (1992, 142-147) contains an exposition of more complicated convergence-of-opinion theorems.

[20] Earman (1992, 149) also mentions the problem of comparing the evidential value of different bits of evidence relative to *a particular theory* as another instance of the inability of merger-of-opinion results to ground the objectivity of scientific inference within the context of Bayesian personalism. This issue, however, seems to be orthogonal to the properties of merger-of-opinion results Earman

From this brief sketch of Subjective Bayesianism one can see that its main virtues are 1) its simplicity and 2) the ease with which it justifies the constraints it imposes, resulting from the denial to impose any kind of *substantive* constraints to prior probabilities (i.e. constraints beyond what logic mandates). On the other hand, its main shortcoming is that it seems to preclude from the outset any hope for an objective account of inductive reasoning, failing to do justice to our rudimentary intuitions. Much later, in chapter 5, I present a re-interpretation of Subjective Bayesianism which I hope does more justice to our intuitions regarding the objectivity of certain aspects of scientific reasoning.

1.2.2 The Logical Interpretation

At the other end of the spectrum lies the *logical interpretation* of probabilities, espoused by Keynes (1921)[21] and Carnap (1950). According to this approach one can fix the prior probability assignment uniquely as a matter of *logic*.

---

discusses. In fact, if there is a problem here at all, it would seem to stem from the value the *prior probability of the evidence*, rather than the prior of the theory, should take. Indeed, whatever the prior probability of the theory under consideration, on the assumption that there is a well-defined value for the prior probability of the evidence and well-defined relationships (deductive or statistical) between the theory in question and the evidence, it is a straightforward matter to determine *objectively* which bit of evidence provides better support to the theory under consideration. For a criticism of Bayesian personalism centring on the problems relating to the determination of the prior probability of the *evidence* see Worrall (2000, 133-135).

[21] It has to be noted that Keynes' ideas were rather different from the ones I attributed to most Bayesians in that 1) he thought that not all probabilities are numerical, and 2) as a result, probabilities could only be partially ordered (Keynes 1921, 27-39). One can easily construct a Bayesian reading of Keynes' views, however, by ignoring his qualms about non-numerical partial belief.

The principle on which the burden of objectification rested according to the logical interpretation, is none other than the (in)famous *Principle of Indifference*. This asserts that "if there is no *known* reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability" (Keynes 1921, 42)[22]. This *prima facie* plausible guideline was accorded logical status by the defenders of the logical interpretation and suggested a way to render prior probabilities definite and objective[23].

Even in Keynes's time, however, it was well known that the Principle of Indifference gives rise to a series of difficulties, all relating to the fact that the principle is not language-invariant, or, in other words, that it is sensitive to the choice of variables[24]. For example, according to the Principle of Indifference the probability of a book, picked at random from a library we have never visited before, being red as opposed to non-red is ½; but at the same time its probability of being red as opposed to being either green or blue or yellow is ¼. Hence, we get

---

[22] Emphasis in the original.

[23] Hence, the logical interpretation of probabilities should not be confused with the logical interpretation of *subjective* probability. The difference lies precisely in the fact that the latter does not admit any kind of constraint like the Principle of Indifference, which goes beyond the standard probability axioms and attempts to single out the one true value. According to proponents of the logical interpretation of subjective probability, such constraints are arbitrary in that they attempt to settle non-trivial questions, while "logic [is] essentially non-committal on substantive matters" (Howson and Urbach 1993, 71-72).

[24] For an exposition of the various paradoxes the Principle of Indifference gives rise to see Keynes (1921, ch. 4), Gillies (2000, 37-42) and Howson and Urbach (1993, 59-62). My discussion follows closely Gillies (2000) throughout.

contradictory results by applying the principle to different partitions of our variable space.

In an attempt to rectify the situation, Keynes suggested the plausible further restriction that the principle be applied only when a *finite* set of mutually exclusive and *indivisible* alternatives is determined. While this takes care of some finite cases like the book-paradox, however, it does *not* generalise to all finite cases. Consider a series of coin tosses, where 0 stands for heads and 1 for tails. For $n = 2$ tosses, we get four possible outcomes: 00, 10, 01, and 11. Assuming that the outcome of one heads in two tosses is divisible, then 10 and 01 are distinct and the probability of each outcome is ¼. If it is not, then the probability of each outcome is 1/3. The interesting thing with this case is that, while the first option may seem the more natural, it results in making learning from experience impossible. In this way, though, it negates the very purpose for which it was selected, i.e. that of implementing the idea of inductive logic.

To see why this is so consider the case where we know the outcomes of n tosses (our evidence E) and we inquire about the probability of the hypothesis (H) that the next toss will be heads. The 'natural' understanding of divisibility yields $\Pr(H) = ½$. Furthermore, $\Pr(E) = 2^{-n}$ since there are 2^{n} possible sequences of heads and tails in n trials and E is just one of these possibilities. The posterior probability of the hypothesis that the next toss will yield heads is then $\Pr(H/E) =$ $\frac{\Pr(E \wedge H)}{\Pr(E)} = \frac{2^{-(n+1)}}{2^{-n}} = 1/2 = \Pr(H)$. It follows that if we still want learning from experience to take place, we have to promote the 'less natural' partition of possible alternatives. Yet there seems to be no other independent, non-*ad hoc*, reason for this choice.

Furthermore, as Ramsey had already noted in his 1922 review of Keynes' treatise, Keynes' proposal fails to accommodate cases where *infinite* magnitudes are involved (Ramsey 1922 [1989], 221). One such case is the infamous water/wine example. Suppose we have a mixture of wine and water and we know that there is at most 3 times as much of one as of the other. Hence, we have $1/3 \leq$ wine/water ≤ 3 and applying the Principle of Indifference yields a uniform probability density in the interval [1/3, 3]. From this, though, it follows that Pr(wine/water ≤ 2) = 5/8 $\neq$ Pr(water/wine $\geq$ ½) = 15/16. In words, the probability of the wine being at most twice as much as the water is different from the probability of the water being at least half of the wine, despite the fact that these sentences are *logically equivalent*.

The general problem can be put as follows: for a continuous parameter θ taking values in an interval $[a, b]$ and a continuous and suitably regular function $f(\theta)$ defined in the same interval $[a, b]$, such that $a \leq \theta \leq b$ is logically equivalent to $a \leq f(\theta) \leq b$, there is no guarantee that applying the Principle of Indifference to both will yield the same results. In fact, the general case is that the opposite will happen, creating further contradictions. As a result, nowadays the consensus has it that, despite ingenious attempts to resolve some individual paradoxes (cf. Jaynes 1973), the Principle has to be discarded as a principle of logic.

So far, then, we have only looked at the opposite poles of the spectrum representing the various interpretational stances one can take towards epistemic probability. Between these two, however, there is a plethora of proposals - all of which try to strike a balance and get the best of both worlds. In order to do so, those proposals must introduce some degree of objectivity while being arguably soundly justified. There are three main rival candidates. In ascending order of strength they are 1) *tempered personalism*, 2) *empirically-based subjective Bayesianism* and 3) *objective Bayesianism.*

1.2.3 Shimony's Tempered Personalism

The idea of tempered personalism is associated with Abner Shimony and is peculiar in that it was not the result of general probabilistic or statistical considerations but was developed specifically for the purpose of attaining a more objective philosophy of science. Its core idea consists in a minor constraint imposed on the prior probability assignments to scientific hypotheses and that is why it only slightly deviates from the strict subjectivist approach.[25]

Shimony put forth the following 'prescription of open-mindedness' as a plausible constraint on prior probabilities:

> "the prior probability ... of each *seriously proposed hypothesis* must be sufficiently high to allow the possibility that it will be preferred to all rival, seriously proposed hypotheses as a result of the *envisaged observations*" (Shimony 1970, 101; my emphasis).

This he calls the 'tempering condition'. Its aim is not so much to objectify the prior probability assignment in the sense of singling out 'the one true value' - he explicitly allows for much subjectivity and variation in this respect (ibid. 102)- but rather to combat one of the problems we encountered when discussing the idea of the 'washing-out of priors'; namely that it only applies to the limit and that limit may be too far away. Shimony expresses this worry as follows: "If personal probabilities were used, unqualified by the tempering condition, then deep-seated

[25] The original presentation in Shimony (1970), nonetheless, contains a much greater extent of deviation, owing to his preferred non Dutch-book justification of the probability axioms.

prejudices … could produce immense (and possibly infinite) prior probability ratios; swamping [i.e. the 'washing-out of priors'] would consequently be impossible with a moderate amount of data, and consensus would be severely delayed" (ibid. 103). What the tempering condition is intended to achieve, then, is that, even with a moderate amount of data, consensus will be reached within a reasonable amount of time, thus ridding the idea of convergence of its dangerously impractical character. The end result is a potentially more effective way to alleviate the effects of arbitrariness in determining the prior probability assignments.

Obviously much rests on what a 'seriously proposed hypothesis' is and on what the 'envisaged observations' mean. Shimony, as might be expected, did not give an accurate characterisation of either notion. With regards to the first, he explicitly acknowledged the role of personal judgments in specifying what counts as a 'serious hypothesis'[26]. It seems that the most one can say in this connection is that the community of scientists creates vague rules as to what counts as a 'serious hypothesis' and what doesn't. Clearly, though, this does not amount to an adequate theoretical foundation for the notion of 'seriously proposed hypothesis'. Similar remarks apply to the notion of the 'envisaged observations'. Though Shimony wants his tempering condition to effect convergence of opinion within a 'reasonable amount' of time and given only a 'moderate amount' of data, he does not specify any further the process which would produce the 'envisaged observations' or how long it should last. Personal judgement and subjectivity, then,

[26] In discussing the need to avoid both the danger of postulating arbitrary and inflexible sharp criteria as well as the one of leaving the choice of what counts as a serious hypothesis entirely up to subjective judgment, Shimony proposes "to formulate some methodologically sensible guidelines for decisions on this question, *without pretending to eliminate entirely the subjective judgement of the investigator*" (Shimony 1970, p. 110; my emphasis).

are bound to enter the picture again, presumably leaving the scientific community to determine internally and in practice the precise form and duration of their investigations. This vagueness in defining the 'seriously proposed hypothesis' or the 'envisaged observations' clearly counts among the weak links in Shimony's theoretical approach. We should also note, however, that any theory is practically certain to admit vague notions and concepts at some point; consequently, vagueness in this respect need not be damning for the 'tempering condition'.

Shimony also thought that his tempered personalism requires a different justification for the probability axioms from the usual Dutch-book argument, because he took the concept of 'rational degree of belief' to be undesirable as an *explicandum* of the concept of probability in the context of evaluating scientific hypotheses. His arguments were a) that a bet on a universal generalisation cannot be decided due to the open-ended character of such a statement, and b) the Popperian idea that the rational degree of belief on such a generalisation should be zero, or, at any rate, very close to zero, which evidently undermines the aims of the 'tempering condition'. His solution was to suggest the notion of 'rational degree of commitment' as the appropriate explicandum for 'probability' and to opt for the Cox-Good-Lucas justification of the axioms (ibid. 94-95 and 104ff.).

In the face of what I said about the character of the Dutch-book argument, however, it seems to me that one need not be that troubled with the *practical* fact that we can't decide the truth-value of a universal generalisation. Assuming one only takes the Dutch-book justification to be an abstract model aiming at conferring normative weight to its conclusions, whether we can *in fact* decide the truth-value of a theoretical hypothesis or not is beside the point of the argument. Furthermore, it is not at all obvious that the probability of a universal law should be zero. In fact there have been convincing criticisms of all of Popper's arguments for his thesis (cf. Howson 1973; Earman 1992, 92-95 and 104-109), which, together with the previous consideration regarding the character of the Dutch-book arguments, suggests that probability as degree of belief is a perfectly adequate notion even within the context of theory-evaluation.

More interesting is the way in which Shimony attempted to justify the tempering condition within the context of a naturalistic epistemology - an approach he dubbed 'Copernican'. In outline the idea is that in justifying the tempering model of inductive inference and confirmation, we have to use the results of precisely those domains of inquiry our model intends to vindicate. In doing so, we employ a prior assumption regarding both the reliability of our pre-scientific intuitions as well as the fruitfulness of the hypotheses we regularly come up with. Then, we are justified in assigning non-vanishing priors to our 'seriously proposed hypotheses', since we believe that our reliable habits, which gave rise to them, also give them a good chance to be the true ones and, hence, make them worthy of a unanimously high degree of belief.

This line of reasoning sounds hopelessly circular. Shimony himself remarks that "the problem of circularity is evidently raised by the fact that a body of natural knowledge, which was itself tentatively established by means of scientific inference, is methodologically significant" (Shimony 1970, 139). Nonetheless, he goes on to dismiss the danger, holding that the circularity under consideration is non-vicious on the grounds that theories are still open to criticism and refutation (ibid. 159-160). Although I think the charge of circularity is well-founded, I will postpone detailed discussion for a later chapter. The reason is that this attempted justification of the 'tempering condition' by endorsing allegedly non-viciously circular arguments is methodologically on a par with some recent popular defences of scientific realism. Hence, since assessing its legitimacy is directly connected with the possibility of offering a Bayesian reconstruction of at least one version of the central realist argument, I will take up this issue again only when I explore the possibility of offering the realist arguments in a Bayesian form.

1.2.4 Empirically-Based Subjective Bayesianism

A stronger constraint on the prior probabilities relates to the way empirical knowledge is taken into account. *Empirically-based subjective Bayesians* propose that, whenever one possesses reliable information about the theoretical (i.e. limiting) frequency of the occurrence of a type of event, then one's degrees of belief that a singular event of this type will occur should be set equal to the theoretical frequency (or objective *chance*) the available data dictate. This idea also originates with Ramsey (1926, 195-196), who, as I remarked earlier, adopted a dual interpretation of the probability calculus, allowing probabilities to stand both for degrees of belief and frequencies. His account, however, remains rather sketchy, since he never managed to give a systematic account of his frequency concept of probability. Empirically-based subjective Bayesians follow Ramsey's pluralistic approach and propose a more rigorous way to constrain the prior probabilities in the light of empirical information.

A viable pluralistic approach to probabilities, of course, demands a multiplicity of fully worked out accounts of probability. We have seen that probability as personal degree of belief is one of them; empirically-based Bayesians also accept the *frequency* interpretation as a legitimate account of probabilities – an interpretation which goes back to von Mises (1928) and Hans Reichenbach (1949).

This interpretation rests on two basic principles. First, the probability of a variable taking a certain value is defined to be equal to the limiting frequency of the occurrence of this value over the total number of trials, when this number tends to infinity. Symbolically:

$$\Pr(E/\mathbf{C}) =_{df} \lim_{n\to\infty} m(E)/n,$$

where **C** stands for the sequence of these trials, *m(E)* for the number of times the variable in question has taken some value E and n for the total number of trials. Von Mises introduced the term *collective* to refer to sequences like **C** and defined it as follows: "[the term 'collective'] denotes a sequence of uniform events or processes which differ by certain observable attributes, say colours, numbers or anything else" (1928, 12). As is obvious by the definition, probability as frequency makes sense *only* after a collective has been specified.

This definition is supplemented by a second principle relating to the explication of the notion of *randomness* within a collective. Von Mises' idea was to identify randomness with the impossibility of producing a gambling system, in other words an algorithm with the ability to produce frequency values for subsequences of the original collective different from those of the collective itself. For a subsequence **C'** of the initial collective that is:

$$\lim^{C'}_{n\to\infty} m(E)/n = \lim^{C}_{n\to\infty} m(E)/n$$

As it turns out, using the resources of the theory of recursive functions this principle can be given a rigorous mathematical expression. Furthermore, coupled with the axiom of convergence, it allows for the standard axioms of probability to be recovered, thus providing an adequate foundation to the frequency theory[27].

A different concern is that, despite the frequency theory being a legitimate interpretation of the calculus, it might not be appropriate as a constraint on our degrees of belief. This worry is founded on the *single-case* character of our degrees of belief, i.e. the fact that they can refer to propositions about *singular* events, and the explicitly *non-single case* character of the frequency approach, which refuses to attach probabilities to such events and can only make sense of the concept of probability relative to a collective. Hence, an account needs to be given of how information about frequencies can inform our degrees of belief.

Howson and Urbach (1993, 345) have responded to the problem by offering a Dutch-book argument for the connection between them. They were able to demonstrate that failing to set one's degrees of belief equal to frequencies leads to sure loss *after some finite time*. In this way, Howson and Urbach attempt to maintain the objective character of the frequency interpretation and transfer it to our degrees of belief. To be sure, our knowledge of frequencies is not always complete. Nothing hinges on that, however, since we can always probabilify hypotheses about what the limiting frequency of a particular collective is (cf. Howson 2000, 234-236).

This process is summarised in the form of (a version of) the Principal Principle, originally introduced in Lewis's (1980) treatment of the notion of 'chance'. 'Chance' is usually understood as objective *single-case* probability rather

[27] For more details see Howson and Urbach (1993, ch. 13), Gillies (2000, 105-109) and Williamson (2005a, sec. 6). A bothersome feature of the frequency interpretation, however, is that it cannot accommodate countable additivity except by *fiat* (cf. Gillies 2000, 110-111).

than frequency and Lewis most certainly follows this tradition (cf. Lewis 1980, 83-84). Empirically-based Subjective Bayesians, however, reject the existence of single-case objective probabilities and identify 'chance' with frequency (cf. Howson 2000, 223 and 230-232). In effect, they adopt a version of the Principal Principle in which 'chance' is understood as limiting relative frequency. With this proviso in place, the principle requires that our degree of belief that some variable will take the value E at some instant given that the chance of E is r be set equal to r. Formally:

$$\Pr_t(E/\text{Chance}_t(E) = \text{r}) = \text{r}.$$

Hence, once we have reliable knowledge of the frequency (or 'chance') of some type of event, empirically-based subjective Bayesians require that we set our prior degrees of belief that the event will occur on the 'next' occasion equal to that frequency. Similarly, if our frequency data specify only an interval, then our prior belief is free to lie anywhere in this interval but not outside of it on pain of inconsistency. In this way, assuming a certain amount of pluralism in the admissible interpretations of the probability calculus, some Bayesians try to reduce arbitrariness in the assignment of priors. As is obvious, however, this can only be achieved when there is frequency data available. When there isn't, either due to practical difficulties or the nature of the subject under discussion, we are inevitably thrown back to pure personalism.

Two things are worth noting here. First, although frequentism is an objective account of probabilities, empirically-based subjective Bayesianism will not be fully objective even in the presence of reliable information about frequencies until it faces effectively the *reference-class problem*[28]. Recall that probability as frequency makes sense only within a 'collective', i.e. a large set of uniform events. Now, there are many ways to assign a particular event within the domain of a collective and it is not at all self-evident that there is a uniquely correct way to do this. For example, inquiring about the probability that Mr Smith, aged 40, will experience heart problems within the next 5 years, it is possible to classify him as a man aged 40, an Englishman aged 40, a Londoner aged 40, a Londoner aged 40 who drinks a lot etc. Clearly, our decision will be based on the factors that we think might be relevant to heart problems, the available statistics etc. The problem, then, is to decide which of the possible classifications to use in order to answer our initial query.

One possible solution to this problem is provided by the principle of *the narrowest reference class*, which states that for collectives **C**, **C'**, **C''**, if **C** is a subset of **C'**, which in turn is a subset of **C''**, one should opt for the narrowest class **C**. Though plausible, however, this principle does not answer to all problems. Occasionally there are more than one narrowest reference classes for which there are reliable statistics. What if, in our hypothetical example, Mr Smith also exercised regularly and there were reliable statistics for Londoners who drink a lot, Londoners who exercise regularly but *not* for Londoners who both drink and exercise? In that case the Principal Principle would yield different degrees of belief for the same proposition, depending on the reference class assignment we would choose, thus failing to introduce a fully objective element into the subjectivist view. Furthermore, our knowledge of what counts as a relevant factor to heart problems changes over time. This means that even if we are able to single out one and only

[28] My discussion of the 'reference-class problem' follows closely the one found in Gillies (2000, 119-125).

one narrowest reference class according to *current* knowledge, future developments may very well show our current probability estimates to be way off the mark. Hence, Gillies' conclusion that "probabilities for single events,...though sometimes objectively based, will nearly always fail to be fully objective" (2000, 120) seems justified.

Secondly, besides the explicitly non-single-case character that Empirically-based Subjective Bayesians attribute to 'chance', their approach is quite different from Lewis' own discussion of objective credence in yet another respect. Empirically-based Subjective Bayesianism makes use of an *independent* account of 'chance', namely that of the frequency interpretation. Lewis, on the other hand, implicitly defines 'chance' in terms of its relation to beliefs (cf. Howson 2000, 222). This strategy, however, leaves the way in which one can sensibly retain talk of *objective* chance entirely mysterious[29].

1.2.5 Objective Bayesianism

An even stronger constraint on prior probabilities has given rise lately to the stance called *Objective Bayesianism*. Objective Bayesians agree with strict personalists that a rational degree of belief must satisfy the axioms of probability, however, like the Empirically-based Subjective Bayesians, they deny the converse: not all sets of degrees of belief that satisfy the axioms of probability are rational. Furthermore, although they agree with empirically-based subjective Bayesians that

[29] An alternative reading of Lewis, which fares better in this respect, would be to reserve the term 'chance' only for the probabilistic assignments derived by means of *probabilistic laws of nature*, understood according to Lewis' 'best-system analysis of laws'. In his (1994, 477-480), Lewis is quite explicit about it, unlike the earlier cryptic remarks found in his (1980, 1986a).

empirical constraints ought to have a bearing on our degrees of belief, they add to this requirement one further 'logical constraint': given the empirical background information, an agent's belief function should be as non-committal as possible, reflecting also one's "lack of knowledge of the world" (Williamson 2006, 13)[30]. The end result is an objective version of Bayesianism, imposing very stringent constraints on our degrees of belief.

Objective Bayesians give priority to the empirical over the 'logical constraints'. They incorporate them into background knowledge and demand that they be taken into account before logical considerations come into play (cf. Williamson 2006, 14). Hence, if reliable frequency information determines an interval of frequency values, our degree of belief should not fall outside this interval. Once this is known, however, logical considerations allow us to single out *one particular belief function*. The 'logical constraint' imposed takes the form of the *maximum entropy principle*, introduced by Jaynes (1957) in the context of statistical mechanics. It requires that from a variety of choices of probability distributions determined by our background knowledge, we choose the one which maximises entropy, i.e.

$$\mathrm{H}(p_i) = -k\sum_{i=1}^{n} p_i \log p_i ,$$

where k is a constant. Being a measure of uncertainty satisfying many intuitive desiderata[31], entropy is taken to characterise precisely the agent's lack of knowledge, which, according to Objective Bayesians, ought to constrain one's degrees of belief. The end result of this process is preference for a belief function,

[30] Any page references to Williamson (2006) and (2007) found in this section correspond to the on-line version of Williamson's papers (the full on-line reference is contained in the bibliography).

[31] See Jaynes (1957, 621) for more details.

which is as middling as possible, or, put differently, as far away from the extremes as possible.

It follows, then, that when our background information suffices to single out one frequency value and not just an interval, the answer Objective Bayesianism gives as to what one's degree of belief should be coincides with that of Empirically-based Subjective Bayesianism. Similarly, when there is no background empirical information at all, the maximum entropy principle's result is the same as the one that the Principle of Indifference would give, i.e. equiprobability of all possible outcomes. In fact, Jaynes himself described it as "an extension of the principle of insufficient reason [i.e. the Principle of Indifference]" (1957, 623), so, given the problems regarding the justification of the Principle of Indifference, it is an urgent question to see how its extension might avoid those problems.

The first attempt at a justification was offered by Jaynes himself in his seminal 1957 paper. Jaynes suggested that "that the maximum entropy distribution may be asserted for the positive reason that it is uniquely determined as the one which is maximally non-committal with regard to missing information, instead of the negative one that there was no reason to think otherwise" (ibid. 623). It should be clear, however, that this kind of approach is question-begging. As Williamson (2007, 6) notes, there is no good *a priori* reason why maximally non-committal degrees of belief ought to be desirable. In fact, the whole edifice of pure personalism rests on the seeming absence of such *a priori* reasons, so simply postulating a desideratum will not do.

In his (1973) Jaynes offered a reformulation of his desideratum as follows: "*in two problems where we have the same prior information, we should assign the same prior probabilities*" (228); and in the next sentence he calls it "psychologically impossible to quarrel", possibly appearing "so trivial as to be without useful content" (ibid. 228). Still, though, one should not rest content with assertions like these. On the one hand the psychological effects of a principle need have absolutely nothing to do with its status as an *epistemic* principle. On the other, if the Principle of Indifference, which is implied by this desideratum, fails to be a trivial logical principle, then the desideratum itself fails to be trivial, in which case a more persuasive argument than mere allusion to appearances is needed.

A different line of argument has been offered recently by Williamson (2007). After reviewing several application-oriented attempts to justify the maximum entropy principle and finding them wanting, Williamson offers his own justification by appealing to *caution*. His point, which he demonstrates formally (ibid. 14-16), is that "the maximum entropy principle is *on average the more cautious policy when it comes to risky decisions*" (ibid. 16)[32], where by 'risky decisions' he refers to ones in which caution matters most since the unfavourable outcomes are very costly[33]. He concludes that one is thereby generally justified in using the maximum entropy principle in fixing one's prior degrees of belief.

While this is an interesting result, it is of both limited applicability and also rests squarely on pragmatic issues. It may well be the case that in everyday life we tend to be cautious in the face of risky decision situations involving costly outcomes, yet risky decisions in that sense are not the only ones we are typically

---

[32] Emphasis in the original.

[33] We must note, however, that Williamson's use of the term 'risky decisions' is rather idiosyncratic. Indeed, as Williamson acknowledges himself, "every course of action has its associated risks" (2007, 12), which means that all decisions under uncertainty may reasonably be held to fall under the concept of 'risky decisions', irrespective of the severity of their unfavourable outcomes. At any rate, in what follows and for purposes of discussion I have adopted Williamson's use of the term.

faced with. *A priori* there seems to be no reason why one should focus only on the 'risky decisions' for purposes of justification of the principle, thus completely disregarding their non-risky and equally frequent counterparts. Furthermore, even if we restrict attention to 'risky' circumstances, it is by no means a universal law that people tend to be cautious when facing risky decision situations but rather a rule of thumb with no self-evident normative character. Finally, regardless of the empirical merit of attitude-to-risk theories, when it comes to non-practical but purely epistemic affairs it is not at all clear how this result can have any significance at all. Consider, for example, the task of determining the prior probability of a hypothesis in an attempt to evaluate its posterior after the data is collected *irrespective of the use* we intend to put it to. The consideration regarding the cautious character of the maximum entropy principle seems to have no application to such a case.

Williamson is, admittedly, explicit that the justification of the principle is bound to go beyond both the evidence and logic (ibid. 9). And argues that adhering to it relative to practical or even scientific affairs is the sensible thing to do. In fact, Jaynes' 1957 paper introduced the principle within the context of statistical mechanics with great practical success, while its use as a heuristic principle in other areas of natural science might well be valuable (cf. Gillies 2000, 48-49). Still, though, arguing this way only provides empirical support for the principle and *only* for its validity within the particular practical domains of inquiry. Hence, it is still not clear how such considerations could constitute a case for its application outside physics and the domain of practical problems[34].

[34] It should also be noted that the maximum entropy principle is not free of technical problems either. The most important of them is that by generalising the principle to apply to uncountably infinite domains uniqueness is jeopardised. There is no guarantee, that is, that the modification of

So far, then, I have sketched the main theses one has to subscribe to in order to qualify as a Bayesian and analysed in some detail the flexibility of this general framework relative to the amount of objectivity one is free to endorse. Schematically, one can arrange the aforementioned interpretations on the basis of their subjectivity as follows:

Subjective Bayesianism → Tempered Personalism → Empirically-based subjective Bayesianism → Objective Bayesianism → Logical Interpretation

For the purposes of the discussion to follow, there is no need to assume that any of the above interpretations is the correct one. Rather, I will explore the extent to which each one of them might be put to use in examining the prospects of a Bayesian reconstruction of the debate regarding scientific realism.

---

the principle will be able to single out one unique belief function any more, which in turn means that subjectivity re-enters the picture. Still, Williamson is quite reassuring about this issue, suggesting that objective Bayesianism need not be fully objective. It would suffice if it is merely *the most* objective Bayesian account of all (cf. Williamson 2007, 24-25). Given that the exploration of this dimension of the maximum entropy principle transcends the scope of this thesis, for present purposes I shall assume that this sort of technical complications can be accommodated within Objective Bayesianism without vitiating its character.

Chapter 2

Bayesianism and 'Theory Acceptance'

The aim of this thesis is to investigate in a systematic way the relationship between Bayesianism and the thesis of scientific realism. The most obvious stumbling block to representing the realism debate in Bayesian terms is the issue of '*theory acceptance*'. Bayesians seem to lack an appropriately strong notion of 'theory acceptance', which, as we shall see in more detail shortly, is often assumed to be an indispensable characteristic of the realist view. If this charge is upheld, however, then the prospects for presenting an illuminating Bayesian reconstruction of the realism debate are slim. Indeed, if the concept of 'theory acceptance' is essential to the realist thesis but impossible to accommodate within the Bayesian framework, it follows that Bayesianism is ill-equipped to express essential dimensions of the realist position and, hence, contribute additional insight to the debate.

In this chapter I argue to the contrary that the Bayesian theory provides the very conceptual tools necessary further to elucidate the notion of 'theory acceptance', which is only *prima facie* transparent and unproblematic, even for a realist. More specifically, I develop and defend a *reductive theory* of 'theory acceptance', according to which Bayesianism helps us *replace* the primitive idiom of 'theory acceptance' with the more precise language of probabilities. In the process, I also discuss and dismiss various alternative Bayesian accounts that might be given of 'theory acceptance'. Finally, I show how my own preferred view is to be used in modelling the realist thesis and its competitors.

2.1 A Challenge to the Bayesian

2.1.1 Scientific Realism - Preliminary Definition

A *strong* conception of scientific realism can plausibly be taken to embody the following three theses, as outlined in Psillos (1999, xix): (R1) the *metaphysical* claim that the world enjoys a definite and mind-independent existential status[35], (R2) the *semantic* thesis that our scientific theories should be taken at face value, i.e. as making assertions about a reality 'underlying the phenomena' that are genuinely true or false rather than mere algorithms/instruments for prediction; and

[35] Put this way (R1) seems to be making use of an old-fashioned dualism, according to which the mind belongs to a different realm from the external world. A more appropriate way to formulate (R1) would be to assert that "there exists a structured and undivided reality of which the mind is part; moreover that far from imposing their own order on things, our mental operations are in their turn governed by the fixed laws which describe the workings of Nature" (Zahar 2001, 56). Expressed in this way, it becomes clear that no such dualism is assumed.

(R3) the *epistemic* thesis that our best well-confirmed scientific theories are reasonably assumed to be at least approximately true – that is, that the unobservable entities they postulate (or perhaps something like them) are real.

The first claim is to be contrasted with the traditional metaphysical doctrines of idealism and phenomenalism, and is taken for granted in all recent discussions. The second is directed against various instrumentalist philosophies of science, which view scientific theories as shorthand notation and mere predictive devices incapable of being true or false. Defending a semantic realist approach to scientific theories was a central concern in the heyday of logical empiricism, given the linguistic nature of the positivist approach to science. Recently, however, the focus has rested squarely on the third *epistemic* thesis, which aims to establish that we are *justified* in believing in the truth, or rather approximate truth, of our best scientific theories, i.e. that we are justified in believing that they reveal how the world actually is.

It has often been claimed that an equivalent way to present this epistemic attitude is to say that 'a scientific realist *accepts* our best current theories'. Richard Boyd, for example, has defined scientific realism as

> "the doctrine that the sort of evidence which ordinarily counts in favour of the *acceptance* of a scientific law or theory is, ordinarily, evidence for the (at least approximate) truth of a law or theory as an account of the causal relations obtaining between the entities quantified over in the law or theory in question" (1973, 1; my emphasis).

Bas van Fraassen has also linked scientific realism with 'theory acceptance' in his famous definition of scientific realism:

"Science aims to give us, in its theories, a literally true story of what the world is like; and *acceptance* of a scientific theory involves the belief that it is true" (1980, 8; my emphasis).

Van Fraassen is, of course, adamant that "acceptance of theories ... is a phenomenon of scientific activity which clearly involves more than belief" (ibid. 12). By this, he means to differentiate the strictly *epistemic* aspect of 'acceptance' from the *pragmatic* one. The former reflects our strictly epistemic dispositions towards a theory. The latter involves consideration of factors like one's personal commitments in terms of time, energy, research projects, institutional behaviour in, say, a university or even large-scale policy decisions regarding fund allocation (cf. van Fraassen 1980, 12-13 and Earman 1992, 194). We can extract from van Fraassen's characterisation, however, a distinct *epistemic notion* of 'theory acceptance' by isolating the epistemic aspects of van Fraassen's wider concept. From the present point of view, i.e. that of normative philosophy of science, it is this epistemic notion of 'theory acceptance' which is of exclusive interest.

Stathis Psillos agrees that the focus of philosophy of science should rest squarely on the epistemic dimension of 'theory acceptance' and offers yet another definition of scientific realism in terms of 'theory acceptance':

"One of the central concerns of modern epistemology of science has been to characterise what should be involved in accepting a scientific theory. Scientific realists suggest that acceptance should be equated with the belief that the theory is approximately true, and that this belief can be warranted and rational" (1999, 249).

It appears, then, that the notion of 'theory acceptance' is often viewed as an essential aspect of the realist position. This, however, raises the first difficulty for our Bayesian reconstruction of the realist thesis.

2.1.2 The Challenge

The debate on 'theory acceptance' is essentially a continuation of the discussion in the 1960s about *rules of detachment* in inductive logic, i.e. the problem of detaching (or 'accepting') the conclusion of an inductive argument. The various statements of realism I have just reviewed suppose that there is nothing particularly problematic about the realist's notion of 'acceptance'. The realist attempts to provide an argument for his thesis[36] and then, on the assumption that it is successful, he 'accepts'/detaches its conclusion without any (or much) qualification. At the same time, however, it is agreed that a Bayesian can uncontroversially accept a proposition only if this proposition is assigned *probability 1* (cf. Earman 1992, 193-194; Worrall 2000, 141-142). Yet, so the argument goes, scientists should not generally believe (except, perhaps, in very exceptional circumstances) that they can be in possession of such conclusive evidence in favour of a scientific hypothesis to assign it probability 1 and, hence, uncontroversially accept it from a Bayesian point of view. Indeed, assigning probability 1 to a scientific hypothesis amounts to saying that the probability of observing any piece of evidence conflicting with that hypothesis is 0[37]. Such a

[36] To be discussed in considerable detail later.

[37] This is because $Pr(E) = Pr(T) \times Pr(E/T) + Pr(\neg T) \times Pr(E/\neg T)$. Since $Pr(T)=1$, $Pr(\neg T)=0$. Furthermore, if evidence E conflicts with T, $Pr(E/T)=0$. It follows, then, that $Pr(E)=0$. A residual problem, of course, is that under these assumptions $Pr(T/E)$ cannot be defined.

dogmatic attitude, however, is hardly sensible when it comes to discussing the epistemic merits of fundamental physical theories. Since the realist's decision to accept his theories is supposed not to be problematic *ex hypothesi*, it is the Bayesian who fails to accommodate an essential aspect of the realist epistemic attitude.

Psillos has raised this issue forcefully in his recent discussion of the prospects of giving a Bayesian rendition of *Inference to the Best Explanation.* As we shall see in considerable detail later, Psillos argues that the realist inference to the approximate truth of our best theories is an instance of the general ampliative inferential rule called Inference to the Best Explanation (IBE). One of the reasons IBE is supposed to serve the realist cause is that it

> "is typically seen as a rule of acceptance. In its least controversial form, IBE authorises the *acceptance* of a hypothesis *H*, on the basis that it is the best explanation of the evidence" (Psillos 2004, 83)[38].

Bayesianism, on the other hand,

> "does *not* have rules of acceptance. On a strict Bayesian approach we can never detach the probability of the conclusion of a probabilistic argument, no matter how high this probability might be. So, strictly speaking, we are never licensed to accept a hypothesis on the basis of evidence" (ibid. 83)[39].

Psillos' wording here is imprecise. Instead of saying that 'on a strict Bayesian approach we can never detach the probability of the conclusion of a probabilistic argument', he ought to say that 'on a strict Bayesian approach we can *seldom* detach the *conclusion* of a probabilistic argument'. First, we have just seen that a

[38] Emphasis in the original.

[39] Emphasis in the original.

Bayesian *can* detach the conclusion of a probabilistic argument if this conclusion is assigned probability 1. And secondly, if there is one thing that a strict Bayesian approach allows us to uncontroversially assert, this surely is the *probability* of the conclusion. Even if we correct Psillos' assertion, however, the problem remains: if scientific realism requires, as has been suggested, that we 'accept' our hypotheses as approximately true descriptions of reality, then Bayesianism will typically not be in a position to capture all the elements involved in the realist thesis. In Psillos' words: "[T]here is a tension between Bayesianism and the standard renderings of IBE. This might make us wary of attempts to cast IBE [and, by default, the realist position] in a Bayesian framework" (ibid. 83).

2.2 The 'Threshold Account'

One natural suggestion for a Bayesian who wants to endorse a notion of 'theory acceptance' is to postulate a 'rule of acceptance' (again analogous to a detachment rule) which makes use of a level of probability less than 1. Such a rule would be of the form:

> a hypothesis *H* is rationally acceptable to an agent on the condition that he assigns to *H* a probability value at least as great as some specified number **r** < 1.

The number **r** represents the 'probability threshold' that must be exceeded if a hypothesis is to be accepted and is usually considered to be close to, but (clearly) strictly less than, 1. This proposal has been adopted for example by Kyburg (1968, 1970), Harman (1986) and by Foley (1992), who even traces its origins back to Locke's *Essay Concerning Human Understanding* (1690).

An obvious initial problem with this proposal is that it is not clear at all what the value of the probability threshold **r** ought to be. A quick answer would be to say that this is a matter of convention and that contextual circumstances will in the end determine the exact value of **r**. This reaction, however, hardly constitutes an adequate answer not only for a scientific realist but also for anyone seeking even a moderately objective analysis of the issue of 'theory acceptance'.

A step towards the precise determination of the threshold value has been taken recently by Achinstein (2001), who argues that the threshold value 0.5 is a necessary condition for acceptance[40]. Suppose, says Achinstein, that one selects a threshold value **r** < 0.5. This value leaves open the possibility that an agent could both 'accept' a theory *H* and, at the same time, 'accept' ¬*H*, since they would both satisfy the threshold requirement (ibid. 116). This, however, is a highly undesirable result, telling in effect in favour of **r** = 0.5 as at least a necessary condition for acceptance.

Achinstein's argument is a strong one. Surely, we could not sensibly be regarded as accepting a theory to which we assigned probability less than 0.5. Surprisingly enough, however, Patrick Maher (1993, 137-139) argues for exactly this possibility[41] on the basis of the following thought experiment[42]. Imagine a

[40] In fact Achinstein talks about 'belief' and not 'acceptance' but we can take the two as synonymous since we are only interested in the epistemic aspects of 'theory acceptance'.

[41] It should be noted that Maher himself does *not* endorse the 'threshold account' but instead a decision-theoretic framework, which I will take up a little later in this chapter. His argument, though, can clearly be used in the context of the 'threshold account'.

[42] The thesis that **r** = 0.5 is necessary for acceptance has also been questioned on the basis of the 'preface paradox', introduced by Makinson (1965) (cf. Foley 1992, 113-114). The 'preface paradox', however, has been formulated in many different ways over the years, "even to the extent

hypothetical encounter with Einstein. Presumably Einstein did 'accept' the General Theory of Relativity (GTR) without serious qualms. Suppose, however, that he were offered two options:

1) World peace if GTR is completely correct.
2) World peace if GTR is false in some way.

Maher claims that Einstein – obviously given his pacifist inclinations- would have chosen (2). Maher adopts the *preference interpretation* of the probability calculus, which suggests that "we understand attributions of probability ... as essentially a device for interpreting a person's preferences" (ibid. 9). On this interpretation one can derive (through a class of results known as *representation theorems*) the usual probability functions on the assumption that people's preferences satisfy certain relatively uncontroversial qualitative conditions. From Einstein's alleged preference for (2), then, one can conclude that he would accord probability less than 0.5 to GTR. Yet, suggests Maher, we cannot really say that the father of GTR did not accept GTR, hence the conclusion that high probability is not necessary for acceptance.

It is not difficult to see what is wrong with this argument. No scientific realist asserts that GTR is *completely* correct. He merely asserts that GTR is *approximately* correct. Notwithstanding the problems with the notion of approximate truth, the choice between the two alternatives would not be that easy if the question were posed in the approximate-truth version. In fact, if Einstein were offered a choice between

---

that one might want to distinguish between various Preface Paradoxes" (Douven 2002, 393, fn. 5). For this reason, and in order to avoid complicating the discussion, I shall not delve into the details of the 'preface paradox' here.

1') World peace if GTR is *approximately* correct

2') World peace if GTR is not even approximately correct,

he would surely have thought of the offer as surprisingly promising for the prospects of world peace and opt for (1').

Hence, it is very hard to deny that the threshold value 0.5 is a necessary condition for 'theory acceptance'. It is equally hard, however, to find an argument specifying the value that constitutes a *sufficient* condition for acceptance. A threshold value as low as **r** = 0.5 intuitively seems too low to warrant acceptance of a hypothesis. At the same time, it is enormously difficult to argue for any particular value lying in the open interval (0.5, 1) in any principled way. Folley expresses this problem as follows:

> "There doesn't seem to be any principled way to identify a precise threshold… . Still we want to be able to say something, even if vague, about the threshold above which our degrees of confidence in a proposition must rise if we are to believe that proposition. What to say is not obvious, however, since there doesn't seem to be a non-arbitrary way to identify even a vague threshold" (1992, 112).

It seems, then, that the determination of what constitutes a probability threshold sufficient to warrant acceptance of a hypothesis must remain vague, a matter of stipulation varying with the context one faces. Foley himself concurs: "we deal with other kinds of vagueness by stipulation. Why not do the same here?" (ibid. 256).

No matter how this may be, for the purposes of our discussion there is no need to determine a unique threshold value for 'theory acceptance'. It will be assumed, that is, that the determination of the threshold on an *ad hoc* basis is permissible and does not pose special concerns. This is because the 'threshold account' must first face a more important and logically prior issue, namely the charge that *any* threshold value less than 1 fails as a sufficient condition for acceptance. The infamous lottery paradox is the paradigmatic illustration of this problem.

2.2.1 The Lottery Paradox and the Conjunction Problem

The lottery paradox, originally due to Kyburg (1961), casts doubt on the acceptance rule that high probability is sufficient for acceptance by applying the 'threshold account' to an ordinary lottery situation. Consider a fair lottery with n tickets, in which there is to be one winner. Each ticket has a probability of $1/n$ to win, which means that each has a probability $(n-1)/n$ of *not* winning. The paradox for the 'threshold account' arises as follows:

1. A hypothesis ought to be accepted if it exceeds an antecedently specified threshold **r**.
2. For any threshold **r**, one can construct a fair lottery by suitably determining n such that for *each* individual ticket i, one ought to accept the hypothesis {H: ticket i will *not* win}.
3. The *closure* principle states that one must accept any deductive logical consequence of any statement one accepts. This implies that if H is accepted with respect to each ticket individually, then it ought to be accepted with respect to their conjunction as well.
4. Hence, we have to accept the hypothesis that *no* ticket will win.
5. However, we know and, hence, 'accept' that one ticket will win the lottery, and hence that the claim that no ticket will win is simply false.

Therefore, we are led to a contradiction. Assuming that premise (3) is unproblematic, this paradox clearly undermines the viability of premise (1), i.e. the 'threshold account'.

All this might sound too contrived to be of interest for the scientific case. But the lottery paradox is in fact just a dramatic illustration of the more general *conjunction problem*. By the probability calculus we know that for two propositions A and B, given that $(A \wedge B) \vdash A$, it follows that $\Pr(A \wedge B) \leq \Pr(A)$. Furthermore, equality holds only when $\Pr(B/A) = 1$. Consequently, increasing the number of conjuncts, which are assigned probability less than unity (given the previous statements in the conjunction) but higher than a pre-specified acceptance threshold, will eventually lead to a probability value for the conjunction *below* that threshold. Now, if we substitute theoretical hypotheses for propositions referring to tickets and assume (quite plausibly) a) that the 'new' hypotheses are not logical consequences of some conjunction of the individual conjuncts that went before, and b) that these hypotheses merit probability less than unity, one can see, for example, that assigning degree of belief 0.8 to each of General Relativity and Quantum Mechanics would make each of them acceptable for, say, $\mathbf{r} = 0.7$, but their conjunction unacceptable. And this of course seems highly counterintuitive. More generally put and independently of the particular value that the threshold for belief in our theories might take, the circumstance might (and typically will) arise that our best theories, although individually acceptable beyond reasonable doubt, fail to pass the threshold of acceptability when conjoined.

One possible reaction to the paradox is to retain the 'threshold account' but suggest that the paradox is a *reductio* of premise (3), i.e. of the principle of deductive closure – rather than a *reductio* of (1). The first to take this line was

Kyburg himself, who claimed that "the lottery paradox shows that one cannot simply *impose* logical closure on bodies of belief" (1968, 118)[43]. For Kyburg problems like the lottery paradox were "fairly straightforward technical problems" in which "nothing of deep philosophical significance is involved" (ibid. 119); consequently, a technical solution of this sort must have seemed the most appropriate. Foley (1992) also went down that path in his attempted resolution of the lottery paradox. He notes that "[the lottery and the (structurally identical) 'preface paradox'] aren't paradoxes at all. They simply illustrate in a particularly dramatic fashion that rational beliefs are not conjunctive [i.e. closed under conjunction]" (117). Finally, a more recent example of this line of reasoning is afforded by Howson, who declares in his discussion of the lottery paradox that "such a closure principle is most certainly not a necessary one" (2000, 213)[44].

As I just mentioned, dropping the closure principle amounts to denying that whatever holds true individually for the members of a set of propositions also holds true for all their (joint) logical consequences, including their conjunction. If so, then the paradox disappears. One cannot infer (4) from (1) and (2) in the lottery case and, similarly, it no longer follows that one should accept the conjunction of individually acceptable theories, which gave rise to the conjunction problem.

There are, however, various problems with this proposal. First, the closure principle – for all Howson's (purely rhetorical) suggestion that it is 'most certainly' not necessary - is, in fact, surely difficult to resist. It is difficult to see intuitively how one could be committed to a and committed to b without being committed to $\{a \ \& \ b\}$. And this (strong) intuition can be reinforced by the following

[43] Emphasis in the original.

[44] Other examples of this line of reasoning include Makinson (1965) and Moser and Tlumak (1985).

consideration due to Douven (2002, 395). He notes that it is very doubtful that there is any psychological or phenomenological difference between asserting propositions a and b individually and asserting their conjunction $\{a \& b\}$. If so, it seems strange to endorse a logical principle which differentiates between these two cases.

A more fundamental problem with any attempt to question the closure principle has been noted by Kaplan (1981, 309). His argument focuses on the logical form of *reductios* in general, of which the lottery paradox is an instance. Kaplan correctly notes that a typical *reductio* makes use of the very same closure condition that those seeking a way out of the paradox allegedly deny. In a *reductio* we derive a contradiction from the conjunction of a number of propositions we already accept individually. Hence, someone who rejects the principle of closure as a way out of the lottery paradox seems to be committed to denying that *we* should be convinced by his way out of the paradox. This is because his argument is a *reductio*, whose force rests on the principle of closure he invites us to deny! In Kaplan's own words: "Not only are [those who deny the closure principle] licensing rational persons to ignore the force of reductios in general; they are licensing *us* to ignore the force of the very reductios they are using to convince us to reject [the closure principle]"[45] (ibid. 309).

Finally, it is not at all clear what 'acceptance' means in the case where deductive closure fails to hold. As I mentioned at the outset of this chapter, when I

[45] Emphasis in the original. One might try to get around Kaplan's objection by distinguishing between those sets of individually accepted sentences which are in fact deductively closed and those that are not. It still remains to be seen, though, how such a distinction can be motivated in a principled, non *ad hoc*, way rather than by a threshold theorist's irritation with the lottery paradox.

presented the challenge posed to the Bayesians, everyone takes 'acceptance' to be an unproblematic notion as long as the probability of a given hypothesis is 1. It is merely symptomatic of this unproblematic state of affairs that virtually everyone *affirms* deductive closure for 'acceptance' when the probability is equal to 1. When it comes to 'acceptance' for levels of probability below 1 though, where the 'threshold account' seems to be necessary, matters are quite different. Deductive closure (allegedly) does not hold anymore, since now we are not certain about the truth of our hypotheses. Doesn't this difference in logical properties, however, merely reflect a difference in the meaning of 'acceptance' between these two cases?

It is quite clear that in the first case we intend to make an unqualified statement while in the second we don't, since uncertainty demands caution. We cannot simply affirm a statement that we think 'as good as' true, in the way we would if we were convinced that it is true and, hence, that its probability is 1. Yet this is a clear case where the same term ('acceptance') is given different meanings under different circumstances. In Lakatosian style we should presumably introduce subscripts: $acceptance_1$ would refer to full acceptance with certainty, while $acceptance_2$ to acceptance despite uncertainty. But this is just another way of saying that we do not mean the same thing in both cases where acceptance-terminology is used. It is, then, quite reasonable to suggest that the allegedly different logical properties of each notion merely reflect this case of meaning variance and that the 'need' to alter the logical properties for acceptance under uncertainty arises out of the inappropriate insistence in using the cover term 'acceptance' for these two different cases. This in turn suggests, however, that we might as well dispense with acceptance_2 altogether and try to explain (away) the concept in some other way. This is the course I will pursue in the constructive part of this chapter.

Abolishing the principle of closure was a response both to the general conjunction problem as well as the more specific lottery paradox. The rest of the existing literature on the 'threshold account', however, focuses exclusively on the latter without any mention of the more general conjunction problem. The common thread running through these approaches is that they attempt to revise the threshold condition in such a way that it circumvents the particular instance of the lottery paradox while also maintaining the core intuition behind the 'threshold account'. Hence, these 'solutions' attempt to bypass the problem by allegedly demonstrating that the concept of rational acceptance, while legitimate in most cases, fails to apply to the lottery case. Pollock (1995), Ryan (1996) and Douven (2002) are typical examples of this strategy.

Douven, for example, proposes to replace the simple rule of acceptance, which requires merely that the probability of a hypothesis exceeds a certain threshold with a more demanding condition reading as follows:

> a hypothesis H is rationally acceptable to an agent on two conditions, 1) that the agent assigns to H a probability value no less than a previously specified number $\mathbf{r} < 1$, and 2) H is not a member of a Probabilistically Self-undermining Set of propositions (PSS)[46].

A PSS is a set of propositions with the following characteristics: a) the probability of each proposition in the set given background knowledge exceeds a certain value

[46] Here I use 'hypothesis' and 'proposition' interchangeably, since on the syntactic approach to theories I adopt hypotheses simply assert certain propositions.

r, but b) each proposition receives a probability value below **r** conditional on background knowledge *plus some conjunction of members of the set itself.*

Douven (ibid. 196-197) shows that the set of sentences about winning tickets in the lottery case constitutes a PSS and, hence, fails to meet the second condition of his proposed acceptance rule. Clearly, the first condition for a set to be a PSS is satisfied in the lottery case - since all propositions describing the event of each outcome not winning are assigned probability higher than any sensible threshold for acceptance. At the same time the second condition is also satisfied. This is a consequence of the fact that the lottery is bound to have a winner, so that each proposition ('ticket *i* will *not* win') conditional on all the others (describing the *non*-winning outcomes) will be assigned zero probability. Hence, the lottery case constitutes a PSS, which in turn means that acceptance talk – on Douven's construal - is inappropriate for the lottery case.

Amongst the problems facing this solution is the fact that it fails to capture basic intuitions regarding 'theory acceptance'[47]. It is not as if the lottery paradox is a very complex example far from ordinary life and so it would seem that if acceptance-talk is appropriate anywhere, it ought to be appropriate in this relatively simple example. And indeed Douven himself acknowledges that "there is some controversy over whether denying that it is rational to accept any of the propositions in [the lottery scenario] is *intuitively* correct" (ibid. 397, fn. 14)[48]. This means, however, that should an alternative account saving those intuitions be shown to be adequate, we would have good grounds to prefer it.

[47] This problem obviously affects all solutions modifying the acceptance rule with an eye to prohibit the application of the notion of acceptance to the lottery case.

[48] Emphasis in the original.

More fundamentally, notions such as the PSS do not even, on close examination, really seem to provide the definitive answer to the lottery paradox that they were initially taken to promise. In fact, Douven and Williamson (2007) have generalised the lottery paradox in such a way that it now renders not only Douven's (2002) proposal but also all solutions, which suggest that acceptance-talk fails to apply to the lottery scenario in particular, at best incomplete. Consider an arbitrary proposition φ, for which $\mathbf{r} < \Pr(\varphi) < 1$. By constructing a series of 'trivialisation arguments' Douven and Williamson show that one can always include any proposition exceeding a previously specified probabilistic threshold in a set which fails to satisfy any amended threshold condition. As a result, "not only lottery propositions, but *all* propositions having non-perfect probability fail to qualify as rationally acceptable" (ibid. 760)[49]. This clearly implies that the strategy of making the threshold condition more demanding fails to provide an acceptable general solution to the lottery paradox.

It seems, then, that the threshold-probability route does not constitute a satisfactory Bayesian account of 'theory acceptance', given the puzzles its various versions give rise to.

2.3 A Decision -Theoretic Account of 'Theory Acceptance'

Some other way of modifying the Bayesian framework would, it seems, need to be found if the realist intuition about 'theory acceptance' that we started with is to be captured. One such proposal, which originates in Levi (1967) and has often been defended in the subsequent literature, proposes to model 'theory acceptance' in *decision-theoretic* terms. This implies, however, that we abandon the direct

[49] Emphasis in the original.

connection between acceptance and (degrees of) belief, assumed so far to be relatively unproblematic within the 'threshold account', and reconstrue acceptance in a different way. The most recent fully fledged decision-theoretic account of 'theory acceptance' is due to Maher (1993) and it is on this account that I concentrate here.

2.3.1 Maher's Decision-Theoretic Account

On Maher's decision-theoretic account a theory is accepted when it is *sincerely and intentionally asserted.* Hence, one cannot accept a theory if he is either insincere about his judgement or this judgement is the outcome of a slip of the tongue. This definition, of course, is not required by the decision-theoretic approach. It only reflects the obvious consideration that merely *uttering* – and seeming to assert - a hypothesis does not mean that you do *in fact* accept it. For instance you might be lying. Hence, defining acceptance this way is also compatible with the 'threshold account'.

But under what conditions is one *justified* in accepting a hypothesis? One might both sincerely and intentionally assert a hypothesis but at the same time have no satisfactory normative grounds for doing so. Such a person would surely be *incorrectly* accepting this hypothesis. Maher's claim is that a decision-theoretic approach is superior to the 'threshold account' from a purely normative point of view on two counts. Decision-theory, we are told, not only avoids the paradoxes of the threshold account, but also sheds significant light on the reasons that determine when one *correctly* accepts a hypothesis by taking into account factors, which although crucial for the normative dimension of acceptance, are neglected by the 'threshold account'.

Two main factors are involved when it comes to theory acceptance, says Maher, namely *truth* and *informativeness*. Obviously we are interested in accepting true theories in science, since truth is (standardly) considered the regulative ideal of rational discourse. At the same time, though, we are not after any true hypothesis. We are primarily interested in hypotheses rich in content and information, which therefore allow us to increase our knowledge. Maher's charge against any theory of acceptance which focuses solely on probabilities – echoing Popper's from long ago (1959, 268-270 and 374)- is that such a theory *is bound* to fail adequately to take into account the second factor just considered, i.e. informativeness. By focusing on probabilities we only pay attention to the issue of the degree of warrant for the assertion of the truth of the hypotheses we consider. Yet, this concern automatically offsets progress relative to the problem of finding informative hypotheses. Consequently, Maher points to a trade-off between "the concern to be right (which would lead one to accept hypotheses of high probability) ... and the desire for informative hypotheses (which tends to favour hypotheses of low probability)" (Maher 1990, 388).

Decision-theory, Maher argues, provides a neat way to underwrite such a trade-off. Alongside Bayesian probabilities, which reveal one's attitude towards the truth of a hypothesis, he introduces the notion of *cognitive utility*, which captures the cognitive value to us not only of truth but also of the informativeness of a hypothesis, and is mathematically represented by an appropriately defined *cognitive utility function*[50]. Maher expresses the intuition behind the introduction of the cognitive utility function as follows:

[50] For the mathematical details of how he proposes to define such a function see Maher (1993, ch. 8).

"Other things being equal, acceptance of a true proposition has higher cognitive value than acceptance of a false proposition, and among true propositions the cognitive value is higher the more informative the proposition is" (1992, 154).

Note in this connection the strictly epistemic character of utility - as Maher construes it. Within the decision-theoretic approach 'theory acceptance' retains its strictly epistemic, non-pragmatic dimension, since utilities refer *not* to material gains of any sort but to the intrinsic cognitive value of truth and of information for the purposes of scientific reasoning. When we have to decide, then, which theory to choose among a set of competitors, all we have to do is estimate the *expected cognitive utility* for each theory and accept the one which maximises this magnitude.

Once a well-defined notion of a cognitive utility function has been given, it is an easy task to compute the various expected utilities, and then choose the theory which maximises it. For an arbitrary theory H, then, the expected cognitive utility of accepting it would be given by

$$\mathrm{E}\,U_H = \sum_{x \in X} \Pr(x) \times U(H, x),$$

where U stands for *cognitive utility*, $\mathrm{E}\,U$ for *expected cognitive utility* and x describes any of the possible *cognitive consequences* of accepting H belonging to some set X. A simple suggestion, due to Hempel (1960) and Levi (1967), defines 'cognitive consequences' as the consequences of accepting a theory relative to its truth-value. Hence, on this approach X would contain two elements, i.e. the states 'accepting H when it is true' and 'accepting H when it is false'. Consequently, $U(H, x)$ expresses the cognitive utility of accepting H in each of the possible states (in this simple case the states that H is true or false), $\Pr(x)$ stands for the probability that these states obtain and the expected cognitive utility of accepting *H* is just the probability-weighted sum of the utilities of the possible consequences (or states) (cf. Maher 1993, 141-143).

Maher, however, takes a more ambitious course and introduces the notion of closeness-to-the-truth (or *verisimilitude*) into the discussion. This is because Maher wants "the utility of accepting a hypothesis [to] depend [not only] on whether the hypothesis is true or false, but also, in case the hypothesis is false, ... [on] how far from the truth the hypothesis is" (ibid. 142). In other words, Maher is not satisfied with the simple suggestion to take as the only possible 'cognitive consequences' the states 'accepting *H* when it is true' and 'accepting *H* when it is false'. Instead, he is after a richer account which allows for more 'cognitive consequences' than these two. Those 'extra' consequences will be determined on the basis of how close to the truth our theories are, on the assumption that we accept them while they are false. To be sure, not much changes with respect to calculating the expected cognitive utility of accepting a hypothesis. The only difference is that on Maher's enriched suggestion, the probability-weighted sum of utilities ranges over the consequence 'accepting *H* when it is true' and those cognitive consequences that will arise from the partition of 'accepting *H* when it is false' on the basis of *H*'s distance from the truth[51].

Maher's allusion to verisimilitude, nonetheless, is quite surprising, since the task of determining anything resembling a formal objective measure of the distance

[51] Note that the introduction of verisimilitude implies that the partition of the consequence 'accepting *H* when it is false' is, in principle at least, non-denumerable. It does no harm, however, to present the situation as denumerable, since all real-life theory comparisons should be given a reasonable answer even under this simplifying assumption.

from truth has often been considered to be virtually impossible. Even if we accept that the concept of verisimilitude is pre-theoretically transparent, there still remain serious problems with respect to a) making it precise within the context of a coherent theory, and b) doing so in a way that allows at any instant the determination of a given theory's degree of verisimilitude. In fact, ever since Popper's seminal 1963 paper in *Conjectures and Refutations*, there has hardly been any remotely satisfactory answer to either (a) or (b)[52].

Popper (1963) construed verisimilitude as approximation to the *complete truth*. Hence, a theory would be *maximally verisimilar* if and only if it were "not only true, but completely comprehensively true: if it [corresponded] to *all* facts, as it were, and, of course, only to *real* facts" (ibid. 317-318)[53]. Popper, of course, was much more interested in the *comparison* of theories in terms of their degree of verisimilitude rather than the conditions under which any theory could be deemed maximally verisimilar. This is because his ultimate purpose was to be able to say that, for two *refuted* theories, one can still be more verisimilar than the other. In this way, he hoped to substantiate the claim that later theories approximate the truth better than earlier ones and, hence, that, through theory-change, scientific knowledge approaches more and more closely to the truth (ibid. 313, 318). To this end, however, the specification of an *objective measure* of verisimilitude is required. Unfortunately, Popper's own proposed measure in terms of the truth and falsity-content of theories had been shown to be empty long ago by Miller (1974) and Tichý (1974), who were able to prove that, on Popper's characterisation, *any*

[52] See Maher (1993, 220-224) for a brief survey of the literature on verisimilitude and further references.

[53] Emphasis in the original.

two false theories have the same verisimilitude. And, as noted a little earlier, no satisfactory objective measure has been proposed so far.

Maher does not deny the force of this problem and that is why he does *not* seek an objective measure of verisimilitude. Instead he proposes a *subjective* definition, which links verisimilitude with human interests and values. Maher's proposal is quite idiosyncratic and rests on the postulate that "a measure of verisimilitude is essentially a measure of the cognitive utility of accepting a hypothesis under different circumstances" (Maher 1993, 226). Since we already possess well-defined cognitive utility functions, which express the various cognitive values scientists take into account when facing a problem of theory-choice, the obvious next step is to use these functions in order to obtain a measure of verisimilitude. People's cognitive utility functions, of course, are bound to differ, since their values and interests need not be identical. Hence, Maher must accept that "verisimilitude is relative to persons" (ibid. 227). The fact that he is only interested in a subjective characterisation, though, makes him think that "we do not need there to be one unique verisimilitude function in order for the notion of verisimilitude to be meaningful" (ibid. 227). All that is required is that the cognitive utility functions we start with are well-defined.

Maher (ibid. 227-231) goes on to measure the verisimilitude of a theory H when state x is the true state, denoted by $v_x(H)$, via a normalised cognitive utility function so that the verisimilitude of a tautology is 0, while that of the complete true theory is 1[54]. When it comes to a problem of theory-choice, then, we can also

[54] Maher's proposed measure of the verisimilitude of a theory H is $v_x(H) = \dfrac{U(H,x) - U_X}{U_T - U_X}$,

where $U(H, x)$ is the cognitive utility of accepting H at state x, U_T is the cognitive utility of the compute the expected verisimilitude of a theory in a similar way to expected cognitive utility, i.e. calculate the quantity

$$\mathrm{E}\,v(H)= \sum_{x\in X}\Pr(x)\times v_x(H)\,.$$

The expected verisimilitude of H (denoted by $\mathrm{E}\,v(H)$) will be the probability-weighted sum of H's verisimilitude in each of the possible consequences (or states). Since verisimilitude is just normalised cognitive utility, it also follows that whenever we maximise expected cognitive utility we also maximise expected verisimilitude.

One might object at this point that the phrase 'a subjective definition of verisimilitude' is a contradiction in terms – analogous to the obvious contradiction in the phrase 'a subjective definition of truth'. The sole reason why we might try to produce an acceptable characterisation of verisimilitude is because we (think we) know that we can't have the complete truth, yet we might feel reasonably confident that we can't be "very far" from it. In other words, we want a measure of how close to the *objective* truth we think we are. Determining this measure on the basis of our agents' interests, though, seems to be entirely irrelevant to achieving this target. Hence, resorting to subjectivity in this respect simply because we tend to think that any attempt to objectively define a measure for verisimilitude is bound to fail is really no different from admitting defeat and, with it, the impossibility of any such measure. In fact, even ardent realists, like Psillos (1999, 276-279), have abandoned this task, treating the notion of 'approximate truth' instead as something like an undefined but still intuitively acceptable primitive concept.

---

completely true theory and U_X the cognitive utility of a tautology. If H is the completely true theory, $U(H, x) = U_T$ and $v_x(H) = 1$, while if H is a tautology, $U(H, x) = U_X$ and $v_x(H) = 0$ (cf. 1993, 228-229).

Let's, despite this objection, grant Maher his definition of verisimilitude for the sake of argument - since such an admission will not affect my central point. Nonetheless, this kind of objection invites a more serious worry for our purposes, namely that the decision-theoretic construction has abandoned the traditional understanding of realism for a more pragmatically-oriented attitude. Indeed, following the example of Levi (1967), most decision-theorists tend to be pragmatist in their general philosophical outlook. This would seem enough to disqualify them from entering the standard debate on scientific realism, since the concept of 'truth' employed in it is one pragmatism disowns. Maher, nonetheless, is adamant that "the notion of truth figures centrally in my account of scientific values" (1993, 208). What is less clear is what conception of truth he adopts.

In his 1993 book (ibid. 208-209) Maher argues that the so-called *transparency property* of truth (also sometimes known as the 'redundancy theory'), which suggests that the statement 'It is true that *p*' means the same as '*p*', suffices for the explication of the notion on the grounds that "since the meaning of the latter sentence [i.e. '*p*'] is unproblematic (if anything is), so is the meaning of the former [i.e. 'It is true that *p*']" (ibid. 209). It might be thought that Maher advocates here some form of a minimalist account of truth, in line with his pragmatist predecessors. Although such a reading is not unreasonable, it is far from being unambiguous. In fact, one can extract a markedly different picture from Maher's peculiar understanding of philosophy of science itself, which brings him much closer to the mainstream assumptions entertained within the standard debates on realism.

Maher's main focus is on scientific values. By 'scientific values' he refers to the epistemic goals science presupposes. Maher, however, construes 'scientific values' in a rather surprising way. Taking his cue from van Fraassen, he writes:

> "Is the issue [about scientific values] a descriptive one about the values scientists actually have? Or is it a normative one, about the values scientists ought to have? It is neither of these. Van Fraassen is answering the question *What is science?...*, and so his claims about scientific values are claims about what values a person must have to be a scientist" (ibid. 241)[55].

But a number of objections can be raised to this. First, the issue of 'what values a person *must* have to be a scientist' is a normative one. Hence, Maher's assertion that it is neither normative nor descriptive is very puzzling[56]. On top of this, van Fraassen's own construal of the issue is, in the end, also normative (and, in fact, much richer than the mere question of defining 'which values a person must have to be a scientist'). After a few rather ambiguous introductory remarks in his (1993, 189), van Fraassen specifies a scientific theory as "an object for epistemic or at least doxastic attitudes – the attitudes expressed in assertions of knowledge and opinion" (ibid. 190). The investigation of the status of our doxastic attitudes towards scientific theories, however, is by definition a normative endeavour.

Maher's main protest, then, is that both realists and anti-realists have failed to appreciate the import of further factors (or 'values', in his terminology) such as informativeness, besides the standard-ones like truth and empirical adequacy, for

---

[55] Emphasis in the original.

[56] In a footnote, Maher attributes to van Fraassen himself the view that the issue regarding scientific values is neither normative nor descriptive (1993, 241, fn. 18). Even if this is a faithful representation of what van Fraassen actually said, though, one is still left wondering about the nature of the debate on van Fraassen's own characterisation.

theory-choice. He claims that his own decision-theoretic account fares better than standard methodologies by explicitly incorporating all the relevant values through the introduction of the concept of 'cognitive utility'. If that is so, though, and in the absence of further illumination from Maher, it seems difficult to see how he could object to the common-sensical correspondence account of truth, assumed among others by van Fraassen himself. In other words, what seems to matter most to him is not what 'truth' is about but, rather, what else besides truth-*standardly-understood* is important for theory-choice. I conclude, then, that the main difference between Maher's decision-theoretic account and the standard approaches is not so much that they employ two different conceptions of truth but, rather, that they disagree on what are the relevant epistemic values for theory-choice.

It follows that Maher's decision-theoretic approach is consistent with the basic realist assumptions regarding the notion of 'truth'. Let's, then, see how it can be used by the Bayesian to counter the realist's charge. The decision-theoretic approach denies that a Bayesian can only accept a hypothesis if it has probability one. Maher claims that "Bayesians can admit that there are more things in this world than subjective probabilities" (Maher 1992, 154). By admitting the existence of epistemic utilities, they explicitly deny that probability is all that matters for 'theory acceptance'. Acceptance of one theory among many competitors merely requires that that theory maximises expected cognitive utility, and this is, of course, quite compatible with that theory's probability being significantly lower than 1. Hence, the assumption that a realist always accepts his theories[57] does *not* commit

[57] It might still be rather unclear what a proper characterisation of the realist position might look like on the decision-theoretic account. Given the central role 'theory acceptance' occupies in this context, the correct formulation is, I think, the following: a decision-theorist will be taking a realist

him to assigning to them probability 1 – if the decision-theoretic account is adopted.

Decision-theory, then, *appears* to offer a way out of the charge that Bayesians cannot accommodate within their analytic framework a notion of 'theory acceptance' that is appropriate for the analysis of the realist position. I think, however, that the decision-theoretic way of reconciling realism and Bayesianism is unacceptable. This is because decision-theory fails to provide a persuasive, or even a coherent, account of the trade-off between informativeness and probability on which it centrally relies.

2.3.2 The Benign Trade-off Between Informativeness and Probability

The only *uncontroversial* case of a trade-off between informativeness and probability is when a theory T *entails* some other theory T'. When this happens, T's logical content (or informativeness) is greater than that of T'. On the other hand, it is of course a theorem of the probability calculus that if $T \models T'$, then $\Pr(T) \leq \Pr(T')$; moreover, the equality holds if either $\Pr(T/T') = 1$, for example when T' logically entails T, or if $\Pr(T') = 0$. Thus, in this special case of entailment, probability and content will generally pull in different directions.

But does this pose special problems for a purely probabilistic approach to theory change that require radical reshaping of the approach and, perhaps, the introduction of a utility function? Let's look at actual science for a case where a theory T entails some other theory T'. A readily available way for the creation of

stance towards his theories if and only if he a) already accepts a theory on the basis of utility-maximisation, and b) assigns to this theory probability greater than 0.5 (more details as to how condition (b) should be construed are contained in section 2.5.1).

such a situation is to isolate parts of general theoretical frameworks and thus create deducible sub-theories. For example, one might consider Newton's theory of *planetary* motion, which is entailed by the general framework of Newtonian Mechanics (NM). It might, then, be argued that although NM is less probable than the sub-theory referring to planetary motion, we still opt for the general framework on the basis of its informativeness.

It is of course true that the probability of NM is lower than that of the sub-theory describing planetary motions in particular; but it is also *false* that these two theories have ever been (or ought to be) considered as genuine rivals. In fact, speaking of *two* theories in this case is already counter-intuitive. It seems more appropriate to describe this situation as a *generalisation* of one or more basic theoretical principles to a larger domain of phenomena, or, in case the principles were employed implicitly, as the extraction of a general principle from what is in reality an instance of its application. Even if we leave aside issues regarding the individuation of theories, though, and take NM and its planetary sub-theory as distinct, it is still far-fetched to suggest that there was ever (or that there ought to be) a question of choosing the one over the other. Instead, the real question is whether one has good reasons to accept *both* NM and its sub-theory or *only* the sub-theory. Obviously the probability of the special case will be at least as great as the probability of the generalised framework, but this is to be expected by pure logic alone. In no way, though, does this constitute an instance of genuine theory choice between proper rival theories. Whenever a theory T entails T', it fails to be a genuine rival to T'.

What happens, then, when it comes to proper rivals, in which the history of science abounds? The replacement of Galileo's and Kepler's laws by the more general NM is one example of genuine theory choice between rival theories. At the same time, it has also been taken to exemplify successful inter-theory reduction *à la* Nagel, which means that it ought to satisfy Nagel's *condition of derivability*[58]. This states that for reduction to take place, the newer (or more fundamental) theory *T* must *entail* the old (or less fundamental) theory *T'* it replaces so that all the laws of *T'* are deductively derivable from *T*. Hence, we seem to have a counter-example to my claim that it is impossible for a theory *T* to entail *T'* and also genuinely rival it.

Appearances to the contrary, this claim is not correct. For as we ought to have known since Duhem's classic treatment (1906, 192-193), the replacement of Kepler's and Galileo's laws by NM in fact fails to satisfy the condition of derivability. Kepler's first law, for example, says that any planet moves on an elliptical orbit around the sun. NM, on the other hand, says that any planet *would* move on an ellipse on the *counterfactual* supposition that the only bodies in the universe were the sun and that single planet. By taking into account the mutual attraction between the planet under consideration and the rest of the planets, however, NM predicts that planets move on a 'perturbed' ellipse, thus *formally contradicting* Kepler's first law. Of course, owing to the massive size of the sun relative to the rest of the planets, NM entails Kepler's law as a *first approximation*, which, nonetheless, fails to render them *logically* consistent. The situation is similar regarding the relation between NM and Galileo's laws. Galileo's theory assumes constancy of vertical accelerations over any fall from a point above the surface of

[58] The other condition Nagel thought necessary for successful reduction was the *condition of connectability*. This dictated that in case there were terms occurring only in the less fundamental theory, then some assumptions or "bridge laws" were required to connect them with the terms of the more fundamental theory in order for the logical derivation to take place. See Nagel (1961, 352-254).

the earth while NM denies this proposition. Acceleration in NM is a function of the distance between the centre of mass of the body and the centre of mass of the earth, and this distance of course changes during the course of the fall. However the distance between the earth's surface and its centre of gravity is so great compared to the length of the fall that the change is very small; consequently, the difference between Galileo's prediction and Newton's is also very small and, in fact, well within observational error. But, of course, a small distance is different from no distance at all. Strictly speaking the two theories conflict and, again, only an approximate relation holds between them (cf. Feyerabend 1962, 57-59; Popper 1983, 142)[59].

The significance of these well-known results should not be underemphasised. They establish that existing genuine instances of rival theories fail to satisfy the condition of derivability, and hence, that they are no counter-examples to my previous claim that when a theory T entails T', it does not genuinely rival T'. We may conclude, then, that in the absence of a more convincing counter-example, the Bayesian has not been shown to face a trade-off between content and probability in cases of *genuine* theory-choice.

One might try to contest this conclusion by noting that it is at least plausible to maintain that, say, NM is much more informative than either Kepler's theory or Galileo's despite the absence of strict entailment relations between them. Although one can derive approximate versions of Kepler's and Galileo's specific theories from NM, the converse does not hold. Hence, although not strictly derivable from NM, Kepler's and Galileo's theories still get very close to being so derivable,

[59] Similar approximate relations can also be derived between NM and special relativity, for speeds strictly smaller than the speed of light.

which in turn would seem to allow us assert (no doubt rather loosely) that NM is much more informative than either of the two.

Even if we grant that NM is much more informative than either Kepler's or Galileo's theory, however, this by itself does not imply that there is a trade-off between content and probability for these theories. Put more generally, for two theories standing in a relation of *approximate entailment* it stills seems plausible to argue that it is at least possible to assert both a) that they are comparable content-wise and at the same time b) that there is no trade-off between content and probability in this case. One reason for this possibility might be sought in our habit of employing an intuitive and imprecise notion of a theory's content alongside the stricter one, which applies only in cases of strict deductive entailment. We seem inclined to make comparisons of content even when, strictly speaking, no such relation can be unambiguously asserted. It's no wonder that such an imprecise notion of content can be made to fit with the equally imprecise notion of approximate entailment. The concept of probability, on the contrary, is nothing like that. Being as rigorously formalised and explicated as it is, it remains mysterious how it is supposed to be influenced by the assertion of relations as vague as 'approximate entailment'.

Still, the possibility of approximating both Kepler's and Galileo's theories from the standpoint of Newtonian theory might give rise to the following possible objection. Those approximate Newtonian sub-theories follow logically from Newton's theory, and hence, are less informative and more probable than the general Newtonian framework. One might be tempted to propose, then, that instead of contrasting the original Galilean and Keplerian formulations with Newtonian theory, what one should do is contrast those approximate Newtonian subconstructions with the general Newtonian framework. If one does this, the trade-off between informativeness and probability re-enters the scene and appears to become a problem again.

The reason the problem is only apparent, though, is not hard to see. What we have done here is traded an instance of genuine theory-choice with one that is *not* genuine. We started with the original contraposition of Keplerian and Galilean theory with NM. These theories genuinely rivalled NM but were irreducible to it. We then substituted Newtonian versions of both Galileo's and Kepler's theories for the original formulations and proclaimed them to be the genuine rivals of NM. This proclamation, nonetheless, is illegitimate. Those Newtonian sub-formulations were *explicitly constructed* on the basis of Newtonian principles with the aim to approximate the original theories. It is surely highly implausible to consider as genuine rivals two theories, one of which is NM and the other one explicitly constructed on the basis of NM itself!

In reality, this objection is just another instance of the general strategy of isolating parts of general theoretical frameworks and thus creating sub-theories entailed by the more general constructions. I examined this strategy early in this section and concluded that it does not give rise to genuine cases of theory-choice. Not surprisingly, the same holds for the Newtonian variants of the Galilean and Keplerian theories. Hence, the main point of this section, i.e. that it is impossible for a theory *T* to entail another theory *T'* and genuinely rival it, stands intact.

Finally, it is worth pointing out that this strategy of sub-theory creation might not even be available for all cases of theory-change. This is especially the case whenever new theories bring with them sharp changes in ontology by introducing novel theoretical entities and relations radically at odds with the ones postulated by the previously dominant theories. In such cases it is very hard to come up with sub-theories, which can be both properly derived from the general framework of the new theory and, at the same time, meaningfully said to approximate the theory they replace. A typical example of this situation, due to Nickles (1973)[60], is the reduction of the Special Theory of Relativity (STR) to NM. The standard way to approximate NM from within STR is to take the limit of certain STR equations as the physical system velocity tends to zero. Yet this is not a case of proper logical derivation of a theory approximating NM from within STR. Strictly speaking, all that STR entails is "conditional statements like '*If* system velocity is negligible relative to the velocity of light, *then*...'"[61] (ibid. 196, fn. 28) rather than a special relativistic variant approximating NM.

The situation is even more striking, if one considers the relation between Quantum Mechanics (QM) and NM with respect to Ehrenfest's theorem. This theorem states roughly that any relation appearing in NM must be valid as a relation between quantum-mechanical expectation values. In other words, wherever NM postulates relations between variables, QM postulates relations between the expectation values of those variables As Nickles (ibid. 196-197) rightly notes, though, this amounts to saying that QM employs a formal calculus which is not shared by NM. Hence, although NM and QM are clearly *somehow* related, it would surely be far-fetched to suggest that through Ehrenfest's theorem one can derive an approximate quantum-mechanical variant of NM, which nonetheless makes use of

[60] Nickles discusses these examples in the course of his argument against Schaffner's (1967) model of inter-theory reduction. His conclusions, nonetheless, are highly relevant for our purposes, since Schaffner's approach crucially rests on the existence of a corrected version of the reduced theory, which *both* approximates the reduced theory and is also logically derivable from the reducing one.

[61] Emphasis in the original.

an altogether different formal calculus from NM itself! These examples illustrate, then, that what appeared to be a ready-made recipe for relating content with probability (i.e. the creation of sub-theories approximating the historical rivals) not only fails in principle, but is also far more difficult to implement in all actual cases of theory-change than initially thought.

We can safely conclude, then, that considerations of content and probability *for genuine cases of rival theories* are logically independent. The absence of a negative relation between them, assumed by decision-theory, shows that the adoption of the decision-theoretic framework is in fact ill-motivated. Since seeking theories with high probability does *not* disqualify one from getting theories with high content, there is no basis on which to motivate the introduction of the notion of epistemic utility alongside that of probability. All we need in order to understand 'theory-choice' is a measure of epistemic security, which, as I argue in the next section, is already available to us in the form of the probability calculus.

2.4 Explaining 'Acceptance' Away

The realist challenge to Bayesianism presented earlier can be reconstructed more formally as follows:

1. The realist offers a *deductively invalid* inductive argument for the approximate truth of our best scientific theories[62].

[62] The term 'inductive' here refers to the realist No-Miracles argument *qua* an ampliative inference rather than to enumerative induction or any other 'inductive rule'. The precise content of this argument will be discussed in detail in later chapters. For the purposes of this chapter it suffices to note that the realist explicitly claims his argument to be ampliative and, hence, that its conclusion does not follow deductively from its premises.

2. This argument allows him to 'accept' these theories on the grounds that it affords them strong inductive support.
3. 'Accepting' the conclusion of an inductive argument is justified if and only if the argument is inductively strong.
4. Hence, the realist *justifiably* 'accepts' his best theories.

At the same time, though, we know that:

5. These theories should be assigned probability strictly less than 1.
6. Furthermore, the Bayesian can uncontroversially 'accept' a hypothesis if and only if it is assigned probability 1.
7. Therefore, the Bayesian lacks the resources to accommodate the realist's *justified* epistemic attitude towards his theories.

Premises (1) and (2) merely articulate what a realist expressly holds, assuming for purposes of discussion that his argument is inductively strong. Premise (3) is implicit in the realist's argument, if his decision to 'accept' his fundamental theories is to be justified at all. Premise (5) is also eminently reasonable. Indeed, no serious scientific theory should ever be assigned probability 1, whatever the evidence for it might be, given the extent to which it transcends the phenomena and its indirect link to them through exceedingly complex ways of experimenting. Finally, premise (6) is supported by the failure of both the 'threshold' and the decision-theoretic account to yield a viable Bayesian theory of 'acceptance' for levels of probability less than 1, as argued for in the previous sections.

In contrast to the 'threshold account' and to the various decision-theoretic approaches, which seek to modify premise (6), the reason I believe that the above argument is unsound is because premise (3) seems clearly false. If so, then there is no basis for (4) and the realist is *not justified* in 'accepting' his best scientific theories. Note, though, that this conclusion does *not* in turn imply that one can never be justified in being a realist. Quite to the contrary, in what follows I argue that a Bayesian approach explains *both* why a realist should never accept his theories *and* why one can still perfectly reasonably claim to be a realist *vis à vis* his best scientific theories. If I am right, then Bayesian Confirmation Theory is not only able to capture fully the essentials of the realist position, it also further elucidates it and becomes a proper means of expressing it.

To start with, we need to identify the relevant aspect of the realist position which allegedly is not captured by the Bayesian view. The reason why the realist asserts that his arguments legitimise detachment of their conclusion and, hence, 'acceptance', is because he wants to claim to possess *knowledge*. Recall his main concern: that of establishing the *epistemic* claim that we have good reasons to believe that our best scientific theories are well-confirmed and (at least) approximately true. We know that when it comes to deductive arguments starting from entirely uncontroversial premises (if there are ever any), accepting (or detaching or asserting) their conclusion is a simple consequence of their validity. For all we know, though, nothing similar holds for the inductive case. The realist's conclusion does not follow from his premises, no matter how uncontroversial they are. Can we detach the conclusion of an (deductively invalid) inductive argument, even if we concede that the argument is inductively strong?

When examining the lottery paradox, I mentioned that we deal with essentially two conceptually distinct notions of 'acceptance'. One of them, I called it *acceptance*$_1$, refers to fully-fledged acceptance, i.e. acceptance when the probability we assign to our propositions is 1. This case is not particularly problematic, since we can arguably claim knowledge regarding such propositions, if we can ever claim knowledge at all. Examples of such propositions are all standard logical truths and (perhaps) some other standardly accepted propositions, e.g. reports of observations collected via our unaided senses, which, although non-trivial, reflect our fundamental epistemological presuppositions. To be sure, endorsing those presuppositions goes beyond subjectivism. This is because we are now in effect asserting that we can obtain objective knowledge of certain non-trivial propositions. The justification of those presuppositions, however, need not occupy us at this point, since they were only intended as an illustration of what is involved in cases of fully-fledged acceptance. The other notion, *acceptance*$_2$, refers to propositions which are assigned lower probability, but still quite high. In this second case, I maintained, we don't employ the same concept. We want to make a qualified statement, which reflects both our confidence in the (approximate) truth of the statement under consideration as well as our recognition that we do *not* have conclusive grounds for it.

The realist's claim is just one such statement. If asked about his epistemic stance regarding his best theories, the realist would *not* respond by saying that he is *sure* that they are true, *nor* even that he is *sure* that they are approximately true. In failing to claim certainty, he also fails to accept his theories in the sense of *acceptance*$_1$. What he would say instead is that he has *reasons to believe* that current theories are approximately true and that is why he only accepts his theories in the sense of *acceptance*$_2$. But, then, what does 'having reasons to believe' mean? What comes naturally (and, I would argue, correctly) to mind when faced with this question is: 'they are probably approximately true' or 'it is very likely that they are approximately true'. In other words, the realist's 'reasons to believe' in the approximate truth of our theories (as well as his grounds for *acceptance*$_2$) are summed up in no more than a probability statement.

A tempting rejoinder for a defender of *acceptance*$_2$ would be to attempt to explicate this notion by alluding to the idea that inductive arguments allow us to 'accept' only *defeasible* conclusions. Psillos, for example, drawing from Pollock's (1986) work, points out that "unlike deductive methods, ampliative methods are *defeasible*. The issue here is ... that further information can remove the warrant for holding the output of the method" (2002, 609). He readily admits, of course, that the defeasible character of ampliative inferences provides only "*prima facie warrant* for an output (belief)" (ibid. 609). Rather than this becoming a reason for refusing to 'hold' or 'accept' the conclusions of ampliative inferences, however, Psillos praises Pollock for "rightly [stressing that] to call a warrant (or a reason) prima facie is not to degrade it, *qua* warrant or reason" (ibid. 609)[63]. In short, the claim seems to be that, by *accepting*$_2$ that our scientific theories are approximately true, we acknowledge that our reasons for 'acceptance' might be removed in due course, yet we think that at present there is no evidence that should make us suspect that this possibility is real.

I still think, however, that this is not an acceptable solution. In fact, alluding to the 'defeasibility' of ampliative inferences further highlights why notions like *acceptance*$_2$ are totally redundant and reducible to a mere probability statement. Indeed, what else does an assignment of a high probability but less than 1 to a given statement express except the recognition of the fact that, despite our present confidence, this statement might turn out to be false, i.e. be *defeated*? Clearly, if there is a feature of ampliative reasoning that a purely probabilistic approach can

[63] Emphases in the original.

express most adequately, then this is its defeasible character. This is because a probabilistic theory incorporates at the outset the salient feature of defeasibility, namely indefiniteness or uncertainty of opinion.

We see, then, that the realist's 'reasons for belief' and the ensuing *acceptance* $_2$ eventually boil down to a probability assignment alone, offering no further information whatsoever about one's epistemic situation. The only thing the realist can achieve by maintaining the cover term 'acceptance' is to convey the false impression that we have somehow got away from the contingent character of the realist claim. And that, as a consequence, we are all of a sudden able to claim some sort of epistemic security for our hypotheses superior to what a mere probability assignment offers. Yet nothing of this sort can happen - and this for logical reasons alone. It follows that probability statements are not only suitable for expressing the realist epistemic situation, they are also mandatory in that they alone express transparently the contingent (and non-certain) character of the realist thesis.

This approach is not new, of course. It was proposed by the master of conceptual elucidation, Rudolf Carnap, in his (1968a). Carnap proposed to apply his notion of *explication* – originally introduced in his *Logical Foundations of Probability* (1950) for the purpose of analysing the concept of 'confirmation' - also to the notion of 'theory acceptance'. By 'explication' Carnap understood "the transformation of an inexact, pre-scientific concept, the explicandum, into an exact concept, the explicatum" (1950, 1). Concepts, he maintained, are typically ranked in ascending degree of accuracy into classificatory, comparative and quantitative ones. Classificatory concepts are typical of a pre-scientific mode of thought, according to which things are assigned only a few mutually exclusive properties. This makes the use of these concepts easy for the layman but at the same time disqualifies them from supplying the fine grained distinctions necessary for the expression of more complicated ideas. Quantitative concepts, on the contrary, are the exact opposite. Despite requiring technical skill, they signify the attainment of a level of conceptual sophistication typically associated with scientific (and, one might add, philosophical) advancement.

According to Carnap, the whole debate about 'theory acceptance' is "not so much a question of right or wrong, but rather of primitive or more refined procedures...[I]n particular, [the problem situation regarding acceptance rules] appears to me [i.e. Carnap] to be a special case of the distinction between three kinds of concepts, namely, classificatory, comparative, and quantitative concepts" (1968a, 147). The concept of 'theory acceptance' is just a typical example of a classificatory concept. Talking in terms of acceptance (in the sense of *acceptance*$_2$, of course) is just one way to characterise one's epistemic disposition towards a theory, yet it is a rather primitive way. Adding a numerical scale makes the characterisation of that very same disposition more detailed and clear. The new and more powerful conceptual tool, then, correctly takes the place of the old and less refined one, allowing one to express his epistemic attitude more accurately and transparently. Sticking with the old conceptual machinery is just an attempt to preserve the psychologically comforting connection with the absolute notions that the realist (rightly) aims at and these primitive concepts rely on, thus creating the illusion that these ideals somehow become easier to achieve if one prefers one linguistic idiom over the other.

I conclude, then, that the realist should never accept his theories. He should not do so in the sense of *acceptance*$_1$, because this sense does not convey the meaning of the thesis he wishes to defend; equally, he should not accept his theories in the sense of *acceptance*$_2$, because employing this terminology offers nothing over and above the probabilistic assignments and at the same time merely obscures the issues involved. Hence, pure Bayesianism both avoids our initial problem and at the same time emerges as a proper standpoint from which to reconstruct the debate about realism.

Finally, let me emphasize once more that this construal of 'theory acceptance' is related exclusively to the epistemic issues debated within normative philosophy of science. It is not an attempt to regulate the everyday use of terms; it is merely a view according to which using seemingly innocuous terms in a philosophical context may create pseudo-problems and unjustifiably prejudice one against otherwise legitimate theoretical approaches like Bayesian Confirmation Theory.

2.5 Modeling Realism and Competitors Once More

How, then, should we model realism and the various alternative positions in purely probabilistic terms, now that we have found good reason to abstain from using any acceptance-related terminology?

2.5.1 Realism

The realist's principal focus is on the *epistemic* claim that we are justified in believing in the approximate truth of our theories. I have just argued that the realist claim naturally translates into the proposition 'we have good reasons to think that our best theories are probably approximately true' and that is why one ought to reduce the *meaning* of the realist thesis to some probability statement(s). But what does it mean, in more quantitative terms, to hold that 'we have good reasons to think that our best theories are probably approximately true'? The natural (and most popular[64]) answer to this question seems to be that, in principle, thinking that we are justified in assigning to our theories any probability more than 0.5 is both a necessary and a sufficient condition for someone to count as a scientific realist. This is because that proposal amounts to saying the obvious, i.e. that a realist thinks that he has good reasons to believe that our current theories *are more likely than not to be approximately true*. And when one thinks that some hypothesis is more likely than not to be true, one can assert, as the realist does, that his theories are 'probably approximately true'.

There are two interrelated issues here. First, note the importance the realist must assign to his *having good reasons to think* that our best scientific theories deserve probability greater than 0.5. Indeed, the epistemic thesis that *we are justified* in believing in the approximate truth of our best theories is an integral component of the realist stance, as defined at the beginning of this chapter (pp. 65-66). Hence, the requirement that the probabilist asserts explicitly that there are certain good reasons which induce his probabilistic assignments is intended to safeguard the normative import of the realist position on its probabilistic understanding. It is not enough merely as a subjective matter to assign a probability higher than 0.5 to some current theory in successful mature science. In fact, one might add that it is not even enough to allude to 'reasons' that are not included in the set of factors, which are taken to be evidentially relevant to a judgement of truth for a scientific theory from the perspective of the modern, empirically oriented, thinker. Examples of such 'reasons' would be divine revelation or the allusion to the opinion of some one authority, e.g. the Pope. As a realist, one must hold that

[64] See, for example, Dorling (1992, 366ff.), Howson (2000, 201) and Magnus and Callender (2004, 330).

one has *good* reason for making that probability assignment, where 'good reason' must of necessity involve those factors that are thought epistemically significant for a judgement of truth from the standpoint of the empirically-minded, scientifically oriented modern thinker.

Secondly, it follows from the realists' insistence on having good reasons for the adoption of their epistemic stance that their probabilistic assignments should be supported by strong philosophical arguments for their correctness. This situation not only takes us beyond subjectivism from the outset but it also sets an, admittedly, unusually strong requirement for having degrees of belief in the truth of certain propositions. I think it reflects, nonetheless, quite accurately how epistemically demanding and interesting the realist claim is. The realist does not wish to issue merely a tentative statement regarding the (approximate) truth of our best theories. He tries instead to present us with a strong normative case, which serves as the foundation for his epistemic stance[65]. Consequently, it seems that the emphasis on justification that realism presupposes accounts quite naturally for the

---

[65] The *philosophical* nature of the realist's argument in favour of his probabilistic assignments (whose content is analysed in considerable detail in the chapters to follow) is also helpful in avoiding a potential infinite regress problem, which might seem to arise out of my probabilistic characterisation of the realist position. Indeed, if the realist's probabilistic assignments are based on our meta-beliefs concerning their justification, don't we need further arguments which would justify those meta-beliefs of ours, and so on and so forth *ad infinitum*? I think the solution to this problem is provided by the fact that the justificatory argument for our first-order beliefs is philosophical, and hence, *foundational* for those beliefs. Like all philosophical arguments, the realist's argument operates by default on the fundamental level of discourse, and hence, breaks a potentially threatening regress at the outset. More details about how such an argument ought to be construed from the point of view of fundamental epistemology are contained in ch. 5.

demanding conditions that must be satisfied if one's degrees of belief are to classify him within the realist camp.

It might readily be accepted that a necessary condition for realism is that we think our probabilities justifiably exceed the threshold of 0.5. It will surely *not* be accepted without protest that this is also *sufficient* for realism. What of the person who thinks that the General Theory of Relativity (GTR) deserves probability 0.500000001? Is he a realist proper? I admit that attributing to him the realist epistemic stance brings with it an air of paradox. It is a paradox, though, difficult to evade. Let's suppose that we refuse to call him a realist. In this case, someone who is not a realist still believes that GTR is more likely to be true than not. Is this less paradoxical? It seems that for limiting cases like this one, whichever way one goes inevitably brings some kind of counter-intuitive consequence. I suggest that, given that any threshold other than 0.5 is bound to reflect some arbitrary choice, Bayesians should accept the surpassing of this particular value as both a necessary and sufficient condition for endorsing the realist thesis. In this way, they can claim for themselves the following two advantages: a) this criterion is at the same time clear, definite as well as very natural and easily justifiable, b) it allows one to count as a realist about different claims without assigning to each one the same probability value. Surely, the fact that our belief in the existence of everyday middle-sized objects is far stronger than our belief in GTR does not entail that we can not be realists about both of them at the same time. Employing the aforementioned standard readily accommodates the obvious fact that realism itself comes in degrees and is not an all-or-nothing affair. In view of these advantages (especially the first one), we are better off if we bite the bullet and allow our peculiar figure who assigns probability 0.500000001 to GTR to count as a realist.

Before I move on to realism's main competitors, however, two things should be emphasised. First of all, setting the threshold at 0.5 is to be understood as an abstract characterisation of the realist thesis and *not* as an argument that realists today ought to believe in our best theories to a degree only slightly over 0.5. In other words, the degree of belief one ought to have in the approximate truth of our theories is to be determined solely through independent argumentation (of the sort presented in the following chapters) and *not* through an abstract analysis of the *concept* of 'realism' undertaken here. Secondly, the threshold probability 0.5 should not be read as a detachment rule, i.e. a rule which allows us to assert without qualification the theory under consideration, since this would reintroduce the conjunction problem. My purpose here is to give a *reductive analysis* of the realist thesis, i.e. define it in terms of some specific set of probability attributions. According to such a view, then, there is nothing over and above the probability attributions just specified in the realist epistemic claim.

2.5.2 Constructive Empiricism

The main anti-realist contender over the last twenty-five years has undoubtedly been van Fraassen's constructive empiricism. Faithful to the post-positivist intellectual atmosphere, constructive empiricism takes issue with the realists' *epistemic* claim. Van Fraassen is in agreement with the realist as far as the metaphysical and semantic theses of the realist doctrine are concerned. From a metaphysical point of view van Fraassen accepts the existence of a mind-independent world and from a semantic point of view he shares with the realist the belief that "the language of science should be literally construed" (van Fraassen 1980, 10). His disagreement centres on the realist's epistemic claim that scientific theories are reasonably believed to be approximately true. Instead, van Fraassen's empiricist alternative suggests that the empirical success of science to-date entitles us to infer only to the 'empirical adequacy' of our theories. It should be carefully noted that by this latter term van Fraassen refers "to *all* the phenomena; these are not exhausted by those actually observed, nor even by those observed at some time, whether past, present, or future" (ibid. 12) [66]. The proper epistemic attitude towards the *truth* of our theories, and therefore towards its trans-observational components, is, van Fraassen suggests, suspension of belief and agnosticism.

There has been much discussion of how, if at all, van Fraassen's agnostic position can be modelled in Bayesian terms. Van Fraassen's own proposal can be found in his *Laws and Symmetry* (1989):

> "The mistake ... is to assume that agnosticism is represented by a low probability. That confuses lack or suspension of opinion with opinion of a certain sort. To represent agnosticism we must take seriously the vagueness of opinion, and note that it can be totally vague" (193-194).

Thus, van Fraassen explicitly denies that agnosticism is to be identified with low probability. The argument for this claim is only found his (1998, 215-216) and is based on the plausible requirement that agnosticism is closed under negation: if one is agnostic about proposition A, then one is agnostic about $\neg A$. But if we attribute to A probability less than 0.5, then we are forced to attribute to $\neg A$ probability more than 0.5. Yet, this would amount to violating the closure condition, for, as we saw (and van Fraassen concurs), assigning probability more than 0.5 to a proposition does not amount to being agnostic about it. Instead, van Fraassen proposes that the right representation is through a *set* of probability functions, which includes 0 (a

[66] Emphasis in the original.

representor, as he calls it). Consequently, agnosticism is identified not with a low point-valued probability but instead with *vague* probability, i.e. an interval $[0, x]$, where x is some probability value other than 0.

An immediate consequence of van Fraassen's proposal, noted by Hajek (1998) and Monton (1998), is that if we model constructive empiricism through interval-valued probabilities, then an agnostic can never change his mind. Indeed, if one conditionalises on any interval $[0, x]$, vagueness will be preserved since the 0-end of the interval will also be preserved in any new interval resulting from conditionalisation. Both Monton and Hajek claimed that it is an unwanted consequence of this way of modelling constructive empiricism that "it has a life-long claim on its devotees" (Hajek 1998, 201).

It might be objected, however, that this consequence of van Fraassen's proposal is not undesirable at all[67]. From the standpoint of philosophy of science and *a priori* analysis, it is quite natural to maintain that the philosophical stance one will take regarding the status of the observation-transcendent parts of theories is not subject to change through conditionalisation on the evidence. This being so, dogmatism on the agnostic's part need not be particularly troubling with respect to the interval-valued probabilistic representation of agnosticism.

Still, even if we grant this, it does not follow that the interval-valued approach succeeds. The main reason to think not is that it seems that interval-valued probabilities fail to convey what is intended by the concept 'agnosticism'. To see this, consider the following situation. Suppose I ask someone, 'do you believe in God?'. If he said 'there is a 90 per cent chance that God exists', he would

[67] Van Fraassen comes close to giving something like the same answer in his (1998, 215) but does not elaborate.

justifiably be classified as some sort of theist or 'realist' with respect to the unobservable entity 'God'. If he said, however, 'my degree of belief in the existence of God lies in the range from 0 to 0.9', then one should think *not* that this person is agnostic, rather that he is just confused with respect to *his own* beliefs! Most probably, he would describe his own epistemic situation with the phrase 'I don't know what to think', instead of 'I am agnostic about God'.

But is there, then, any probabilistic construal of agnosticism? As I argue below, the answer to this question is affirmative. The agnostic position *can* be faithfully modelled in probabilistic terms. I wish to argue, nonetheless, that the model that works shows that *once the probabilistic idiom is adopted*, the identification between agnosticism and constructive empiricism becomes very misleading. This is not to say, of course, that van Fraassen has been misled from the outset into identifying his theory with agnosticism, since this identification was part of an informal (and, hence, vague) explication of his position. It is only to suggest that it is much more fruitful to leave behind this identification when one already operates within the context of the quantitative, probabilistic language. To see this, let's go back to the essentials of the constructive empiricist position.

Constructive empiricists assert that scientific theories are to be interpreted literally. This means that they can freely assign probability values to their truth, unlike instrumentalists, who don't think that scientific theoretical talk admits of truth-values. Furthermore, constructive empiricists deny the realist claim. Assuming that the construal of realism in the previous section is the right one, they don't assign probability greater than 0.5 to our theories. If they did, then they would rightly be called realists. I have also argued that they do not assign interval-valued probabilities either, since constructive empiricists do *not* want to say that we are confused regarding the issue of realism. Instead, they want to say that they can understand the realist claim, that they don't share the realist's optimism *but* at the same time that *they can't exclude the possibility that our theories are in fact true*. Indeed, this is what suspending judgement means.

Agnosticism, on the other hand, can be informally understood as saying that there is no reason to believe one thing or the other in the face of evidence. Hence, being agnostic about a scientific theory arguably amounts to assigning probability 0.5 *exactly* to the claim that the theory in question is true. In this case, it's safe to say that the reasons for the truth or falsity of the theory are in perfect balance. Granted, a sharp probability value equal to 0.5 makes it also possible to express all the main elements of the constructive empiricist's position. But, of course, or so I claim, we can also faithfully express the fundamentals of the constructive empiricist stance when we assign *determinate* probabilities to our theories lying in the interval (0, 0.5][68]. By assigning a value of, say, 0.3 one expresses disbelief in the realist claim but at the same time acknowledges the possibility that the theory is in fact true. This possibility leads to what can be informally characterised as 'suspension of judgement' but not to agnosticism according to the aforementioned proposal. I conclude, then, that insisting on identifying constructive empiricism with agnosticism within the probabilistic idiom leads one to a very narrow and implausible representation of the former, i.e. as requiring that one assigns to his theories probability exactly equal to 0.5.

[68] In fact van Fraassen might even be happy with the interval [0, 0.5]. Since he is not an orthodox Bayesian and allows changes of belief without conditionalisation, one can assign 0 probability to a hypothesis today and change his mind in a non-conditionalisation fashion tomorrow (cf. 1998, 219).

One way out of this problem might be to claim that instead of assigning probability equal to 0.5, an agnostic should simply suggest that there are propositions for which we have no rational degrees of belief at all. Arguably, if we have no degrees of belief at all in the truth of certain propositions, we can be no other than agnostic about them. Despite the fact that such an epistemic (non) attitude can plausibly count as agnosticism, though, it is highly counter-intuitive that constructive empiricism amounts to having no degrees of belief at all in the truth of our scientific theories. It is very sensible to understand van Fraassen's view as denying that our reasons for belief in truth are as strong as the realist suggests. On the contrary, it just does not seem plausible at all that van Fraassen's pessimism with respect to our knowledge of the unobservable realm extends so far as to say that we do not even have the slightest clue about it.

These sentiments are reinforced by van Fraassen's own occasional presentation of constructive empiricism as a fall-back (i.e. logically weaker) position relative to full-blown scientific realism. Here is van Fraassen on the merits of constructive empiricism compared to realism:

> "[B]elief in the truth of theories is supererogatory. Suppose that nothing except evidence can give justification for belief. However flexibly this is construed, it means that we can have evidence for the truth of a theory only via evidential support for its empirical adequacy. The evidence then still provides some reason for believing in the truth, *a infirmiori* so to say ... but the additional belief is supererogatory" (1985, 255).

This clearly demonstrates that van Fraassen himself has *not* intended constructive empiricism as implying that we can have no rational beliefs at all in the truth of our theories. On this understanding, constructive empiricism is only a 'safer' option than realism on the basis of logic alone. If belief comes in degrees, then, and we can have beliefs in truth, we can also have degrees of belief in the truth of our theories. Restating van Fraassen's argument in probabilistic terms makes this point even more transparent. Since constructive empiricism is logically weaker than realism, i.e. since belief in truth implies belief in empirical adequacy but not *vice versa*, it follows from the probability calculus that the probability for truth can never exceed that of empirical adequacy. This for van Fraassen is reason enough to opt for his account rather than realism. Irrespective of whether he is right in his assessment[69], though, it follows from his argument that a constructive empiricist has well-defined degrees of belief in the truth of our theories.

In effect, I have argued for a low probability understanding of constructive empiricism and for denying its identification with agnosticism within the probabilistic language. This line of thought is also perfectly consistent with the positive element of this account, i.e. the thesis that our current theories are empirically adequate in the interesting sense that they save all the phenomena, whether past, present or future, observed or unobserved. Surely, this claim is not trivial but requires some kind of inference. Additionally, this inference has to be inductive in kind and, at best, will make van Fraassen's claim 'probable enough' rather than established. At any rate, we can't possibly expect van Fraassen's particularly powerful claim of empirical adequacy to be established with anything other than some degree of probability. Very much like realism, then, van Fraassen's position amounts to saying that 'although we don't have good reason to believe in the (approximate) truth of our best theories, we do have reasons to think they are

[69] For a similar probabilistic reconstruction of van Fraassen's argument and criticism see also Psillos (1999, 204-105).

empirically adequate', which translates quite naturally into 'although it is not very probable that our theories are (approximately) true (although they might be), there are good reasons to think that it is quite probable that they are empirically adequate'[70]. Hence, van Fraassen's position amounts to assigning point-valued probabilities between (0, 0.5] to the (approximate) truth of our theories, and greater than 0.5 to their empirical adequacy.

Finally, we can also see why van Fraassen's argument against constructive empiricism being identified with low probability has already been neutralised. This argument, recall, poses as a plausible requirement that 'agnosticism' is closed under negation: if one is agnostic about A, one ought to be agnostic about $\neg A$ also. But we have already denied that it is advisable to identify constructive empiricism with 'agnosticism' within the probabilistic idiom. This being so, the road is open for a low-probability understanding of constructive empiricism.

I conclude, then, *contra* van Fraassen, that constructive empiricism is best understood as assigning *point-valued* probabilities lying in the interval (0, 0.5] to the (approximate) truth of our theories, and *point-valued* probabilities greater than 0.5 to their empirical adequacy. And the same argument I offered earlier for the troublesome case of someone assigning probability 0.500000001 to GTR, I am willing to offer for his constructive empiricist counterpart who thinks that GTR deserves probability 0.0000000001. Probabilistic reasoning expresses indefinite knowledge even assuming point-valued attributions. Within this indefiniteness,

---

[70] Hence, when van Fraassen asserts that according to constructive empiricism "*acceptance of a theory involves as belief only that it is empirically adequate*" (1980, 12; emphasis in the original), his notion of 'acceptance of a theory as empirically adequate' is again reduced to a mere probability statement for precisely the same reasons as the realist's notion of 'acceptance of a theory as true', outlined above in section 2.4.

there ought to be room for both realism and constructive empiricism. And as we have seen, there is.

Van Fraassen's own reaction was quite unclear. In his (1998) he admitted that his interval-valued account was not adequate as an explication of the concept 'agnosticism' but maintained that it was still good enough for the purposes of describing the constructive empiricist's epistemic attitude. Van Fraassen explains:

> "[I]n the initial context, the important desire was not to be agnostic about theories which postulate unobservable entities - the desire was to be agnostic about the existence of the entities postulated. I submit that this desire can be satisfied. But what has now become clear is that we need to distinguish between these two related concerns" (1998, 217).

This is surely a surprising remark. Indeed, how else can we be agnostic about the unobservable entities our *theories* postulate, except by being agnostic about the theories themselves? Van Fraassen seems to be repeating the fundamental misconception of *entity realism* here, i.e. he wrongly supposes that one can hold differing epistemic attitudes with respect to the entities we can access *only* with the mediation of our theories and the theories themselves. However that may be, though, I have maintained that one can reasonably model constructive empiricism quite faithfully, without any reference either to 'agnosticism' or interval-valued probabilities.

2.5.3 Epistemic Structural Realism

The last epistemic rival to full-blown realism I wish to consider briefly is *Epistemic Structural Realism* (ESR). ESR is a moderate realist position, apparently standing in-between full-blown realism and constructive empiricism. It aims to take into account both the realist's reasons for his optimism as well as one of the main reasons for the anti-realist's (though not van Fraassen's own) reservations.

The realist's optimism is based on the No-Miracles Argument (NMA), which claims that unless we take our theories to be true, their undeniable predictive success would amount to a miracle of a near-cosmic scale. This is the argument on the basis of which the realists have been proclaiming, misleadingly as we saw, that they 'accept' their theories in an epistemically special way. The exact content and success of the argument will be the subject of much to follow; for the time being let us accept that it does carry considerable force. The sceptic's reservations, on the other hand, derive in great part from the famous 'pessimistic induction', an argument usually attributed to Laudan (1981), but whose essentials are found already in Poincaré (1905) and Duhem (1906). The 'pessimistic induction' aims to undercut the NMA. The NMA asserts that there is a close link between predictive and experimental success and truth. In response, Laudan offered a list of theories in his (1981), which were once considered true by virtue of their empirical success, yet were subsequently abandoned as false. Laudan's conclusion is that there is in fact no connection between success and truth, a claim he famously termed 'the confutation of convergent realism'.

ESR tries to combine NMA and the 'pessimistic induction' by asserting that science reveals (approximately) the *structure of the world,* preserved within the continuous development of the mathematical laws of physics, while remaining silent about "the *nature* of the basic furniture of the universe" (Worrall 1989a, 162)[71]. Hence, the epistemic structural realist still adheres to the realist theses (R1)

[71] Emphasis in the original.

and (R2) (p. 64 above) but modifies (R3) to the effect that only (approximate) *structural* knowledge is accessible to us humans.

In effect, ESR falls in-between realism and constructive empiricism[72]. The epistemic structural realist sides with the latter in that we are not entitled to the realist's epistemic optimism. He further claims, though, that approximate structural truth is attainable, its possession being evidenced in the continuous development of the mathematical expression of physical laws, which manifest the structure of the world.

It has been pointed out (cf. Howson 2000, 39-40) that structural realism can also be viewed as offering a detailed account of approximate truth. Indeed, it is quite reasonable to suggest that if ESR is correct and science does reveal the (approximately) true structure of the world, then we can claim that our theories are approximately true[73]. And since ESR is a well developed theoretical approach, it enjoys the additional advantage of being much clearer than most competing accounts of approximate truth. Hence, ESR not only combines the best realist and anti-realist arguments into one coherent thesis, it also provides a workable account of the elusive and obscure notion of approximate truth.

---

[72] There has been considerable discussion lately regarding whether ESR is any different from either realism or constructive empiricism (cf. Demopoulos and Friedman 1985, Psillos 1999, Worrall and Zahar 2001, Ketland 2004). For present purposes I assume, although I don't argue for it, that ESR does succeed as a freestanding position.

[73] Note, however, that ESR claims that science is *approximately* structurally true. This means that although ESR offers a more precise account of approximate truth when compared to full-blown realism, it still makes use of a similarly vague notion of 'approximation' when referring to *approximate* structural truth.

While this line of reasoning is perfectly sensible, it is nonetheless logically independent of how ESR and full-blown realism are to be modelled in Bayesian terms. Regardless of whether they express themselves in Bayesian terms, full-blown scientific realists tend to have a stronger notion of approximate truth in mind than ESR allows for. Accordingly, they claim that their position is markedly different from ESR[74]. Whether their disagreement is presented as one over the precise content of the notion of 'approximate truth' or over the proper scope of application of an antecedently agreed upon notion of 'approximate truth' is largely a linguistic issue, just a different way of phrasing the same substantive point of disagreement.

Taking this into account, I follow standard usage in my Bayesian reconstruction of ESR and focus on the *extent* to which it claims that an antecedently understood concept of 'approximate truth' can be claimed for our theories. Very much like the constructive empiricist, then, the structuralist does not share the realist's epistemic optimism. Hence, he assigns to the *unqualified* approximate truth of our theories point-valued probabilities lying in the interval (0, 0.5]. On the other hand, he does not merely claim that our theories are empirically adequate. The structuralist is optimistic regarding our attaining (approximate) *structural truth.* Following precisely the same rationale as in the case of realism, the structuralist's position amounts to assigning probability greater than 0.5 to our theories being approximately structurally true.

[74] A structuralist might wish to contend this by arguing that ESR provides the *only* workable notion of approximate truth. Such a strong impossibility result, nonetheless, needs independent argumentation.

Finally, it is interesting to note the connections between the various philosophical standpoints we discussed in this chapter. It seems quite reasonable to suppose that if a theory is true, then this theory ought to be both structurally true as well as empirically adequate in van Fraassen's strong sense. Hence, a realist might be expected to hold by default that his theories are also structurally true and empirically adequate. Nonetheless, no serious realist claims that total truth is forthcoming, since even today's science is clearly not complete. This is why realists typically speak of 'approximate truth'. The problems regarding the definition of this notion notwithstanding, it is not obvious that approximate truth entails either structural truth (even approximate structural truth) or full empirical adequacy. Especially with respect to the later, it seems clear that approximate truth does *not*, in fact, entail full empirical adequacy. Indeed, the main reason why a realist argues only for *approximate* truth is the reasonable expectation, in the face of the historical record as well as the open problems current science faces, that there will be some evidence, which current theories will not in fact be able to deal with (at least properly). The same remarks hold for structural realism. No structuralist infers to the total structural truth of our theories; instead, the notion of approximation must be invoked here too. Hence, even if empirical adequacy follows from total structural truth, this seems no longer to be the case when we consider *approximate* structural truth.

In what follows I do not attempt to resolve this issue, since it is largely tangential to my purposes in the rest of the thesis. It should be noted in passing, however, that this consideration of the logical connections between the main competitors in the realism debate reveals that the notion of empirical adequacy van Fraassen has in mind is too strong. A more appealing version of constructive empiricism would, perhaps, result from substituting reference to full empirical adequacy with *approximate* empirical adequacy. To the best of my knowledge, however, van Fraassen has never proposed such a revision.

2.6 Conclusion

My aim in this chapter was to show that a standard preliminary objection to the Bayesian is misguided. The objection amounted to the claim that the Bayesian lacks the conceptual apparatus to capture the subtleties of the realism debate, and in particular those revolving around the issue of 'theory acceptance'. I maintained that this claim is in fact misguided since it depends on our *unjustifiably* strong pre-theoretic intuitions that there is an epistemically significant concept of 'theory acceptance' for non-trivial propositions. Rather than modifying Bayesianism in order to do justice to these intuitions, I suggested that they are abandoned in favour of the more precise, accurate and sophisticated probabilistic language. The resulting probabilistic reconstruction of realism and its main competitors shows, I think, how the probabilistic language best captures all their subtleties and essential details relevant for the debate among them. Hence, now that we know that the Bayesian can fully appreciate all the crucial concepts of the debate, we can proceed to a Bayesian reconstruction of its *content* and see whether we can learn anything useful from this point of view.

Chapter 3

Bayesianism and the No-Miracles Argument I – Incompatibilism Examined

The discussion so far has revolved around the adequacy of the Bayesian framework to provide a formal reconstruction of the realism debate. It has, in other words, been mostly concerned with whether the *concept* of realism can be faithfully captured within Bayesianism, a question we answered affirmatively. Nothing has been said with respect to the *content* of the debate, i.e. which of the various positions outlined earlier, if any, is judged the correct one from a Bayesian point of view. On this issue, there are a number of claims to consider. First, we need to identify the most serious informal argument for the realist thesis. Once this is done, we can proceed to its probabilistic reconstruction. The problem is that there is no universal agreement with respect to the precise understanding of this argument nor is there only one probabilistic reconstruction of its various interpretations. In this chapter I commence an attempt to bring some order to the discussion of these issues. My tasks will be to: a) single out the most frequently-cited informal argument for realism – that is, the No-Miracles Argument (NMA) and its possible interpretations, and b) discuss and disarm a set of arguments claiming either that Bayesianism shows that all interpretations of NMA are unacceptable or that Bayesianism fails to capture the NMA's essential features. This will prepare the way for the positive Bayesian reconstruction of NMA I attempt in chapter 4.

3.1 The 'Ultimate Argument for Scientific Realism'

The most popular argument for Scientific Realism is the so-called *No-Miracles Argument.* Its name derives from Putnam's remark that "the positive argument for realism is that it is the only philosophy that doesn't make the success of science a miracle" (Putnam 1975, 73). However, the claim itself has a long pre-history and has been understood in two markedly different ways. According to one conception, NMA is just another plausibility argument resting largely on our intuitions, while according to the other it represents an application of the distinct and allegedly sound mode of inference called *Inference to the Best Explanation.* It is unfortunate, however, that the terminology occasionally used tends to obscure the situation. Thus, Musgrave (1988), for example, offers a reconstruction of NMA as a *plausibility argument* despite the fact that he makes heavy use of the term 'best explanation'. In what follows and for the sake of clarity I shall reserve the term 'best explanation' only for the non-intuitive version of NMA.

3.1.1 Plausibility Considerations

The intuition behind the plausibility version of NMA (hereafter also referred to as the 'intuitive' NMA), put forth explicitly by Worrall (1989a) but with some of its elements going back to Smart (1963), Maxwell (1962, 1970) and, most notably, as Worrall himself insists, Poincaré (1905), is conveyed by the following *prima facie* obvious consideration. It is widely accepted that current science is predictively very successful. Some of its predictions are born out to an extraordinary degree of accuracy, as testified through sophisticated and reliable processes of testing and measurement. Given this relatively uncontroversial fact, it would surely be a miracle if scientific theories failed to be fundamentally correct in at least some important ways in their theoretical description of nature. Miracles, however, are commonly thought of as our last resort, after all attempts to provide a rational account of some phenomenon have failed. That science's predictive success is due to its being 'essentially correct' in its description of the unobservable reality is both a plausible as well as a non-miraculous claim, and, hence, ought to be preferred to the miraculous alternative.

Despite the intuitive appeal of this understanding of the NMA, the concepts employed in it are not transparent. One issue has to do with the notion of predictive success. Although it sounds innocuous to say that science is a predictively successful enterprise, it has been forcefully argued that the empirical success of a theory should *not* be equated merely with its ability to *entail* the phenomena, since any theory can trivially satisfy this requirement. Indeed, this is just a consequence of the celebrated Duhem thesis, i.e. the idea that a theory is always tested in conjunction with initial conditions and auxiliary assumptions. Given enough ingenuity, any recalcitrant datum can be made to follow from a system based on any core theoretical claim by suitably modifying some one of the various parts of the conjunction under test. Surely, though, the realist wants his theories to be empirically successful in a more genuine way than the aforementioned truism allows for.

Perhaps the most widely accepted non-Bayesian characterisation of *genuine* empirical success is due to Elie Zahar (1983) and John Worrall (1985, 1989a, 2002) and is usually referred to as 'the use-novelty' account. It suggests that a theory is genuinely empirically successful if and only if the successfully predicted data were *not* used in the construction of the theory under consideration. This allusion to the 'construction of a theory' in the clearest case means that the theory initially left unspecified some free parameter, which is evaluated only by means of the data in question. The claim of 'use-novelty' is that, when these circumstances are in place, the data can't be taken to provide any (or at least full) support to the core theoretical claim of these systems relative to any other, which predicts them independently. The claim of the realist is that our current best theories do in fact exhibit this demanding form of predictive success.

Even if one grants, though, that our best theories are genuinely predictively successful, this does *not* entail that they are approximately true. Of course, no serious proponent of the NMA claims that the argument proves the realist thesis. It is still logically possible that our genuinely predictive theories are false in their description of the unobservable parts of nature and that, in due course, new theories will take their place in the scope of the NMA. Despite this possibility, however, there is an intuitively clear sense that it would be too far-fetched to believe that our world is not even approximately as our theories say it is, even if that is logically possible. Although there is no deductive proof from indubitable first principles that the NMA delivers its conclusion, our 'inductive intuition' assures us that all the future, logically possible competitors to our theories are indeed not worth worrying about. The only *plausible* view is that it is very likely that our current mature science is (approximately) true.

How is this plausibility claim justified in its turn, though? The answer provided by its defenders is that it is not justified by means of an argument. For any further argument would involve non-trivial premises, which would need further arguments in their support and so on to an infinite regress or a logical circle – in fact, this is no other than an elaboration of Hume's problem of induction. Nonetheless, we seem to possess some sort of 'inductive intuition', which guides and to date has guided successfully our ampliative reasoning, be it everyday or scientific[75]. No one seriously questions the truth of 'obvious' but non-trivial propositions assuming the existence of an external world revealed to us through the senses. Affirming such propositions, though, depends precisely on a consideration of the same kind as NMA: it would surely be a miracle if our everyday experiences, feelings etc. failed to correspond to an objective external reality, despite the fact that there is no demonstrative proof of the truth of these propositions *vis à vis* the infinity of logically possible metaphysical alternatives. For someone who accepts the everyday version of NMA, then, it would arguably be inconsistent to deny the scientific one.

Of course, as noted a little earlier, the above argument does not establish the realist thesis, since it assumes that one justifiably accepts everyday extrapolations on the basis of inductive intuitions of the NMA kind. These extrapolations are non-trivial claims too, and hence, non-demonstrable. The intuitive force, nonetheless, of

[75] The term 'inductive intuition' was used by Carnap (1968b), yet his own understanding of this term is quite different from the use I make of it in the main text. For Carnap 'inductive intuition' refers first and foremost to our intuitive knowledge of the *axioms* of the system of inductive logic, which are logical in nature, rather than our ability to acquire empirical knowledge. Furthermore, and despite the fact that he regards inductive logic as *a priori*, he makes explicit reference to circular but admissible ways of reasoning, something modern *a priorists* deplore.

our everyday inductive judgements is so overwhelming that it allows us to regard our 'inductive intuitions', everyday or scientific ones, as fundamentally sound on *a priori* grounds (cf. Worrall 1989b, 384;1999, 350 and 360-361). This approach to the problem of induction is very well expressed by Pierre Duhem (who was in turn echoing of Blaisse Pascal): "We have an impotence to prove, which cannot be conquered by any dogmatism; we have an idea of truth which cannot be conquered by any Pyrrhonian scepticism" (1906, 27).

It has been argued, however, that this line of defence, no matter how honest in terms of taking into account our logical limitations, is not satisfactory. This complaint originates in the work of Richard Boyd, who held the view that the 'intuitive' NMA[76] "does not address the crucial epistemological claim of the empiricist argument: that since factual knowledge is grounded in experience, it can extend only to observable phenomena" (1984, 43). Though Boyd readily admits that the 'intuitive' NMA "is probably the argument that reconstructs the reason why most scientific realists are realists" (ibid. 49), he still doubts that its 'intuitive' formulation is strong enough to underwrite the realist claim and decisively defeat all anti-realist arguments. Boyd's central worry seems to be that plausibility considerations based on *a priori* grounds are too weak to offer a satisfactory epistemological foundation for the realist thesis. In his own words, "[the 'intuitive NMA provides us] with a reason to suppose that realism is true, but [does not provide us] with any epistemology to go with that conclusion" (ibid. 50).

[76] Though Boyd never uses the term 'intuitive NMA' in his (1984, 49-50), his preliminary discussion makes it clear that he treats NMA as a plausibility argument. As is well-known, of course, he then goes on to 'strengthen' the argument by subsuming it under the general scheme of *Inference to the Best Explanation.* An outline of this version of NMA is contained in the next section.

Consequently, if we wish the NMA to become an effective argument against the various forms of anti-realism, we should try to enhance its epistemological foundations. Boyd's own suggestion was to supplement the epistemological basis of the NMA by regarding the argument as an instance of abductive reasoning, suitably understood as *Inference to the Best Explanation* (IBE). This is in turn taken to be an independent and sound form of non-deductive reasoning.

3.1.2 NMA as an 'Inference to the Best Explanation'

Boyd (1984) argued that in order for realism to neutralise the anti-realist challenge, what is required is a neutral basis on the background of which the relative merits of the realist and anti-realist positions can be assessed. Boyd found this neutral ground in the way each philosophy *explains* what nearly everyone takes to be an uncontroversial fact, i.e. that our scientific methodology is predictively and experimentally very successful in the demanding way of the 'use-novelty' approach outlined earlier (cf. Psillos 1999, 106-107).

According to the IBE-version of NMA (hereafter also referred to as the 'explanationist version') one should be a scientific realist because realism provides us with the best explanation of the instrumental reliability of current mature science – indeed Boyd implies that it is the *only* real explanation of that phenomenon (Boyd 1984, 66). As Psillos explains, however, Boyd's programme rests on the assumption that "a realist epistemology of science should employ no methods other than those used by scientists themselves" (1999, 78). This *naturalistic* thesis implies the controversial admission that realism offers a *scientific* explanation of the reliability of science, presumably of the same kind with the explanations of natural phenomena provided by ordinary scientific theories. Postponing further investigation of this point until the next paragraph, the explicit explanationist argument, succinctly reconstructed by Musgrave (1988) and repeated in Psillos (2006, 139-140), goes as follows:

(IBE–1) Science is a predictively successful enterprise.

(IBE-2) Realism enjoys the advantage over anti-realist philosophies that it best explains the predictive success of science.

(IBE-3) It is reasonable to infer the truth of a theory from the fact that it best explains the evidence in hand.

Therefore, realism is the correct philosophical perspective[77].

Neither Boyd nor Musgrave nor Psillos of course simply asserts the explanatory superiority of realism in premise (IBE-2). They also provide an intuitively very powerful detailed comparison of realism to the significant anti-realist alternatives, ranging from Berkeleyan phenomenalism to van Fraassen's constructive empiricism (cf. Boyd 1984, 58-63, Musgrave 1988, 60-68 and Psillos 1999, 90-97).

Clearly (IBE-1) is true. However any number of doubts can be raised about (IBE-2). First, it presupposes that we have some clear notion of scientific explanation - a presupposition rendered very doubtful by a glance at the vast and heterogeneous literature on this topic produced over the last 50 years. Secondly,

[77] Musgrave's own understanding of IBE is *not* that it is a distinct, non-deductive from of inference, since he notes that "I [Musgrave] prefer to construe so-called 'inductive arguments' as deductive arguments with 'inductive principles' of one kind or another among their premises. This conduces to clarity and *obviates the need for any special inductive logic*" (p. 54, fn. 69; my emphasis). My understanding is that he simply uses the idiom of explanation to convey what in essence is a plausibility judgement about the merits of what we take to be a well-explaining theory. It remains a fact, however, that most defenders of IBE don't think in the same way, as I explain below. Hence, Musgrave's use of terminology tends to be rather misleading.

even if we agreed on what a scientific explanation is, could we really expect realism to provide a *scientific* explanation of the success of science, as Boyd's naturalism invites us to do? Surely on any (sensible) account this would require independent testability and it seems impossible to see how realism could be independently tested. Moreover, even if we were to accept (IBE-2) so that the inference depended only on (IBE-3), the explanationist would still not be home free since there are important issues about (IBE-3) too.

Since the explanationist defenders of NMA are not satisfied with the *a priorism* of the plausibility version, they can't possibly claim that (IBE-3) is shown to be true by means of a plausibility argument. This would simply collapse their case to the plausibility strategy we encountered in the previous sub-section. What they need to claim, instead, is that they can provide a convincing argument, which shows that IBE is *in fact* a reliable inference pattern, allowing us to infer truth on the grounds of explanatory superiority (cf. Psillos 1999, 79).

This takes us to the most controversial part of the explanationist defence of NMA, since the argument which allegedly demonstrates that IBE conveys epistemic warrant to its conclusions is *circular* and depends explicitly on the success of the NMA. The *motivation* for this kind of approach can be traced back to Goodman's classic (1954) exposition of his 'solution' to the old problem of induction. As I noted earlier, in attempting to justify induction, one finds oneself confronted either by circularity or infinite regress. Goodman (ibid. 63-64), drawing on a parallel with deductive logic, reasoned essentially as follows in his 'solution' of the traditional problem:

1. There can be no non-circular justification for *deductive logic*.
2. We typically assume, however, that we are justified in using it.

3. Therefore, once we are able to formulate a justification for *induction*, which, though admittedly circular, is arguably equivalent to the case of deduction, we should be justified in using it.

If we are willing to tolerate circularity in the case of deductive logic, doesn't consistency require that we do the same for induction?

Determining the truth-value of premise 1 in this latest argument, of course, is not a trivial issue. Those who are suspicious of the justificatory merit of all circular arguments will look for some other way to account for our allegedly justified use of deduction. And, as we have seen, allowing for *a priori* knowledge easily accommodates this requirement: if our strong inductive intuition warrants belief in the NMA, then our even stronger deductive intuition surely warrants belief in deductive logic.

On the other hand, those who accept premise (1) face the task of arguing for the counter-intuitive claim that circular arguments can serve justificatory purposes. The next step, then, is to characterise precisely the notion of circularity involved in the case of deduction and investigate whether it can indeed be harmless.

It seems incontestable that an argument is *viciously* circular when its conclusion is contained *explicitly* among its premises. Such an argument fails to provide reasons for the truth of the conclusion, since it merely reasserts the truth of what is already taken to be a premise. When we try to justify deductive reasoning, though, there is another kind of circularity involved, or so it is claimed, which is *not* vicious. In the deductive case we typically are asked to prove a meta-theorem in the meta-language that, for example, *modus ponens* in the object language is truth-preserving (a soundness theorem). In doing so, though, *we have to use modus ponens* (cf. Psillos 1999, 86). Here the conclusion, we are told, is not among the premises. Instead the argument 'merely' rests on a rule of inference, which would not be sound if the conclusion arrived at using it were not itself true. Since the conclusion of the argument is not *modus ponens* itself but that *modus ponens is truth-conducive*, merely using modus ponens does not make the argument viciously circular. It would have been so only *if we had assumed in the first place that modus ponens is truth-conducive*. This kind of circularity is often referred to as *rule-circularity*[78].

Attempts to justify induction using rule-circular arguments go back to Braithwaite (1953) and, more recently, Van Cleve (1984)[79] and Papineau (1993). Psillos (1999, 82-83) has more recently argued that the IBE version of NMA also belongs to the class of merely rule-circular, rather than outright viciously circular, arguments. According to him, scientists typically construct their theories using IBE as their mode of inference. NMA then asserts that these theories are approximately true on the basis of their predictive success, also using an IBE at the meta-level. Since these theories have (allegedly) *both* been shown to be true *and* reached in the first place by means of IBE, one is allowed to conclude that IBE is a reliable rule of

[78] I should note that despite the fact that this line of reasoning follows Goodman's in *structure*, my characterisation of circularity is *distinct* from Goodman's own 'solution' to the (old) problem of induction. For although Goodman's proposed 'solution' also relies on the notion of acceptable circular arguments, he does not endorse the reliabilist framework. Instead he put forth the idea that, very much like deduction, the justification of inductive rules consists in their being 'fine-tuned' with the particular inductive inferences we are ordinarily willing to sanction and vice versa (1954, 63-64), a process which is also known as 'reflective equilibrium'. This is the reason why I have taken Goodman's argument to provide *only* the motivation for the reliabilist solution.

[79] Neither Braithwaite nor Van Cleve considered their attempted justification as an extension of what holds in deductive logic to the inductive case. Instead they applied it independently to induction.

inference. The argument is claimed not to be premise ('viciously') circular because the conclusion that our theories are approximately true is not among the assumptions of NMA and, therefore, since it is 'only' rule-circular, it can legitimately be taken as vindicating the mode of inference involved - in this case IBE.

Hence, IBE is a reliable rule of inference and allows us to reach the conclusion that our theories are approximately true because this very conclusion shows that IBE is reliable. Or, put differently, (IBE-3) gets support from its conclusion in a non-vicious way and that is why NMA is a permissible and forceful argument.

There are some serious problems with this line of argument. First of all, this version of NMA makes use of a rather obscure distinction between particular theories and scientific methodology in general. In particular, since the NMA is a meta-IBE, it is supposed to refer to scientific methodology in general and *not* to particular theories. This move allegedly allows us to validate IBE as a general rule of inference through validating the many different IBEs, which gave rise to the "acceptance" of particular theories of current science. But scientific methodology in general is surely nothing more than the union set of the methodologies used in the many particular sub-disciplines of mature science. Hence, it is odd to refer to NMA as an argument distinct from the particular IBEs it is supposed to validate. In fact, particular versions of NMA can perfectly naturally be run with reference to the predictive success of each particular theory, thus making the reliabilist's NMA merely a condensed and economical way to refer to all mature scientific disciplines and their respective particular NMAs in one go (cf. Worrall, forthcoming).

Furthermore, in order to use IBE in the first place, according to this reliabilist account, do we not need to *know* that IBE is reliable? And if the only way to do so is by knowing that our theories are true, do we not need to assume this also? But clearly, assuming that our theories are true would make the argument viciously circular. In the face of these difficulties Psillos has insisted that dismissing this danger requires a radical shift in one's epistemological standards (1999, 84-85). Traditional internalist epistemology requires that one *knows* or *has reasons to believe* that the rules one is using are reliable before one is justified in using them. Rival externalist accounts do not. They insist that the rule need simply *be* objectively reliable, regardless of whether we know it or not.

Put this way, however, the reliabilist position still faces considerable difficulties. The whole recent debate on realism revolves around the *epistemic* problem of whether *we* have reasons to believe in the truth of our theories. Removing completely the knowing subject from the picture by insisting that the rule he uses simply be objectively reliable does not of course give any reasons for belief. Psillos tried to amend this defect by insisting that the proper way to interpret the reliabilist thesis is by requiring that "one should have no reason to doubt the reliability of the rule" (ibid. 85). This formulation sneaks the subject back into the picture. Quite obviously, though, it also reintroduces problems of circularity (or infinite regress) – how are these reasons to be judged? and so on and so forth[80].

Both renditions of the realist NMA, then, face some serious philosophical challenges. The 'intuitive' version, on the one hand, defends scientific realism on

[80] The prospects of an externalist theory of justification as the background for the explanationist NMA are discussed in more detail in section 5.2.2 of chapter 5.

the basis of the dictates of our 'inductive intuition', which are only assertible on *a priori* grounds, and, hence, seems to rest on an overly weak justificatory foundation. The explanationist version, on the other, in its attempt to supply a sounder epistemological basis for the defence of the realist thesis, faces serious problems of internal coherence. In the next chapter I offer probabilistic reconstructions for *both* versions of the NMA in order to investigate their character and relative merits from the perspective of Bayesian Confirmation Theory. I shall argue that Bayesianism not only further illuminates the logical structure of each version, thus helping us to decide which one best serves the realist cause; it also contributes a lot towards clarifying the peculiar philosophical nature of the realism debate as a whole by bringing to light the interplay between the various considerations that can be adduced in favour of or against realism. Before we proceed with our probabilistic variants of NMA (and, also, as a means of going about it), however, we had better take a close look at some arguments claiming that Bayesian reconstructions of NMA *demonstrate* that the argument, in any of its versions, is fallacious.

3.2 Some Incompatibilist Claims Examined

There have been a number of arguments purporting to show that the NMA is entirely unacceptable from a Bayesian perspective. Some such incompatibility arguments have focused solely on the IBE version of NMA; van Fraassen, for example, claims that "[anyone] who becomes converted to the use of some sort of probabilistic IBE [of which the NMA is an instance] ... quickly discovers that he is led into incoherence" (1989, 161). Others have questioned the general intuition behind the NMA, irrespective of the particular interpretation one gives to it;

Howson, for instance, claims that "as it is usually presented [the No-Miracles argument] is fallacious, thoroughly fallacious… ." (2000, 59). Hence, before we commence our Bayesian reconstruction of NMA, it is necessary to examine and counter all these incompatibilist views. Let's focus on IBE first.

3.2.1 IBE Is Incoherent

The idea that IBE is incoherent is due to van Fraassen (1989) and is part of his attempt to discredit the realist thesis in comparison to his favourite constructive empiricism. Van Fraassen's arrives at this conclusion by modelling IBE in Bayesian terms and then demonstrating that it is susceptible to the *diachronic* Dutch-book argument for conditionalisation that we encountered and discussed in detail in chapter 1[81]. If sound, his argument would totally neutralise the second interpretation of NMA, irrespective of the merits or demerits of reliabilism.

As Van Fraassen sees it, we should think of IBE as proceeding as follows:

1. Start with all the alternative hypotheses you have plus your data and use Bayesian Conditionalisation in order to determine the posterior probabilities of these hypotheses in the light of the available evidence.
2. After you conditionalise, though, add *bonus points* to your hypotheses according to their explanatory merits.

---

[81] The probabilistic rendering of IBE, van Fraassen thinks, is the last defense line of the explanationist in the face of further arguments he had already offered against IBE in his (1989, 142-150). Although I shall not discuss them here, one can find extensive discussion and criticism in Lipton (1993), Psillos (1999) and Okasha (2000).

3. Hence, the posterior probability of a hypothesis is determined by normal Bayesian Conditionalisation *plus* the bonus probabilities based on explanatory considerations (whatever these may be) (ibid. 166).

Assuming this probabilistic construal of IBE, van Fraassen argues that, since it violates Bayesian Conditionalisation, IBE is incoherent on the basis of the diachronic Dutch-book argument.

Van Fraassen's argument has been challenged from many different angles. Day and Kincaid (1994) and Okasha (2000) admitted the force of the dynamic Dutch-book argument but questioned the way van Fraassen chose to model IBE, i.e. in terms of bonus points. They suggested that his modelling option is neither compelling nor even 'natural' - claiming that it is in fact "an idiosyncratic way of representing IBE in probabilistic terms" (Okasha 2000, 703). Okasha has argued quite convincingly that, when engaging in explanatory considerations of the sort IBE tries to codify, there is nothing like the two stage process on which van Fraassen based his reconstruction. Instead of adding explanatory considerations to normal Bayesian response to new evidence, one merely uses explanatory considerations in order to decide how to handle this evidence (ibid. 702-703).

This reaction is, I think, fundamentally correct. IBE was never intended as a supplement or addition to Bayesian Confirmation Theory – one that requires such a probabilistic reconstruction as van Fraassen's. The following problem still emerges however. Ever since Harman's seminal (1965) paper, IBE has been systematically championed as an *independent* form of inductive inference (cf. Lipton 2004, Psillos 2002). Van Fraassen's proposed reconstruction, although not faithful to the common understanding of the workings of IBE, *does* convey this intuition that most of its defenders share, i.e. that IBE is in fact an independent form of inference.

Rejecting van Fraassen's interpretation of it does not imply, of course, that IBE loses this status. It does imply, nonetheless, that if IBE is to maintain its independence, then the probabilistic reconstruction that should replace van Fraassen's own has got to reflect it somehow. Okasha touches on this issue when he discusses possible objections to his own reconstruction (to be discussed in the next chapter), only to conclude that "a fundamental and unresolved question is whether the Bayesians are *explaining*, or just *representing* [the old tradition of non-formal accounts of scientific method, in which IBE belongs, when probabilistically reconstructed]" (ibid. 706)[82]. As I explain in the next chapter, I think this question is far from unresolved. In fact, it seems that the character of all *viable* probabilistic reconstructions of IBE tells decisively against any claim of independence its defenders can make. All this, however, presupposes the discussion of some compatibilist accounts between Bayesianism and IBE and so will be postponed for the present.

Douven (1999) offered an alternative, decision-theoretic way of defending IBE, which concedes to van Fraassen his modelling option in terms of bonus points, yet denies the validity of the diachronic Dutch-book argument for conditionalisation. His strategy depends crucially on a literal reading of the betting situation and the ability of the agent to foresee the Dutch-book. In particular, Douven argues that van Fraassen's probabilistic construal of IBE can be combined with a method of calculating the initial probability of a statement such that some of the bets giving rise to the diachronic Dutch-book would no longer seem fair. Hence, the agent would avoid the trap by refusing to buy these bets in the first place (ibid. 429-433).

[82] Emphasis in the original.

I have already maintained in the first chapter that such a literal reading of the betting situation misses the whole point of the argument by transforming a challenge directed against our inferential methods *per se* into a challenge directed against the agent's intelligence or prudence. To repeat briefly: such a literal reading of the betting set up fails to appreciate the original intention of finding a normative basis (no matter how idealised) on which to ground considerations of consistency. This line of thought, which treats Dutch-books as mere symptoms of inconsistency, not only fails to impair the normative force of Dutch-book arguments but also results in a clear and systematic framework for the treatment of consistent partial belief. It is quite natural, then, to conclude that a decision-theoretic denial of the dynamic Dutch-book argument hardly stands up to the *normative* challenge presented by van Fraassen[83].

Having already denied the validity of the dynamic Dutch-book argument myself in chapter 1, though, I hasten to say that the alternative strategy of blocking van Fraassen's charge through rejecting the validity of this argument is sound. It is important, nonetheless, to have in mind the *correct reasons why it fails*. Unlike Douven, who thinks that the dynamic Dutch-book argument fails on the basis of the contingent computational abilities of rational agents, I have already endorsed in chapter 1 a different line of criticism. This centres on the observation that consistency is only a synchronic notion. Hence, one ought to be free to update diachronically one's degrees of belief at will, under the sole constraint that he

[83] Douven himself came close to acknowledging the implications of the literal reading of the betting scenario in footnote 8 of his paper (1999, S433), yet he dismissed the challenge that they pose for his preferred solution on the grounds that "[it] is not at all clear [that the decision-theoretic solution is shown to address an inessential problem]" (ibid. S433, fn. 8). Except from some very brief remarks, though, Douven presents nothing like an argument for his stance.

adjusts his new synchronic beliefs accordingly, so as to avoid proper synchronic Dutch-books. Of course this way out of the problem does not imply that van Fraassen's modelling option is indeed the correct one, it merely shows that, even if it were, IBE would still have not been shown incoherent.

This way of putting the issue also sheds some light on van Fraassen's views on *rule-following*. Van Fraassen's original intention was to model IBE as a *rule* of belief revision. The diachronic Dutch-book argument, then, was supposed to establish that any such rule other than conditionalisation was bound to lead to inconsistency. Van Fraassen, in effect, invites us to consider the following problem:

> "Could there be a...view, that we have rules which (*a*) we are rationally compelled to follow, which (*b*) leave nothing to our choice when we proceed rationally, and yet (*c*) give us new expectations that are not logically implied by our old opinions leavened by new experience?" (1989, 171).

Deductive logic most clearly satisfies (a) and (b), and also most clearly fails to satisfy (c). When it comes to IBE, though, things are less clear. It seems reasonable to suggest that IBE supporters intend their inferential method as 'rational' *in some sense*, which, again *in some sense*, is supposed to dictate the correct answer to our inferential problems. It is also clear, though, that these, hitherto unspecified, senses of rationality and compulsion are markedly different from the case of deductive logic. IBE is intended as a non-deductive, and hence *logically non-compelling*, rule of inference. At best it is hoped to be a factually *reliable* rule of inference, leading us to *contingently* true propositions about the world. This being so, and despite the apparent vagueness surrounding the notion of IBE, all one can say against IBE is that it is not a factually reliable rule of inference, or that it is not reliable enough to talk about the unobservable realm, or other such things. What one *cannot* say is that IBE is incoherent, in the sense of *forcing* us to hold incoherent degrees of belief.

It follows, then, that van Fraassen's conclusion that "*if and when we commit ourselves to a rule* for the revision of opinion, it [the rule] must be non ampliative" (ibid. 174)[84] reveals an overly strong conception of inductive rules, which tends to equate them with the canons of deductive inference. Indeed, if it is a necessary requirement for the adoption of a rule that it provides absolute epistemic warrant, then surely no inductive rule meets this qualification. What is even worse for van Fraassen, though, not even conditionalisation meets this qualification! Of course, it remains to be clarified in which sense IBE (or any other inductive rule for that matter) is rationally compelling. Surely, though, failure to do so does not make this rationally-compelling-in-some-sense rule incoherent.

It might be argued that van Fraassen can easily restate his case with reference not to conditionalisation and $\Pr_{new}(H)$ *after* the evidence is known but with reference to the synchronic relation $\Pr(H/E)$. In other words, *qua* empiricist, he would simply ask what the probability of the hypothesis is in the light of the evidence irrespective of when this evidence becomes known, following in effect Howson's latest proposal (unpublished) that we encountered in chapter 1. He would then be able to show that violating the axiom $\Pr(H/E) = \Pr(H \wedge E)/\Pr(E)$ by adding bonus points to $\Pr(H/E)$ would make one susceptible to a Dutch-book proper (i.e. to a synchronic Dutch-book). In fact, this was van Fraassen's own initial rendition of the problem (1989, 166-167), only to change his mind apparently on the basis that confirmation is a forward-looking relationship between data and theory most faithfully captured by conditionalisation. And since van Fraassen believes in the

[84] Emphasis in the original.

validity of the diachronic Dutch-book, it is no accident that he eventually charged IBE with incoherence on the basis of this very argument. But we have seen that the diachronic Dutch-book argument fails, because conditionalisation is not a logical principle[85].

Van Fraassen's attack, then, misfires on two counts: a) it is based on a model of IBE which is not faithful to the way IBE is supposed to operate, and b) even if this modelling option were provisionally accepted, IBE could not reasonably be suspected of incoherence, since the diachronic Dutch-book argument ultimately fails. Van Fraassen is surely right when he concludes that "rationality does not require conditionalisation, nor does it require any commitment to follow a rule [outside those of deductive logic] devised beforehand" (ibid. 174). He should add, though, that rationality does not require conditionalisation even if we decide to commit ourselves to a rule of belief revision, unless of course this rule is conditionalisation itself.

3.2.2 The NMA Commits the 'Base-Rate Fallacy'

The claim that NMA is fallacious because it embodies the 'Base-Rate Fallacy' has been forcefully put forth by Colin Howson in his (2000), and repeated by Magnus and Callender (2004). This argument attacks the general orientation of the NMA without distinguishing between its 'intuitive' and explanationist

---

[85] Needless to say, had van Fraassen opted for the synchronic case, the argument would in fact be successful, showing in effect that IBE cannot be modelled in terms of bonus-points. Conceived as a diachronic relation, though, van Fraassen's argument does not go through.

construals. If sound, it defeats *both of them* on the grounds that they violate valid probabilistic reasoning.

Howson (ibid. 36) proposes the following reconstruction of NMA:

1. If a theory *T* is predictively successful yet not approximately true, then its success can only be attributed to chance.

2. The chance, though, that *T* is false and at the same time predictively successful is extraordinarily small.

3. Granted (2), one should reject the hypothesis that the success of *T* is to be attributed to chance, especially in light of the fact that there is an alternative non-miraculous explanation, the realist one.

4. Therefore, we can infer that T is approximately true.

Since Howson's argument makes use of Bayesian probabilities, we have to translate steps (1) – (3) into probabilistic terms. Let us first of all assume that talk of approximate truth is precise enough to allow us express the various relations between an approximately true theory and the predicted evidence *E* without major problems. Let *H*, then, stand for the hypothesis that '*T* is approximately true'. This being so, (1) and (2) suggest that Pr(*E/H)* is 1 or, at any rate, very close to 1, while Pr(*E/¬H*) is on the contrary very close to zero. From these two likelihoods, (3) invites us to conclude – assuming that indeed evidence *E* turns out to hold – that the probability of *H* (i.e. the claim that *T* is at least approximately true) is high. *H* is the natural, non-miraculous alterative to the suggestion that *H* simply got *E* right 'by chance'.

Whatever intuitive appeal this argument might have, Howson claims it is as fallacious as it could possibly be since *all three* of its premises can be shown to be false. His main complaint[86] against the first two refers to the value to be assigned to $\Pr(E/\neg H)$. The NMA assumes, as we saw, that $\Pr(E/\neg H)$ is very close to zero. Goodman's exotic predicates 'grue' and 'bleen', however, allow us see that this assumption is not as trivial as it might seem. Knowing that there are infinitely many alternatives to *T* which save the phenomena equally well ('grue' and 'bleen' simply illustrate this fact), it would appear that $\Pr(E/\neg H)$ equals "the combined chance of all the infinitely many worlds in which *E* is true and *T* not approximately true"(ibid. 45). Such a chance is hardly definable mathematically. To sidestep this problem, one has to *assume* that the *prior probability* of all hypotheses H_i, whose union set constitutes $\{\neg H\}$, is very small. By the probability calculus it can be shown that $\Pr(E/\neg H)$ is proportional to the product of the priors $\Pr(H_i)$ times the likelihoods $\Pr(E/H_i)$ summed over all *i*. Hence, the value of $\Pr(H_i)$ will have to be small enough to outweigh the high value of $\Pr(E/H_i)$, which in its turn is high simply because the empirically adequate alternative hypotheses H_i by definition save the phenomena. If we do this, however, we end up reasoning in a circle, for "there seems to be no way to compute this chance [i.e. $\Pr(E/\neg H)$] without begging the very question that the exercise of computing it is supposed to answer" (ibid. 47). In other words, we already need to have strong prior beliefs in the theory *H* the NMA allegedly supports in order to use the miracles consideration in its favour. But then, one wonders, what good is the argument?

Even if we were to grant as justified the assumption of the low value of $\Pr(E/\neg H)$, this is not sufficient to guarantee the inference to the overwhelmingly

[86] The rest of Howson's (technical) objections to the first two premises can be found in his (2000, 43-45).

probable truth of *H*, as premise (3) invites us to do. Howson's main claim on this count is that the NMA commits the 'Base-Rate Fallacy', i.e. the fallacious habit of ignoring the base-rates (or prior probabilities) in the computation of the posterior probability of a hypothesis[87].

Howson illustrates the 'Base-Rate Fallacy' by means of a medical example, sometimes known as the 'Harvard Medical School test'[88]. Suppose that a diagnostic test for some one disease has a zero *false negative rate* (i.e. that the probability that the test gives a negative result given that the patient really has the disease is zero) and, say, only a 5% *false positive rate* (i.e. the probability that the patient tests positive even though he does not have the disease is 5%). What we want to know is the *posterior probability* of the hypothesis, say *D*, that the patient has the disease given that he tested positive. Despite (allegedly) deceptive appearances[89], the correct answer is *not* 95%, nor even 'very high' but is given by Bayes' theorem,

---

[87] There are many studies providing evidence that most people (but certainly not everyone) tend to neglect base-rates when engaging in probabilistic reasoning. See for example Lyon and Slovic (1976), Casscells *et al.* (1978) and Kahnemann and Tversky (1982). Bar-Hillel (1980) contains an informative overview of the debate regarding the 'Base-Rate Fallacy' as well as a psychological explanation of its persistence.

[88] This test was actually performed in the form presented in the main text at four Harvard Medical School teaching hospitals by Ward Casscells and his collaborators, who published their results in Casscells *et al.* (1978). Their sample consisted of 20 house officers, 20 fourth-year medical students and 20 attending physicians serving in these hospitals.

[89] Casscells *et al.* (1978, 1000) report that only an 18% of the participants in the test answered the question correctly. The most common answer, given by 45% of them, was that the posterior probability was 95%.

which in its simplest form reads $\Pr(D/E) = \frac{\Pr(D) \times \Pr(E/D)}{\Pr(E)}$, and thus clearly depends crucially on the prior probability (or base-rate) Pr(*D*).

Rewriting Bayes theorem as

$$\Pr(D/E) = \frac{\Pr(D)}{\Pr(D) + \Pr(\neg D) \times {}^{\Pr(E/\neg D)}\!\big/\!{}_{\Pr(E/D)}} = \frac{\Pr(D)}{\Pr(D) + f\Pr(\neg D)},$$ where $f = \frac{\Pr(E/\neg D)}{\Pr(E/D)}$ is the *Bayes factor* in favour of $\neg D$ and against *D*, makes the problem transparent. No matter how small the numerator relative to the denominator in Bayes factor is (in the medical example it is just 0.05), this by itself says nothing about the posterior probability of the hypothesis we are after. Whoever answers that the posterior probability of the hypothesis *D* that the patient has the disease is 95% commits the 'Base-Rate Fallacy' for he ignores the prior probability in his computation. If the prior probability is sufficiently low, then the posterior probability will still be low despite an event's having occurred (viz. the patient testing positive) that would be overwhelmingly unlikely (95% unlikely) to occur if the hypothesis that he has the disease were false.

Recall the claims of the NMA on Howson's reconstruction: from the fact that a theory is very unlikely to be predictively successful given that it is false, while at the same time it is very likely that it is predictively successful given that it is approximately true, we are asked to infer that the theory is indeed approximately true. In other words, from the fact that the theory has got a low false positive rate and a close to zero false negative rate, i.e. *a low Bayes factor*, we are asked to infer that the posterior probability of the hypothesis *H* that theory *T* is approximately true in the face of the evidence is very high, or at any rate higher than 0.5 to justify the realist claim. Very much like the medical school case, reasoning this way exemplifies once again the 'Base-Rate Fallacy'. What proponents of NMA do, it is claimed, is precisely to ignore the prior probability of the hypothesis in question, thus committing an elementary probabilistic mistake.

It has become obvious, then, that modelling NMA in probabilistic terms requires one to take into account prior probabilities. Prior probabilities, though, seem foreign to the miracles consideration. The NMA, be it an IBE or just an inductive plausibility claim, describes the alleged relationship between a theory's predictive success and its truth. *Prima facie* it appears that this relationship can only be captured through the likelihood values Pr(E/H) and Pr($E/\neg H$), which are nonetheless shown to be insufficient to determine the required posterior probability.

Careful commentators have reacted to this line of reasoning by suggesting that all Howson's arguments show is that in reality the NMA is just an *enthymeme*, i.e. an argument containing an implicit premise (Lipton 2002, 583; Worrall 2007, 144-147; forthcoming). Why, then, not simply make this implicit premise explicit in the form of adding the prior probabilities into the calculation? Howson does precisely this in his own reconstruction of 'a sounder no-miracles argument', only to conclude that the NMA is still not valid after all (cf. Howson 2000, 57). The reason for this latest claim has to do with the character of these prior probabilities. In the case of the medical example we are normally in possession of fairly reliable statistical frequency data, which can then be plausibly interpreted as 'objective' base-rates or prior probabilities, and which inform us about the overall incidence of the disease. In the scientific case, though, no such prior estimate of the relative frequency of true theories in the population can possibly become available. The only remaining option is that these probabilities are *subjective* estimates. Being subjective, it is standardly argued, they can't possibly satisfy the claims to objectivity the proponents of NMA think it underwrites. Howson's comment is revealing of this line of thought:

> "The proponents of the No-Miracles argument regard it as sound reasoning *as it stands*, without need of any further assumptions, and in particular not estimates, which by their nature must be highly subjective, of how probable types of large-scale world are" (ibid. 55)[90].

Subjective estimates are, then, foreign to the intent of NMA and Howson concludes that the argument is either fallacious or simply fails to deliver.

There are two ways to respond to this dilemma. First, one can deny that the argument is fallacious in the first place. Since the probabilistic reasoning is impeccable, the grounds for such a denial can only be *either* that Howson's particular way of modelling the argument in probabilistic terms is mistaken *or* that the general decision to embed the discussion within a Bayesian framework is ill-motivated. Since these alleged grounds are incompatibilist in spirit, I shall discuss them in the next subsection of the present chapter. Secondly, one can follow Howson's advice and accommodate prior probabilities in one's Bayesian reconstruction of the NMA so that probabilistic validity is restored, but claim that the resulting argument is still a powerful consideration for realism. This possibility will be explored in considerable detail in the following two chapters.

3.2.3 Should the NMA Be Modelled Probabilistically at All?

One way out of the 'Base-Rate Fallacy' is to deny that a probabilistic reconstruction of the realist argument captures its essential aspects. Stathis Psillos

[90] Emphasis in the original.

(2004) is the main defender of this reaction to the fallacy[91]. Recall, first of all, his understanding of NMA as an instance of the general inference pattern of IBE. Psillos readily admits that the 'Base-Rate Fallacy' demonstrates that NMA and the explanation of science's predictive success should not be equated with the likelihoods in Bayes' theorem (2004, 87). At the same time, however, he thinks that subjective prior probabilities fail to do justice to the NMA's claims to objectivity. Furthermore, he is very sceptical about the prospects for 'objectifying' those prior probabilities, calling it a task which "we, presumably, know cannot be done" (ibid. 88).

Hence, Psillos is facing the following dilemma: he *either* has to show that there is a different way to model NMA in probabilistic terms – one that avoids both the 'Base-Rate Fallacy' and the issue of the subjectivity of the priors, *or* deny outright that it was ever appropriate even to seek a probabilistic reconstruction of the NMA.

As remarked a little earlier, Howson's modelling option in terms of likelihoods is natural, since it conveys NMA's focus on the relationship between predictive success and truth. Howson has also established beyond doubt that the likelihoods do not suffice for the realist cause. Hence, someone of Psillos' persuasion is only left with the option of denying that a probabilistic reconstruction can ever do justice to the argument.

In his attempt *independently* to motivate this sort of incompatibilist stance, Psillos focuses his attention on the allegedly independent character of IBE. IBE, we

[91] In his (2004) Psillos discusses the prospects of casting IBE in a Bayesian framework without mentioning explicitly NMA. Of course, the discussion generalises straightforwardly on the assumption that NMA instantiates IBE.

are told, is an independent mode of inference, which tries to strike a balance between amplification and epistemic warrant. As such, it contrasts at least with Subjective Bayesianism, which is supposed to be a non-ampliative theory concerned with the synchronic consistency of our degrees of belief. Thus conceived, Subjective Bayesianism lacks, while IBE possesses, the resources to capture the increase in knowledge that NMA offers (ibid. 88-89).

This argument, though, is surely weak. Psillos claims that Subjective Bayesianism fails to capture the ampliative character of NMA because it is just a non-ampliative extension of deductive logic to partial belief. If he is right, though, one would expect the similarly non-ampliative deductive logic to also fail as a framework within which to cast the informal, non-probabilistic, version of NMA. This conclusion, nonetheless, is untenable. Deductive logic is clearly an adequate framework within which the informal NMA can be systematically reconstructed. In fact, in section 3.1.2 we saw how the explanationist NMA can be expressed as a deductive argument. This is not to say, of course, that the conclusion of the NMA is derivable from indubitable first premises - such a proof is explicitly acknowledged by the arguments' proponents as unavailable. Since the NMA is ampliative, all our reconstruction has of necessity included is (at least) one inductive premise. But none of this, of course, disqualifies deductive logic as the framework within which one ought to work out the details of the argument.

The same situation holds with respect to Subjective Bayesianism. The fact that we cannot axiomatically prove that NMA underwrites a high probability for our successful (particular) theories, does not imply that we cannot express it in a probabilistic form – by invoking the grounds that the general framework is non-ampliative. All we have to do is insert an inductive premise, exactly as in the case of deductive logic, and this can, of course, easily be done. This inductive premise is none other than an assertion about the *prior probability* of the hypothesis in question. What allows for amplification in the Bayesian model is the value of the prior probability one will affirm and which enables the evidence to boost the posterior of the hypothesis appropriately. This being the case (which Psillos hints at but does not insist on), there seems to be no reason why Bayesianism cannot accommodate quite legitimately the NMA despite its non-ampliative character. This need not imply, of course, that Bayesianism also provides the means to *justify* the claim that the prior probability of some one hypothesis ought to be 'reasonably high'; this is a separate issue. And although it surely is the all-important one in the debate, it is also logically independent of the false claim that Bayesianism cannot capture the ampliative character of the NMA.

3.3 Conclusion

There is, I have argued, no reason to think that the enterprise of trying to find a probabilistic reconstruction of the NMA is a fundamentally misguided one. The final option available to the incompatibilist would be to deny the very fundamental idea of Bayesianism that partial belief is probability. To the best of my knowledge, no one has seriously entertained this idea, especially in the face of the foundational merits of Bayesian Confirmation Theory, as outlined in chapter 1. It seems safe to conclude, then, that none of the challenges examined in this chapter succeeds in showing that Bayesianism is in some fundamental way incompatible with the NMA. This in turn invites us to consider what a proper probabilistic rendition would look like, an issue I tackle immediately in the next chapter.

Chapter 4

Bayesianism and the No-Miracles Argument II – The Prospects of Compatibilism

Having rebutted claims of incompatibility I now address the task of showing how a Bayesian reconstruction of the argument is in fact possible and also explaining the philosophical insight that is provided by such a reconstruction. In this chapter I explain how *both* versions of the NMA examined earlier can be reconstructed probabilistically. The main theses I shall put forward are: a) Bayesian Confirmation Theory can be seen as a general framework within which the standard informal arguments become transparent; b) Subjective Bayesianism in particular emerges as the most suitable framework of analysis, but c) in doing so, it undermines the explanationist understanding of the NMA, while at the same time gives credence to its plausibility counterpart; and, finally, d) although such a Subjective Bayesian reconstruction of the NMA commits one to no particular epistemic attitude towards our best scientific theories, it brings out vividly the assumptions each such attitude tacitly presupposes. The question of whether the price of this construal of the NMA is the endorsement of a form of relativism is considered in chapter 5.

4.1 Bayesian Reconstructions of the NMA

4.1.1 The Explanationist Version

The first attempts to investigate possible ways of casting the NMA in Bayesian terms were made by defenders of the argument who took it to be an instance of IBE. This is explained by the fact that these attempts were part of the general reaction to van Fraassen's charge (examined in the previous chapter) that IBE is probabilistically incoherent. This incorporation into mainly defensive responses to van Fraassen's views, however, has meant that these Bayesian renditions of IBE and NMA were fragmentary in nature, typically confined to a few scattered remarks in different pieces in the literature[92].

Peter Lipton's (2004) contains the first *sustained* attempt to show how IBE can be incorporated within the Bayesian framework. It should be noted at the outset that Lipton presents his case as it concerns the general inference pattern of IBE, rather than the NMA as a particular instantiation of that pattern. This is not accidental, of course, since his book's main focus is IBE rather than the NMA. Still, one can easily apply his proposed reconciliation of IBE and Bayesianism to the explanationist NMA on the grounds that the former is standardly taken to instantiate IBE. This is *not* to say, however, that Lipton also embraces the explanationist NMA. In fact, he is quite explicit that the NMA *qua* an over-arching

[92] The most important papers dealing with the Bayesian rendition of IBE are Day and Kincaid (1994) and, especially, Okasha (2000).

inference to the best explanation, distinct from the first-order inferences to particular theories, is *unsuccessful* (cf. 2004, 192ff; 2005, 354). In the previous chapter I, too, suggested that the distinction between first- and second-order IBEs is dubious but added that one can focus on the first-order inferences to particular theories and re-run the argument on a 'particularist' basis. Lipton would surely be more favourably inclined to those particular versions of the explanationist NMA, given that his own case for realism focuses on those first-order inferences to particular theories. At any rate, what follows is not intended to attribute to Lipton beliefs he does not possess; rather, it is an attempt to bring his more systematic approach to bear on the question of how to reconstruct the argument from a Bayesian standpoint.

Before coming to the details of Lipton's proposal we need to look into some more features of IBE. As I emphasised in the previous chapter, IBE is considered by its defenders to be an independent rule of inference. Its focal point is the notion of explanation, which is intimately connected with the notion of understanding (cf. Friedman 1974, 6). IBE, then, is identified by its defenders as a mode of inference which picks out of many competing hypotheses the one "which would, if correct, be the most explanatory or provide the most understanding" (Lipton 2004, 59). This feature of IBE is emphasised frequently in the literature (cf. Psillos 2002, 617) because it is tantamount to the claim that IBE describes an independent mode of inference. In particular, special care is taken so that IBE is *not* understood as inference to the likeliest or most probable explanation. In Lipton's own words:

> "[T]he more we must appeal to likeliness [or probability] analysed in non-explanatory terms to produce a defensible version of Inference to the Best Explanation, the less interesting that model is. Conversely, the more use we

> can make of the explanatory virtues, the closer we will come to fulfilling the exciting promise of Inference to the Best Explanation, of showing how explanatory considerations are our guide to truth" (2004, 62).

Of course there is still a link with truth. If it were not there, it could hardly be maintained that IBE is reliable. Hence, it is presumed that "loveliness [or the quality of explanation] and likeliness [or probability] will tend to go together, and indeed loveliness will be a guide to likeliness" (ibid. 61). In other words, the quality of explanation serves as a symptom of truth and that is why it is claimed that, in the end, both IBE and methods concerned solely with the probability of hypotheses will largely be in agreement.

The dominant approach to inductive reasoning from a probabilistic point of view is, of course, Bayesianism. In Lipton's idiom, Bayesianism assesses the 'likeliness' of a hypothesis, i.e. its probability. It follows from our discussion of IBE a little earlier that IBE and Bayesianism are two distinct reconstructions of inductive inference, not to be conflated with each other. In the previous chapter I denied that the two are incompatible. The question which arises, then, is how they relate and what consequences this relation has for NMA.

Lipton expresses the main aspect of what he takes to be the relationship between Bayesianism and IBE as follows:

> "Bayesianism and Inference to the Best Explanation are broadly compatible...[In fact], not only are [they] compatible but... they are [also] complementary. Bayesian conditionalisation can indeed be an engine of inference, but it is run in part on explanationist tracks" (ibid. 106-107).

Put this way, it is not entirely clear how exactly IBE complements Bayesianism. In fact, the text reads as if IBE *constrains* Bayesianism in certain respects. A little later, though, we get the opposite impression. This is Lipton again, elaborating on his previous quote:

> "One way of putting [my proposal] is that explanatory considerations provide a central heuristic we use to follow the process of conditionalisation, a heuristic we need because we are not very good at making the probabilistic calculations directly" (ibid. 106).

Hence, it seems that IBE is here taken to be something like an auxiliary device to probabilistic reasoning aimed at everyday practice due to ease of implementation rather than a foundationally sound and independent mode of inductive inference. As we shall see shortly, deciding on this issue is crucial for assessing the prospects of the explanationist NMA.

In attempting to build a bridge between Bayesianism and IBE, Lipton singles out the following four elements which he believes link the two approaches: 1) explanatory considerations supply the values of both the prior probabilities and the likelihoods found on the right hand side of Bayes' theorem; 2) they also help us determine which pieces of evidence count as relevant for conditionalisation; 3) IBE sheds light on the context of discovery, whereas Bayesianism operates solely within the context of justification; and 4) explanation-based reasoning can be seen as a useful heuristic, which replaces quite abstract and cumbersome probabilistic reasoning.

Elements (2), (3) and (4) can indeed plausibly be seen as complementing Bayesian reasoning. The reason for this, though, is that they operate on different planes from Bayesianism, i.e. either within the context of discovery, as in (3), or on the descriptive level of everyday habits, as in (4), or finally, as auxiliaries to the inference process itself, as in (2). Bayesianism, nonetheless, is a theory of inference, a rational reconstruction of how people ought to reason. As such, its place is within the context of justification and the realm of normativity. Now, what makes the investigation of the relations between IBE and Bayesianism interesting is precisely the fact that IBE claims to be *normatively binding* as well. IBE is not intended simply as a faithful *description* of our inferential practices. Instead, it professes to strike a balance between amplification and epistemic warrant that other approaches to inductive reasoning fail to capture and that is why it is alleged to provide good normative reasons to follow its dictates. From the normative point of view, the only element which carries any real significance is (1), that is, how explanatory considerations influence the determination of the terms on the right hand side of Bayes' theorem.

In claiming that explanatory considerations determine the values of the prior probability and the likelihoods in Bayes' theorem, Lipton follows Okasha's proposal, analysed in his (2000, 702-704)[93]. In that paper, Okasha suggested that if a theory T is a better explanation of the evidence E than T', then either $\Pr(T)>\Pr(T')$ or $\Pr(E/T)>\Pr(E/T')$, or both. Given that realists typically argue for *approximate* rather than total truth, however, we should extend Okasha's reconstruction as follows: let H stand for 'T is approximately true' and H' for 'T' is approximately true', as introduced in chapter 3. Okasha's proposal would now suggest that if a theory T is a better explanation of the evidence E than T', then either $\Pr(H)>\Pr(H')$ or $\Pr(E/H)>\Pr(E/H')$, or both.

As we saw also in the previous chapter when discussing the 'Base-Rate Fallacy', it is quite natural to identify at least some aspects of explanation with the likelihoods in Bayes' theorem. *Prima facie* it may seem that the likelihoods are

[93] A similar picture arises out of Day and Kincaid's (1994, 285-286) (more sketchy) analysis.

occasionally easy to determine. In case an explanatory theory T deductively implies the evidence E, $\Pr(E/T) = 1$ and, hence, the explanation is as good as it can get. In discussing the issue of realism, however, we are typically interested in arguing that T is probably 'approximately true' rather than fully true. Letting H stand, as before, for 'T is approximately true', the question arises whether when T deductively entails E, H also entails E. Admittedly, this is a very hard question to answer unequivocally in the absence of a precise notion of 'approximate truth'. It seems that realists want to argue that, intuitively at least, it is quite plausible to maintain that when T entails E, then $\Pr(E/H) \approx 1$. This contention, nonetheless, can be doubted. Though there is clearly a problem here for further investigation, for our purposes in this thesis I shall grant the intuitive pull of the realists' case and accept that when T entails E, $\Pr(E/H) \approx 1$.

Moreover, there are cases where, even though T does not entail E, "accepted" statistical theories (e.g. those of statistical mechanics) fix $\Pr(E/T)$ (and so arguably $\Pr(E/H)$) unequivocally. The same cannot, of course, be said for the 'false positive' rate $\Pr(E/\neg H)$, which is also relevant in the determination of a hypothesis' posterior. Here, one might indeed seemingly plausibly claim that explanatory considerations of some sort relating to the quality of the explanation are involved. Of course, we know from the 'Base-Rate Fallacy' objection that even if the likelihoods can be unequivocally determined, they alone do not suffice to determine posterior probabilities, so that we are still unable to infer anything with respect to the (approximate) truth of our theories.

IBE and NMA (as an instantiation of it) though, focus on (approximate) truth, which in turn means that assumptions about prior probabilities need to be made as well. It might be thought that resorting to prior probabilities is indeed a natural way to codify the so-called *a priori* explanatory virtues of scientific theories. Such virtues are standardly taken to include our theories' ability to unify seemingly diverse phenomena under the same underlying theoretical principles (thus simplifying our world-picture), and their relations, deductive or inductive, to other highly confirmed theories. These virtues seem to reflect global features of scientific theories, independently of the particular piece of evidence they are called forth to explain. Lipton's suggestion, then, is to take explanatory considerations as a means to evaluate these *a priori* virtues in terms of understanding, and on this basis assign the relevant prior probabilities (Lipton 2004, 113)[94]. In this way, not only do we avoid the 'Base-Rate Fallacy', but we do it on the basis of the very explanatory considerations IBE is preaching. Some aspects of what makes a theory explanatory are reflected in the likelihoods, others in the prior probabilities. Although he would seem to have a better case concerning the priors than concerning the likelihoods, the crucial point of Lipton's proposal is that what determine the assignment of whatever probabilities we need are explanatory considerations.

On this approach, the explanationist NMA, then, can be reformulated as follows: explanatory considerations support the non-miraculous conclusion that our (individual) best theories are (approximately) true on the basis of creating the rationale for 1) a low Bayes factor, which, informally speaking, captures the intuition that it would be a miracle if our theory were (wholly) false but predictively successful, and 2) a relatively high prior probability, which reflects the 'miracle' of theories characterised by the aforementioned *a priori* virtues being (wholly) false.

[94] In fact, Wesley Salmon in his (1990, 283-285) had already suggested the association of these *a priori* virtues with the prior probabilities in Bayes' theorem.

This might sound plausible but there are in fact some serious problems. In assessing Lipton's proposal, we have to distinguish between two cases: a) IBE and explanation as a *descriptive* account of how people reason in real life (tacitly or explicitly) in determining the values of the prior probability and the likelihoods in Bayes' theorem, b) IBE as a *prescriptive* constraint on prior probabilities and likelihoods. It might very well be the case that people do think in terms of some (rather vague) notion of explanation in engaging in inductive reasoning in their everyday affairs and that scientists and even philosophers do the same when they try to codify NMA in terms of probabilities. If so, then IBE becomes a contribution to the investigation of human psychology and actual mechanisms of thought. Construed this way, however, there is no normative import of the sort the defenders of IBE profess. Recall that IBE is supposed to be an independent and *reliable* rule of inference. The only way to show this is to present an argument that explanatory considerations of the sort presented above suffice to *uniquely fix* the prior probability in Bayes' theorem to a sufficiently high, and the Bayes factor to a sufficiently low, degree. Only then would we be warranted to infer to the (approximate) truth of our theories through the explanationist NMA. And clearly it is quite possible that descriptively speaking some agents may fail to assign such values to the priors and to the Bayes factor. The reliability claim can only be defended on *normative* grounds.

Surprisingly enough, neither Lipton nor Okasha have anything to say in defence of the normative reading of the relation between IBE and Bayesianism. Lipton in particular is adamant that his purposes are only *descriptive*:

"My suggestion is that explanatory considerations of the sort to which Inference to the Best Explanation appeals are often more accessible than

> those probabilistic principles to the inquirer on the street or in the laboratory, and provide an effective surrogate for certain components of the Bayesian calculation. On this proposal, the resulting transition of probabilities in the face of new evidence might well be just as the Bayesian says, but the process that actually brings it about is explanationist" (ibid. 114).

It appears that Lipton intends explanationism as an auxiliary device for Bayesian reasoning, one that allows the layman to reason, while at the same time relieving him of excessive computational burdens. But if this is how we ought to understand the relationship between Bayesianism and IBE, then one wonders how the explanationist can still uphold the claim that IBE is a distinct and fundamental mode of inductive inference.

Okasha, on the other hand, does not take a stance on the issue, thinking it "a fundamental and unresolved question ...whether the Bayesians are *explaining*, or just *representing* [IBE]" (2000, 706)[95]. If the Bayesians only represent IBE, then they admit that IBE plays a more fundamental role than Bayesian techniques. On this reading IBE presumably constrains Bayesianism by fixing the prior probabilities. If, on the other hand, they explain IBE, then they "deny that there *is* a more fundamental sort of inductive reasoning [than the Bayesian apparatus]" (ibid. 706)[96]. Though Okasha refuses to take sides, it remains true that if he intends IBE as a normatively binding and distinct mode of inference, he has to join those who think that Bayesianism only represents IBE.

The moral of the discussion so far, then, is that the explanationist NMA can be given a probabilistic representation faithful to the claims of objectivity IBE

[95] Emphasis in the original.

[96] Emphasis in the original.

raises only if one can provide a convincing argument to the effect that explanatory considerations somehow impose normative constraints on our assignment of prior probabilities and likelihoods to our best theories. We have seen that Lipton's and Okasha's approaches fall short of this requirement. It is quite striking, though, that no recent defender of the explanationist NMA has seriously considered Shimony's account of tempered personalism, introduced in chapter 1, as a possible way to bring together the normative import of *this particular understanding of NMA* and Bayesianism. As we shall see next, Shimony's proposal, although put forth much earlier than the explanationist NMA, seems *prima facie* perfectly tailored to it.

4.1.2 Tempered Personalism and the Explanationist NMA

As we saw in chapter 1, Abner Shimony's tempered personalism endorses a mild constraint on the prior probability of a hypothesis. More specifically, if a scientific hypothesis has been 'seriously proposed', it deserves a prior probability sufficiently high so that convergence of opinion can occur within a 'reasonable amount' of time. Although the exact value of the prior is not (and, presumably, cannot be) specified, prior probabilities are not subject solely to the constraints of the axioms of probability.

Evidently, much rests on what a 'seriously proposed hypothesis' is, on what a 'reasonable amount of time' is and how these can be identified in particular circumstances. Shimony tried to avoid outright subjectivism in the identification of such 'serious hypotheses', proposing "to formulate some methodologically sensible guidelines for decisions on this question" (Shimony 1970, 110). Such guidelines include the clarity of exposition of a scientific theory and the amount of 'intellectual freshness' it brings to the community of experts. He refused, nonetheless, to turn these guidelines into a sharp set of conditions "because of the danger of arbitrariness and of diminishing the flexibility of the scientific method" (ibid. 110). Consequently, some amount of subjective judgement is preserved in Shimony's account. Furthermore, as we noted in chapter 1, the situation is similar with respect to what a 'reasonable amount of time' is. Shimony never specified the form or duration of the process, which would produce his 'envisaged observations'. Presumably a fair amount of subjective judgement would also have to be involved in specifying and conducting the research required to yield the necessary observations, so that, ultimately, convergence of opinion can take place. It is arguable, however, that in today's science at least we are frequently in possession of quite strong intuitions regarding what constitutes serious science and how it should be conducted. Grant this and the main issue becomes that of the justification of the tempering condition.

This is the potential meeting point between Shimony's approach and the explanationist NMA. Recall the main line of defence on behalf of the latter: NMA instantiates IBE and IBE is shown to be reliable by the success of NMA. This circle has been declared non-vicious by the explanationists and this is why they claim that the explanationist NMA has normative force. Shimony, on the other hand, has offered a similarly circular argument in his attempt to justify his proposed constraint on the priors. The tempering condition allows the posterior probabilities of 'seriously proposed hypotheses' to converge quickly. In its turn, the tempering condition is itself justified on the basis of the presumption that these 'seriously proposed hypotheses' which our science regularly produces somehow track the truth and, hence, deserve a higher prior probability. Shimony's grounds for this last claim seem to rest on the evolutionary success of scientific and everyday induction, thus making his reasoning circular. Much like the explanationists, Shimony declares this circularity harmless. Here is his rationale:

> "I claim, however, that the circularity is non-vicious in the following sense: *the theory as a whole is open to critical evaluation in the light of experience, for the reciprocal support of a methodology and a scientific world picture does not render it impregnable to criticism*" (ibid. 159-160)[97].

Leaving the viability of such proposals aside for the moment, Shimony's argument provides a hint about how to codify the explanationist NMA in Bayesian terms in a way that fills the normative void in Lipton's approach. This void, remember, consisted in the absence of an argument showing how explanatory considerations *constrain* the 'reasonable' values of the prior probabilities and the likelihood $Pr(E/\neg H)$. As we shall see in a moment, the explanationist can invoke general epistemological considerations in order to determine the value for $Pr(E/\neg H)$. Consequently, the only open question concerns the way that he can also constrain the values of the prior probabilities. In order to begin answering this last question we must keep in mind that the normative import of the *informal* explanationist NMA relies on an (allegedly) non-viciously circular argument for the approximate truth of a theory *T* on the basis of its empirical success. The contribution of Shimony's proposal, then, lies in the fact that it shows that circularity of this alleged non-vicious sort *can* be captured by the Bayesian framework. Consequently, the explanationist only needs to formulate an (allegedly) non-vicious circular argument with reference to the prior probabilities and show how such an argument introduces the necessary constraints for normativity.

[97] Emphasis in the original.

Before we come to the priors, however, we must first see how general epistemological considerations provide reasons in favour of a low value for $\Pr(E/\neg H)$. Those reasons derive from the explanationist's decision to adopt an empiricist epistemological standpoint and seek an explanation for science's *empirical success*. Empirical success cries out for explanation only because it carries considerable epistemic significance. Its attainment is thought to be a remarkable feat and, hence, a valuable source of information regarding the epistemic merit of our theories. This implies, however, that empirical success is also thought to be a virtue that is very hard to reach. Recall from chapter 3 that the main worry for NMA comes from the possibility that our empirically successful theories are (radically) false and that, in due course, other empirically successful theories, hitherto unknown, will replace them. Yet, if it were fairly uncontroversial that all (or even most) of the logically possible future competitors to our theories will be empirically successful, that would amount to saying that empirical success is so easy to achieve that it can hardly be of any exceptional epistemic value to us in the first place.

We can express this thought more formally as follows: by the probability calculus we know that $\Pr(E/\neg H) = \frac{1}{\Pr(\neg H)}\sum_i \Pr(H_i) \times \Pr(E / H_i)$, where H_i refers to the (mutually exclusive) hypotheses whose union set is equivalent to $\{\neg H\}$. It has just been argued that adopting the empiricist standpoint amounts to maintaining that a theory's achieving empirical success is not a very frequent phenomenon of scientific life. This, however, is equivalent to saying that most of the future competitors to current successful theories will, in fact, fail to be empirically successful. Hence, the explanationist is free to think that $\Pr(E/H_i)$ is low for most of the possible future competitors to our theories comprising $\{\neg H\}$. To conclude that $\Pr(E/\neg H)$ will also be low we need only further assume that the sum of the prior probabilities of the few (mutually exclusive) *successful* theories belonging in $\{\neg H\}$ is strictly smaller than $\Pr(\neg H)$. This, however, is an assumption that we can grant the explanationist. Indeed, denying this assumption amounts to assigning to the few successful but still unknown theories a significantly higher prior probability than to the many unsuccessful ones. But, surely, given that $\{\neg H\}$ is partitioned by theories hitherto *unknown*, there is hardly any reason for such a differential treatment. Consequently, the explanationist may reasonably assume $\Pr(E/\neg H)$ to be low.

Let's now move on to the prior probabilities for our current theories. Here there are no independent considerations deriving from the adoption of an empiricist epistemology to help us. For this reason, the explanationist must allude to what he thinks is the main source of normative import for the informal version of his NMA, namely (allegedly) non-vicious circularity. Shimony's 'tempered personalism' showed us that this sort of circularity *can* be accommodated within the Bayesian framework. Here, then, is how the explanationist may attempt to constrain prior probabilities through the invocation of (allegedly non-vicious) rule-circularity:

A Bayesian-explanationist NMA

Assign theories proposed through IBE non-negligible priors. Empirical success will eventually raise their posterior probability to a degree that will allow us to hold that they are more likely to be (approximately) true than not. Since the posterior is now high enough, IBE is probably reliable. Hence, our initial assignment of a non-negligible prior is justified on the grounds that IBE has been shown probably reliable.

This Bayesian-explanationist NMA is as rule-circular as the original informal NMA. IBE is used essentially in the determination of the prior probability, since it becomes the criterion for determining which hypotheses will be assigned non-negligible priors. In this way it influences crucially the posterior of the hypotheses, whose value the reliability of IBE itself depends on. Its success, therefore, also depends on the legitimacy of rule-circular arguments in general. What we have shown so far, though, is that within Bayesianism *there is* room for the formulation of such arguments. Hence, not only are there good reasons to doubt the claims of the incompatibilists, but we now also have a positive argument for compatibilism.

This is not to say, however, that rule-circular arguments are, in fact, legitimate. In chapter 3 I examined some serious problems regarding their prospects. The main worry such arguments have to face refers to their *professed* property not to include their conclusion in their premises. I now wish to argue, *in tandem* with all those critical of such arguments[98], that one *cannot* legitimately use a rule of inference without prior knowledge of its reliability. In the case of the explanationist NMA this means that one is *not* entitled to use IBE without knowledge of its reliability. Here is the reason. The explanationist claims that NMA shows IBE to be reliable. NMA, though, already *uses* IBE. Hence, the explanationist ends up supporting the following paradoxical schema: we claim *knowledge* of the fact that science tracks truth on the basis of a rule of inference we use *without knowing* whether it is reliable or not in the first place! This schema, I

[98] See, for example, Howson (2000, ch. 2), Worrall (1999, 2000b), Musgrave (2006, 314-315).

think, is truly paradoxical. NMA can only go through if we *know* that IBE is reliable. And if only the conclusion of NMA can establish the reliability of IBE, then the (approximate) truth of our theories must be a premise in the argument.

This line of reasoning makes rule-circular arguments as viciously circular as it gets. The explanationists will of course respond that, on their conception, all we need is that it be objectively the case that our rule of inference is reliable regardless of our knowledge. This response, nonetheless, merely retains the mystery of inferring knowledge from ignorance, since the knowledge of whether our theories are approximately true is inferred by means of a rule of inference whose credentials are unknown to us[99].

These general doubts on the possibility of benign circularity apply equally well to the Bayesian-explanationist NMA. Recall, this argument is structurally identical to the informal NMA. Hence, it is not permissible to assign non-negligible priors to theories invented through IBE on the basis of its reliability, since its reliability hinges on the posterior probability the prior itself gives rise to. To fix the prior we need independent reasons of a sort circularity is unable to provide. If that is so, then Shimony's own defence of harmless circularity, mentioned earlier, fails. Recall his main idea: despite the 'reciprocal support of a methodology and a scientific world picture', scientific theories are still open to criticism and 'critical evaluation in the light of experience'. If my criticism of rule-circularity is well-

[99] Peter Lipton seems almost to concur, despite the fact that he clearly does not intend to. Towards the end of his discussion of the circularity objection to the reliabilist version of NMA he writes: "So I conclude that, while the [explanationist] miracle argument is no argument against the inductive sceptic or the instrumentalist, the circularity objection does not show that realists are not entitled to use it. The argument is circular against non-realists, but not for realists themselves" (2004, 192). But, surely, if one is already a realist, he needs no further arguments to turn into one!

founded, though, it remains mysterious how a thesis supported by a circular argument is open to criticism in the way that Shimony asserts. It simply can't be that the explicit inclusion of an argument's conclusion among its premises leaves the argument open to criticism.

So far, then, I have argued 1) that a Bayesian formulation of the explanationist NMA, which captures all its essentials, is possible, and 2) that the explanationist NMA, no matter how it is formulated, faces insuperable difficulties. (1) is both an amendment to Lipton's normatively weak account and a counterexample to the incompatibilist claim advanced by Psillos (2004) and discussed in the previous chapter, i.e. that the explanationist NMA ought not to be modelled in Bayesian terms if it is to be successful. It was shown that the decision to formulate the argument probabilistically has nothing to do with its success, since our Bayesian reconstruction misses none of its essential elements. It was further maintained, nonetheless, that explanationism fails on independent grounds, i.e. on the basis of its inability to yield a satisfactory defence of the epistemic claim that we have good reasons to think that science yields approximately true theories.

One final remark: I have argued that the informal explanationist NMA and its Bayesian counterpart stand or fall together. I have also noted that it is very surprising that none of the explanationists has seriously considered the prospects of a Shimony-type Bayesian explanationist NMA. The reason for this neglect can perhaps be traced to the fact that, when probabilistically formulated, the circularity involved in the argument becomes too obvious for the argument to be taken seriously. Hence, Psillos' strong remark that "an objectivisation of Bayesianism is ... something that we, presumably, know cannot be done" (Psillos 2004, 88). Incompatibilists have tried to salvage the argument by renouncing probabilistic talk altogether. It has been shown, nonetheless, that Bayesianism should be the least of their concerns. The Bayesian framework is flexible enough to capture the main message of all strands of epistemology. It is also rigorous enough to expose (at least some of) their defects. The explanationist NMA is a case in hand. A consistent explanationist ought either to defend or reject both formulations of the argument. There were already good reasons for rejection based solely on the analysis of the informal version. A careful analysis and contraposition to it of the Bayesian version makes these reasons even more evident.

4.1.3 Can Frequencies Help Us Out?

Rejecting the explanationist NMA on the basis of our inability to justify non-trivial constraints on prior probabilities through explanatory considerations is bad news not only for explanationism but also for those with high hopes for the NMA. In its explanationist version the NMA is intended as an argument with heavy normative import, a feature which also explains its appeal. Is there perhaps some other way to constrain the priors in order to bring these hopes to fruition? It has long been known that the Principle of Indifference faces important conceptual obstacles, already discussed in chapter 1 – sufficiently many, I would say, so as not to merit any further discussion. It remains to be seen whether Empirically-based Subjective Bayesianism or Objective Bayesianism, also discussed in chapter 1, can offer a satisfactory answer.

Both these theories propose to introduce further objective constraints into Bayesianism with the help of empirical data in the form of fairly reliable statistical information. Hence, fixing the prior probability of a theory's being (approximately) true means that we would need information with respect to the frequency of the occurrence of (approximately) true theories among (some specification of) the relevant population of scientific theories. But since this is the objective of NMA itself, i.e. knowledge of which theories can be justifiably considered to be approximately true, such estimates cannot be had. Magnus and Callender correctly note in this connection that "if we had [independent grounds for thinking that a theory is very likely true], the no-miracles argument would be superfluous" (2004, 328). Indeed, obtaining reliable estimates regarding the frequency of the occurrence of approximately true theories means that a method is available, which allows us to reliably judge whether a given theory is approximately true or not. But surely, *if* there is such a method, the obvious thing to do would be to apply this very method on our current theories in order to find out whether they are (approximately) true rather than invoke the NMA.

The only kind of information available to us, through which we might attempt to constrain the priors, is statistical information about the predictive success of past and present "accepted" theories. This proposal originates in Salmon (1990), who claims that "[prior probabilities] can be understood as our best estimates of the frequencies with which certain kinds of hypotheses succeed" (270). The use of such information to fix the priors, however, would surely turn the NMA into a *non-sequitur*. What the NMA attempts to do is somehow allow us to infer (approximate) truth from predictive success. By letting information regarding the frequencies of success of certain theories to fix the prior probability of (approximate) truth, though, we are implicitly presupposing that predictive success is, in fact, a reliable indicator of truth, i.e. precisely what the NMA aspires to conclude[100].

[100] Similar critical considerations to the ones in the main text can be found in Howson (2000, 46-47) and Worrall (forthcoming).

Even if such information were allowed (somehow) to fix the priors, though, there would still be a remaining issue: 'the reference-class problem'. Recall from chapter 1 that frequencies require a collective, and there are many collectives that any given event might be considered to be an element of. Consequently, many different collectives can serve as the basis for measuring the frequency of empirically successful theories, and a question naturally arises as to which one is the most appropriate. As it turns out, this question is extremely hard to answer.

To begin with, consider the collective including *all possible rivals* of a given theory *T*. The problem with this specification is that the resulting collective seems to include the (non-denumerable) infinity of all the 'grue-like' variants of *T*, which makes the determination of the relevant frequency mathematically impossible. Another option would be to constrain attention to all rivals of *T actually* articulated in the history of science so far. This proposal, however, is ambiguous as it stands. Are we to include all theories examined under the vague heading 'history of science' or only the ones classified as 'mature science'? If we opt for the first alternative, we run the danger of underestimating our chances due to the large number of predictively unsuccessful 'wild guesses' articulated in the course of history. Even if we opt for the second alternative, however, things are no better, since theories falling under the label 'mature science' are by definition empirically successful (cf. Magnus and Callender 2004, 326). We would then reach the other extreme, i.e. that of overestimating our chances. Hence, even if frequencies of success could somehow fix prior probabilities, the task of selecting the reference-class these frequencies necessarily assume in a meaningful and, at the same time, *non question-begging* way seems insurmountable indeed.

It seems fair to conclude, then, that the prospects for justifying non-trivial constraints on the prior probability of a theory being (approximately) true through frequency information are also very dim[101]. This in turn means (in the absence of some further and so-far-unarticulated argument) that the sole available interpretational context that remains for one who seeks a probabilistic reconstruction for NMA is Subjective Bayesianism. Indeed, the impossibility of constraining prior probabilities through frequency information means that Empirically-based Subjective Bayesianism in particular fails to add anything to pure personalism with respect to this particular issue. So, is NMA in any way compatible with Bayesian personalism? And if so, what does this mean for NMA?

4.1.4 NMA as a Plausibility Argument

Our discussion of explanationism a little earlier led us to the conclusion that the explanationist way of objectifying prior probabilities is impermissible. Consequently, the explanationist's attempt to move beyond Subjective Bayesianism and towards a more objective account has failed too. It appears, though, that the alternative understanding of NMA as a plausibility argument is much more within the spirit of Subjective Bayesianism. Recall the crux of that version of the argument: our 'inductive intuition' somehow assures us that the empirical success of our best theories is not to be attributed to chance, despite the fact that, in principle at least, one cannot rule out the possibility that future theories, radically different from ours, will eventually overthrow them and take their place. Thus

[101] My argument is confined to the context of NMA. All too frequently we are in possession of statistical data that it would be too far-fetched not to call reliable. Hence, there are cases one might reasonably claim that the quest to objective priors is in fact successful.

stated, it becomes obvious that empirical success does not suffice alone to establish the realist thesis. In fact, this is nothing other than the informal counterpart of the 'Base-Rate Fallacy'. What is required is a presumption in favour of *our* theories, which is afforded by our 'inductive intuition' on the basis of what I earlier called the *a priori* virtues of our theories (simplicity, fruitfulness etc.). This presumption, though, claim proponents of the plausibility version, is not founded on any argument, for any such argument would require further non-trivial premises and so on *ad infinitum*. The only way to break the regress is by taking a stance 'dogmatically'. Our strong intuitions on this issue 'justify' the realist stance.

It seems quite convincing to suggest that Subjective Bayesianism in particular *is* in a position to convey the fundamental idea of the plausibility understanding of the NMA. Our optimistic intuition towards our theories is captured by assigning relatively high prior probabilities to the approximate truth of our (particular) theories and, correspondingly, low priors to that of the as-yet-unknown but always possible future challengers. These prior probabilities are, as Colin Howson has put it, "[highly subjective estimates] of how probable types of large-scale world are" (2000, 55). There is no further argument that can be given in their favour. But, similarly, there is no further argument which can be given to support the dictates of our 'inductive intuition'.

A natural reaction to this line of reasoning is that it makes the NMA essentially powerless. This worry revolves, I think, not so much around Subjective Bayesianism[102] as it does around the notion of 'inductive intuition'. How can such a vague and intuitive notion form the basis of the substantive claim the NMA seeks

[102] Many people think that there is also a problem with Subjective Bayesianism. I deal with this problem in the next chapter in detail.

to support? The only available answer to this question is that, although this version of the argument might indeed be an unsatisfactory defence of realism when compared to the aspirations of the committed realist, it still remains the only defensible construal of NMA. John Worrall (forthcoming) has argued persuasively that the NMA *qua* a plausibility argument merely sets realism as the 'default position' in the debate. No more than this can be expected from the argument, since NMA was never intended as a *demonstration* of the truth of realism. On the contrary, being inductive in character, it rests on certain presuppositions which, on pain of infinite regress, at some point have to be taken 'on faith'. Indeed, having already seen the inadequacy of trying to justify these assumptions on the basis of allegedly non-viciously circular arguments, retorting to intuition remains the only open option. Furthermore, this situation is structurally no different from any content-increasing inference we perform in our everyday lives, and for whose legitimacy we are seldom in doubt. If so, then there remains no reason to be suspicious of the workings of intuition in scientific matters either[103]. Worrall concludes:

> "The intuitions set some form of realism as the default position: there is no more to the 'argument' than that. The realist might like to say something stronger, but there is nothing stronger to say that anyone should like" (ibid.).

This 'default position', I have further claimed, can be faithfully reconstructed within Subjective Bayesianism and takes the form of subjective prior probabilities. Without them, empirical success has no bite and the argument has no impact. The nature of 'inductive intuition', though, is no different from the nature of subjective

[103] A fuller assessment and defence of the normative weight of the 'intuitive' NMA is contained in chapter 5.

priors: both are primitive and cannot be argued for any further. Worrall seems to concur with this, saying:

> "This assignment of a prior is, of course, on the personalist Bayesian approach, to be thought of as simply reflecting a personal judgement about the plausibility of the theory, and not as any sort of reflection of some 'objective chance' that the theory is true... Of course, as will be very clear, I share this intuition ... ." (ibid.).

This concludes the argument for points (a), (b) and (c) stated in the beginning of this chapter, i.e. that Bayesian Confirmation Theory broadly understood makes both versions of the argument transparent (and, thus, supports a *compatibilist* stance) and that Subjective Bayesianism in particular, being the most suitable interpretational context, vindicates the plausibility version of NMA rather than its explanationist counterpart. If all this is true, should one be a realist or not?

4.2 Adjudicating the Competitors

4.2.1 Can the NMA be Resisted?

The tenability of realism centrally relies on the success of the NMA. Having already argued that only the plausibility variant delivers its (modest) promise, we have to examine the issue of whether there are good grounds for denying the dictates of our intuition. *Prima facie* it might seem that no further argument is possible, since the validity of the content of our 'inductive intuition' is to be decided on 'dogmatic' grounds. It might seem, that is, that the realist simply 'decides' about the validity of his intuition and asserts the conclusion of the argument, while the anti-realist merely disagrees. Put this way, the realist and the anti-realist end up talking past each other without any prospects for mutual understanding and fruitful debate.

Although this picture is *to some extent* an accurate depiction of the situation, there are still kinds of considerations which reasonably influence one's evaluation of his intuition. One such consideration has been put forth by van Fraassen (1980) and forms the backbone of his constructive empiricism. The reason, van Fraassen claims, which allows us infer only to the empirical adequacy of our theories rather than their (approximate) truth is that our 'inductive intuition' is reliable only with respect to the observable realm. In other words, there is a difference in kind between the observable and unobservable domains of the world, which appears to be all important when it comes to our content-increasing inferences. This in Bayesian terms means that there is a relatively high prior probability that a given theory is empirically adequate and a low prior probability that it is actually (approximately) true on the basis of the qualitative difference between what counts as observable and what as unobservable.

Arguably, such a consideration might influence our evaluation of the reach of our 'inductive intuition'. This consideration, however, faces sharp objections. The standard realist response, for example, is to deny that there is a meaningful observable vs. unobservable distinction. What counts as observable and what as unobservable, we are told, forms a continuum, such that any sharp line across it is inevitably philosophically ill-motivated (cf. Maxwell 1962; Psillos 1999, 193-200). But this response is not without its own problems. On the one hand, the dichotomy is intuitively evident, despite the existence of some limiting cases which are by nature more problematic[104]. On the other hand, it is unclear how denying this dichotomy will serve the positive aim of the realist to provide an argument for his thesis. Indeed, if there is no clear-cut line between the observable and the unobservable, then, in principle at least, the very basis of NMA, i.e. the *empirical* success of science, is not an uncontroversial issue any more. I doubt, though, whether the realist will welcome this consequence.

A more appropriate response to van Fraassen's challenge ought to proceed as follows. Even if we accept that there is a meaningful distinction between the observable and unobservable domains of reality, there is still reason to doubt that the reliability of our 'inductive intuition' fails to extend beyond the observable realm. Our inferences to the truth of theoretical claims are *structurally* no different from inferences we perform in our everyday lives for more mundane affairs. They are also no different from those we engage in at the purely empirical level, in order to arrive at van Fraassen-style empirical adequacy claims. After all, we learned from Goodman that underdetermination threatens the purely inductive generalisations that we all are inclined to make just as it threatens 'still more ampliative inferences' to universal theories. Both involve consideration of the *a priori* virtues of our hypotheses, like simplicity and fruitfulness, where 'inductive intuition' comes into play. Contrary to van Fraassen's claims, then, it seems natural to suppose that transferring an ampliative mode of inference, which has proved very successful so far in its applications to the observable realm, to that of the unobservable, should enhance, rather than decrease, our confidence in the

[104] Van Fraassen's own response to such limiting cases is that they only show 'observable' to be a vague predicate, not that they undermine the principled distinction between the observable and the unobservable (cf. 1980, 16).

(approximate) truth of the hypotheses regarding the latter. Consequently, the epistemic significance of the distinction between the observable and unobservable parts of the world seems insufficient on its own to turn the balance in favour of constructive empiricism and set this rather than realism as the default position[105].

However, the most celebrated argument in this discussion along with the NMA, namely the 'pessimistic induction' threatens exactly to drive a wedge between scientific inferences at the observational and at the theoretical level. Science seems to be cumulative at the empirical level, each theory-change adding to our predictive powers at the observational level, and yet, so the 'pessimistic induction' supposes, at the level of the 'deep structural' claims that our theories make we see radical change in the history of science. Consequently, the only sensible inductive inference at that theoretical level is that even our current theories are likely to be (radically) false. Indeed, this has been repeatedly presented as the main counter-argument to the NMA and, if successful, would be sufficient to remove realism from being the 'default position'. What would be put in its place, of course, is not easy to see. The 'pessimistic induction' aims merely at denying the force of the NMA and not at putting forward any alternative positive account of scientific knowledge. Hence, its success could be used to motivate constructive empiricism as much as it has been explicitly used to motivate Epistemic Structural Realism[106].

[105] Similar considerations can be found in Psillos (1999, 211-215), though with reference to IBE rather than 'inductive intuition'.

[106] It is quite striking how very little use van Fraassen has made of the 'pessimistic induction' throughout his writings. To the best of my knowledge, he refers explicitly to, and embraces, this argument (still without naming it) only in his (2006, 288ff.), on his way to criticising structural realism. One possible explanation of this continuous absence is his strong aversion towards

The 'pessimistic induction' aims to undercut the inductive strength of NMA through a list of alleged radical theory-changes, explicitly offered for the first time by Laudan (1981). These counter-examples to the realist inference consist of theories once considered (approximately) true on the basis of their empirical success, but which were subsequently abandoned as (radically) false. On this basis, the 'pessimistic induction' is usually reconstructed as a *reductio*[107] of the claim that there is a connection between empirical success and truth, a claim which much (but not all, as we have seen) of NMA rests on.

In an interesting recent paper, however, Peter Lewis (2001) suggested that a probabilistic reading of the argument shows that the 'pessimistic induction' is not a valid *reductio* of the connection between success and truth, since, like the NMA according to Howson, it also commits the 'Base-Rate Fallacy'. As a result, Lewis concludes, the realist is not threatened by the argument. Here is how he reconstructs the *reductio*:

(1) "Assume that the success of a theory is a reliable test for its truth.

(2) Most current scientific theories are successful.

(3) So most current scientific theories are true.

(4) Then most past scientific theories are false, since they differ from current theories in significant ways.

(5) Many of these false past theories were successful.

---

metaphysics. If any inference to anything unobservable (including structure) "[sinks] us into this metaphysical morass that swallows all seekers for the true foundations of being" (ibid. 303), one (presumably) hardly needs a historical argument from discontinuity to discredit (any version of) the realist thesis.

[107] For example see Psillos (1999, 102-103) and Lewis (2001).

(6) So the success of a theory is not in fact a reliable test for its truth" (ibid. 373) [108].

Lewis' strategy, very much like (but independently of) Howson's with reference to NMA, is to translate this argument into probabilistic terms. Hence, reliability is to be expressed in terms of the *false positive* $Pr(E/\neg H)$ and *false negative* rate $Pr(\neg E/H)$. Saying that 'the success of a theory is a reliable test for its truth', as the realist does, amounts to saying that these two probabilities are small. Lewis' ultimate aim is to show that Laudan's *reductio* in fact fails to use the history of theory-change in science to establish that these probabilities are, on the contrary, high.

Lewis extracts a value for the posterior probability $Pr(\neg H/E)$ from Laudan's assertion that "for every highly successful theory in the past of science which we now believe to be a genuinely referring theory, one could find half a dozen once successful theories which we now regard as substantially non-referring" (1981, 123). This value (which, of course, Lewis does not expect us to take too seriously, but simply uses as a 'ball park figure' given Laudan's formulation) is $Pr(\neg H/E) = 6/7$. Following Magnus and Callender (2004, 326), $Pr(\neg H/E)$ is best understood as the posterior probability that an arbitrary member of the set of *past* theories merited (as judged by current lights) on the basis of its past empirical success. Lewis argues that it is fallacious to infer from this probability anything with respect to either the false positive $Pr(E/\neg H)$ or the false negative rates $Pr(\neg E/H)$, since doing so would

[108] Throughout his paper Lewis omits any reference to '*approximate* truth' and talks of 'truth' *simpliciter*. In the main text I follow his usage for ease of exposition, since my main criticism revolves around his appreciation of the logic of the argument rather than the concept of 'approximate truth'. It suffices to note that the argument stands intact by substituting 'approximate truth' for Lewis' 'truth' (cf. Psillos 1999, 102-104).

amount to committing the fallacy of ignoring the base-rates or prior probabilities, exactly as in the NMA case examined in chapter 3. Indeed, re-using the formula for the posterior probability which involved the *Bayes factor* $\frac{\Pr(E/\neg H)}{\Pr(E/H)}$, i.e.

$$\Pr(H/E) = \frac{\Pr(H)}{\Pr(H) + \Pr(\neg H) \times \Pr(E/\neg H) / \Pr(E/H)}$$, it becomes obvious that a small posterior probability Pr(*H/E*) = 1- Pr(¬*H/E*) = 1/7 does not by itself imply that the Bayes factor is large enough to signify that empirical success is an unreliable test for truth. What is also required is consideration of the relevant prior probabilities. It could be the case that the prior probability of our theory is so small that even a low Bayes factor fails to raise the posterior probability to any significant level. Hence, Lewis concludes that the move from (4) and (5) to (6) is invalid.

Lewis in fact suggests that his analysis "provides a natural way for a realist to explain Laudan's historical evidence" (2001, 376). In probabilistic terms, all the realist needs to do in order to avoid the *reductio* is blame the prior probabilities of those past theories for the low value of the resulting posterior for truth instead of the Bayes factor. In words, "the realist can interpret Laudan's historical cases, not as evidence against the reliability of success as a test for truth, but merely as evidence of the scarcity of true theories in the past" (ibid. 377). Put more simply, science is better today and that is why it is natural for our theories to enjoy higher prior probabilities than it used to be the case.

Lewis is surely right that from a low posterior probability for truth it is not valid to infer to a high Bayes factor (i.e. infer that success is an unreliable test for truth). He is also right that the line of defence he suggests is open to the realist. This is *not* to say, however, that the way he has presented and evaluated the situation does justice to the 'pessimistic induction'.

Lewis claims that the move from (1) and (2) to (3) is valid. He writes:

> "In affirming that success provides a reliable test for truth, the realist is claiming that the rates of false positives and false negatives are low. If this is indeed the case, then if most current scientific theories are successful, it follows deductively that most current theories are true, as required by the realist" (ibid. 375).

The only way for (3) to follow deductively from (1) and (2) is to somehow read (1) and (2) as fixing the prior probability of a current theory at a sufficiently high degree. Indeed, Lewis interprets premise (2) as asserting that the prior probability of success Pr(*E*) is sufficiently high. As Magnus and Callender, however, correctly note,

> "on the assumption that success is a reliable indicator of truth, [asserting that Pr(*E*) is high] is tantamount to assuming that any arbitrary member of the population is likely to be true. If [the prior probability] is low (and how can we know if it is not?), then [the assumption about a high Pr(*E*)] fails and the conclusion [that our theories are probably true] does not follow" (2004, 325).

To see why this is so one has only to consider the expansion for Pr(*E*)= $\Pr(H)\Pr(E/H)+\Pr(\neg H)\Pr(E/\neg H)$. If success is a reliable test for truth, as assumption (1) dictates, both the false positive Pr(*E*/¬*H*) and the false negative Pr(¬*E*/*H*) rates are low. Hence, the only way for Pr(*E*) to be sufficiently high is that the prior probability Pr(*H*) is also high. Without this assumption, as we have seen, the NMA commits the 'Base-Rate Fallacy'.

Lewis in fact extracts the prior probability assumption from (2), i.e. the fact that most current theories are successful, combined with (1). This, of course, is the indispensable inductive assumption the realist must affirm if his reasoning is to go through. I suggest, nonetheless, that the most plausible way to understand Lewis' rendition of the 'pessimistic induction' is on the basis of this very assumption of high prior probability. In other words, what Lewis ought to have taken Laudan to be really suggesting from the very beginning through his counter-examples is the following: *assuming that the prior probability of our past theories being true was as high as that of our current theories*, evidence from the history of science entails that empirical success is not a reliable test for truth.

Reformulated this way (with, in effect, an extra premise about the priors), the argument becomes *valid*. As in all good *reductios*, the anti-realist must assume *all that the realist assumes* and create a contradiction. By adopting the realist's inductive assumptions in the form of the prior probabilities, Lewis' pessimistic inducer merely extends the realist's current inductive optimism to past theories. In this way, he is able to present his counter-examples in the form of past theories with the following two characteristics: a) they had been highly accredited by our 'inductive intuition' (i.e. assigned high prior probability) at the time that they were entertained seriously, and b) they were later (allegedly) proven false. With (a) and (b) in place, the *reductio* is valid: empirical success is no longer a reliable test for truth.

Consequently, Lewis' original reconstruction of the 'pessimistic induction' has missed the following premise (call it MP):

> Those past successful theories which have been declared false by current lights were justifiably highly accredited by our 'inductive intuition' at the time they were entertained seriously.

With (MP) alongside premises (4) and (5) in Lewis' reconstruction (6) follows logically and the *reductio* delivers its intended result. In more informal terms, (MP) further explicates the obvious consideration that for the counter-examples to have any force against the claim that success is a reliable test for truth, those past theories must bear some essential epistemic similarities with their forerunners. One of them, codified in assumption (5), is that they were empirically successful. Another one, codified in (MP), is that they were also worth taking seriously on the basis of inductive grounds in the first place.

To be sure, the 'pessimistic induction' is as much an inductive argument as the NMA is. Hence, there is ample room for questioning its logical force. The realist can respond in three (mutually compatible) ways: a) he can attack premise (4) and suggest that it is not the case that past theories were (wholly) false, thus restoring historical continuity on the theoretical level. This response, pursued over the years by Hardin and Rosenberg (1982), Kitcher (1993) and Psillos (1999, 108-114), is well-documented, so nothing else will be said here. b) Alternatively, the realist can attempt to resist premise (5) and the claim that these past theories were genuinely empirically successful. This kind of reaction has been suggested by Worrall (1988, 1989a, 1994) and has also been extensively discussed[109]. What is important for present purposes is that both (a) and (b) aim to cut down the size of Laudan's list of false but successful theories and, thus, on this probabilistic rendition of the argument, reduce the value of $\Pr(\neg H/E)$. Finally, there is another

[109] Worrall, of course, does not believe that this strategy takes all force from the PI but only that Laudan has tended to overstate his case (cf. 1989a, 154-155; 1994, 335). After all, it is this remaining force of the 'pessimistic induction' that led Worrall to put forth Epistemic Structural Realism as the only defensible realist position.

escape route available to the realist, namely Lewis' option. So, c) the realist can also resist (MP) and claim that the prior probability of those successful but false past theories was not as high as the prior probability of current scientific theories. The success of this move, nonetheless, presupposes that some principled distinction can be drawn between today's theories and past theories with respect to their fundamental inductive merit *before* any evidence comes in. Lewis' analysis has little to say towards this goal, except for the assertion that "[the realist's interpretation of Laudan's counter-examples] commits the convergent realist to the empirical claim that successful theories were rare in the past and are common today..." (Lewis 2001, 377). It is certainly true that Lewis' recommended reaction to the 'pessimistic induction' commits the realist to the claim that past theories deserved low prior probabilities. Whether such a commitment is *justified*, though, is a separate issue. On this count, the failure of all attempts so far to single out an uncontroversial measure of the *a priori* virtues prior probabilities codify, like simplicity or unity, casts doubt on the prospects of finding a convincing positive argument for distinguishing between present and past theories in terms of their fundamental inductive merit.

Far from being fallacious, then, the 'pessimistic induction' can indeed pose a threat to the realist position in the following two ways:

1. Assume that one accepts that Laudan's counter-examples suffice to generate a high $Pr(\neg H/E)$ and, thus, a low $Pr(H/E)$ for past theories.
2. Further assume that one fails to find a principled distinction between past and present theories, which can give rise to differential prior probabilities, but,

3. One is also unwilling to give up the belief that empirical success is a reliable test for truth.
4. It follows, then, that one ought to hold *low* prior probabilities of truth for *current* theories, which amounts to a withdrawal from the realist 'default'.

Alternatively:

1`. Again take Pr(¬*H/E*) to be high and, thus, Pr(*H/E*) to be low.

2`. Also assume that there are no grounds for differential prior probabilities, but,

3`. Take these prior probabilities to be uniformly *high* across the history of science.

4`. It follows, then, that empirical success is no longer a reliable test for truth, as Lewis' original *reductio* had it.

It is also interesting to note that it is not clear at all whether Laudan himself intended the 'pessimistic induction' as a *reductio* of the reliability of success as a test for truth rather than a *reductio* of the realist 'default'. Laudan asks the following question: "is there any plausibility to the suggestion ... that explanatory success can be taken as a rational warrant for a judgement of approximate truth?" (1981, 121). His answer is, of course, negative. As stated, however, it is unclear where exactly he thinks that the problem lies. Is success somehow inherently deficient as a test for truth or is it just insufficient to yield rational warrant for a claim of approximate truth, due to the extreme prior implausibility of theoretical science? It is hardly surprising, of course, that we cannot tell from Laudan's presentation what he thinks the main problem is. This is because his argument consists of an informal inductive claim, whose structure becomes transparent only under the lights of Bayesian Confirmation Theory.

Though interesting in its own right, I shall not offer an exegesis of Laudan's paper and his 'real' thoughts. I think, however, that the 'pessimistic induction' is more convincing as a *reductio* of the realist 'default' rather than the reliability of testing. Indeed, it would be too far-fetched to deny the basic empiricist presumption that empirical testing is the only reliable tool we can use in order to increase our knowledge of the world. Consequently, what the past record of false but empirically successful theories seems to suggest is that we should be overly cautious when evaluating the prior plausibility of current scientific theories.

As I said earlier, it is by no means trivial which epistemic position the success of the 'pessimistic induction' commits one to. All that can be concluded from this line of argument is that one has no reasons to be a realist. If this is the case, however, the 'pessimistic induction' has successfully, though by no means conclusively, undermined the NMA.

4.2.2 Should we be Realists, Constructive Empiricists, Structuralists or what?

In their recent paper (2004), Magnus and Callender attempted to give an explanation of "the feeling of futility in the realism debates" (327) by offering a probabilistic reconstruction of these debates. By endorsing the charge that both the NMA and the 'pessimistic induction' commit the Base-Rate Fallacy, their main diagnosis of this pathological state of affairs is that realists and anti-realists talk past each other:

> "anti-realists responding to the no-miracles argument seek to increase [$\Pr(\neg H/E)$] ... [while] realists responding to the pessimistic induction seek to lower the value for [$\Pr(E/\neg H)$]" (ibid. 327).

We have seen, nonetheless, that one can reconstruct both arguments by taking into account prior probabilities or base-rates in a way that allows both sides to talk about the same thing and (at least) agree on what premises they disagree upon.

In fact, one can be even more optimistic than this. By revealing the presuppositions of each argument, the foregoing analysis also shows, I think, the extent to which there can be fruitful debate concerning the issue of realism. As we have seen, the success of NMA hinges to a large extent on the prior probability one will assign to the approximate truth of our best theories. I have argued that there are kinds of considerations which *can* influence our prior degrees of belief. There are, for example, perfectly legitimate philosophical considerations which tell against the observable vs. unobservable dichotomy as a suitable motivation for constructive empiricism. It has also been shown that the 'pessimistic induction' is still a serious threat to realism, although by no means a presuppositionless one either. Quite inevitably, these considerations take the form of plausibility arguments and, hence, rest on further non-trivial assumptions. This fact notwithstanding, a Bayesian reconstruction of the debate contributes a rigorous conceptual apparatus which allows us to see the picture more clearly and appreciate the import of all the tacit presuppositions the standard informal arguments are committed to.

Having said this, though, we reach the *real* source of the 'feeling of ennui' that Magnus and Callender perceive in the debates. This is no other than the unfortunate (re)discovery of the fact that it is always possible to attribute the disagreement between realists and anti-realists to something on which *unequivocal* agreement is just not possible! This thing is no other than prior probabilities. In this chapter I maintained that, among the Bayesian variants, only Subjective Bayesianism can do justice to the real import of the main arguments in the debate.

Nonetheless, Bayesian personalism abstains from *definitively* constraining the prior probabilities beyond the restrictions set by the formal axioms of the probability calculus. This being so, the option is always open to blame the priors for any difference of opinion, under the sole constraint that the formal axioms of probability are not violated.

This fact, of course, trivially entails that a Bayesian reconstruction of the realism debate *fails to decide* the issue. In other words, Bayesianism cannot dictate what stance we have to take. If one's 'inductive intuition' is strong enough, then belief in the reliability of success as a test for truth is enough to make him a realist. A negative assessment of the efficacy of one's 'inductive intuition', though, would not suffice for the realist thesis. Which stance it would result in is a separate question. A constructive empiricist would hold a high prior probability that our best theories are empirically adequate while structuralists that our theories are only (approximately) structurally true[110]. Which of these 'default positions' is the right one, though, on the basis of some external objective criterion is a question that *cannot* be answered by Subjective Bayesianism. All that Bayesianism can do is accurately describe the situation, reveal hidden assumptions and bring to light the logic of the situation. When it comes to substantive issues, it has to remain silent.

It is, of course, possible, within the context of Bayesianism to reconstruct faithfully various episodes in the history of science and explain the reasons that made the adherents of one or other philosophical position change their minds on the basis of evidence. A classic example of such kind of historical reconstruction is

[110] An implicit (but permissible) assumption at this point is that the constructive empiricist and the structuralist recognise success as a reliable test for empirical adequacy and approximate structural truth respectively.

Dorling (1992), in which we find, among others, an extensive discussion and reconstruction of the debate on atomic realism. Dorling explains how using Bayesian conditionalisation accounts for actual cases of conversion from positivist/instrumentalist views of the atom to atomic realism on the basis of evidence. Note, though, that Dorling's task is essentially *descriptive*. What he is after is an explanation of the behaviour of actual scientists in the history of science rather than a definitive, *normatively* binding response to the question 'should we be realists?'. I do not doubt that reconstructions such as Dorling's are perfectly legitimate within a Bayesian framework of analysis. Yet, it should be equally clear that they do not address the normative issue of whether belief in the (approximate) truth of our theories is warranted or not, which is the central question of the realism debate.

This conclusion seems to contradict Magnus and Callender's own assessment of Dorling's endeavours. After praising Dorling for presenting "retail arguments about particular entities" (2004, 329) they continue:

> "In Dorling's cases, there may well be some plausible set of priors available, priors that realists and anti-realists could have agreed on before all the evidence came in. In the present wholesale case, however, where the entire fate of realism or anti-realism seems bound up with the priors, we can't imagine how one could find a reasonable set of priors" (329).

Magnus and Callender's conclusion rests heavily on the assumption that there is a significant difference between what they call 'wholesale arguments' for and against realism in general and other 'retail' ones, which allegedly refer only to particular entities. I have already endorsed Worrall's (forthcoming) criticism of this distinction when discussing the IBE-version of the NMA, which also makes use of (a version of) it. To repeat briefly, it is very difficult to see what the alleged 'wholesale' element could be that Magnus and Callender (following Boyd and Psillos and their talk of 'scientific methodology in general') have allegedly identified except for the union of all the 'retail' arguments for particular entities. If we grant this plausible understanding of 'wholesale arguments' or 'scientific methodology in general', both the NMA and the 'pessimistic induction' can be re-run with reference to particular entities only, as Magnus and Callender want it. Hence, if there are cases in the history of science where a reasonable assignment of priors could be agreed on, then the same ought to hold when one discusses the realism issue in more abstract terms.

But even if it is possible to discuss the issue of realism from a general perspective with 'priors that realists and anti-realists could have agreed on before all the evidence came in' in place, it is still mysterious how such treatment can have any normative import whatsoever. Indeed, all that is asserted here is that at a particular point in time, some number of scientists actually agreed upon some set of priors that seemed plausible at the time. Again, though, this descriptive statement by itself does nothing towards resolving the normative issue of whether a realist epistemic stance is warranted.

In short, Bayesianism can only (re)affirm that the only *compelling* reasons for one to change his mind on the realism issue are reasons of internal coherence. The realist who views the history of science in the same way as Laudan *and* sees no difference in kind between past and current theories *and* who is a committed empiricist can convert to anti-realism or structuralism. This will happen, though, only because clinging onto realism is not coherent any more with the rest of the details of his *own* scientific worldview and not because those other stances have somehow proven 'superior to' the realist view in some objective sense or against an objective common measure. Saying this much, though, is not incoherent with my claim that a Subjective Bayesian reconstruction of the NMA vindicates the 'intuitive' understanding of it as a plausibility argument rather than its IBE counterpart. It is just another way of recognising the impotence of NMA to provide anything like a *proof* of (the probability of) its conclusion from first principles.

One might protest at this point that I am merely rephrasing the problem of induction in its general form, i.e. the impossibility of providing conclusive reasons for the truth (or probable truth) of any non-trivial proposition. Well, this is indeed accurate but I don't see why it should be the reason for protest. All that is asserted here is that in the end, it is always possible to blame the priors for any *prima facie* irreconcilable difference of opinion. In the absence of an argument for *demonstrably* objective priors (some of the proposals to this aim were discussed and found inadequate in the course of our treatment), such disagreements are bound to be irreconcilable. It is true that the only remaining interpretation, i.e. Subjective Bayesianism "also respects Hume's argument that there is no sound inductive argument from experiential data that does not incorporate an inductive premise, and it also tells us what the inductive premise will look like: *it will be a probability assignment that is not deducible from the probability axioms*" (Howson 2000, 134)[111]. Hence, it is true that on a Bayesian analysis, much of the difference of opinion which is observed in the realism debate is to be attributed to the all time classic problem of induction. Does this present special problems for our analysis?

Some would argue that it does. There is a wide presumption that "most realists [do not] see the no-miracles argument as solving the problem of induction;

[111] Emphasis in the original.

rightly or wrongly, that problem is being bracketed here [i.e. in their paper] (assumed 'solved' or 'unresolvable')" (Magnus and Callender 2004, 323). It would seem to follow from this widespread sentiment that my analysis misses the point. Nonetheless appearances are deceptive here. Ignoring the problem of induction amounts to transforming both the NMA and the 'pessimistic induction' into instantiations of the 'Base-Rate Fallacy'. Correcting this fallacy means that one has to take these inductive assumptions into account in the form of prior probabilities. Now, if the problem is recognised as 'solved', then there ought to be a unique prior probability assignment. We have seen that there isn't. If it is recognised as 'irresolvable', there ought to be at least a consideration of the consequences of this situation for the problem in hand. Typically, though, such consideration is missing from the debate. One can only assume that confusion reigns. A Bayesian reconstruction of the realism debate is itself a very convincing *reductio* of the decision to 'bracket' the problem of induction. Instead of doing so, one ought to take it seriously into account and evaluate its import. Even if this decision brings with it feelings of 'ennui'.

This concludes my argument for thesis (d), i.e. that, although Bayesianism fails to decide the question of the appropriate epistemic attitude towards our best scientific theories, it still contributes a lot towards bringing out assumptions the standard competing stances are tacitly committed to. If anything, it is the inductive nature of these assumptions, codified in the prior probabilities, rather than fallacious reasoning, which is responsible for the often-expressed sentiments of futility regarding the prospects of the debate.

4.3 Conclusion

Taking stock, the main message of the last two chapters is that Bayesianism enjoys some kind of qualified success in its involvement in the realism debate. Although it is a legitimate and valuable tool for analysing and evaluating the main arguments for and against realism, it does not add much to our (already informally stated) reasons for deciding which side to take. I do not think that this qualification renders the Bayesian framework useless. But it is also true that we should not expect more from it than it can promise. There remains one more issue to examine, namely the possibility that our Subjective Bayesian reconstruction of the debate has somehow undermined all prospects for a solution and, in this way, surrendered the case to the relativist. In the next chapter I investigate this accusation arguing against the common tendency to relate Subjective Bayesianism with some form of relativism.

Chapter 5

Subjective Bayesianism, Relativism and Epistemology

In the last two chapters I argued that a reconstruction of the realism debate within the framework of Subjective Bayesianism, though possible, fails to decide the issue; indeed, it fails even add to the informal considerations for or against the various possible doxastic attitudes towards our theories. In this chapter I wish to examine the following possible objection to my conclusion. It is no wonder, it might be claimed, that Subjective Bayesianism fails to support any of the competing attitudes towards our best theories because Subjective Bayesianism *precludes* from the outset any such solution. Instead of failing to add anything to our informally stated reasons for belief, a Subjective Bayesian analysis rather undermines the strength of any argument we can possibly come up with. Hence, the widespread sentiment amongst contemporary philosophers that endorsing Subjective Bayesianism amounts to endorsing some form of relativism. My first aim, then, will be to deny that there is any special, *a priori* relationship between

Subjective Bayesianism and relativism. The problem of relativism is a much more fundamental issue and has to be countered on general epistemological grounds rather than on the basis of one's interpretational stance towards the probability calculus. Still, there is an inverse connection between relativism and one's interpretation of the probability calculus. Should an effective answer to the problem be found on general epistemological grounds, we can expect this answer to have an impact on our ways of interpreting partial belief. My second aim, then, is to explore the nature of this relationship by critically reviewing the main answers to the relativist challenge from the standpoint of fundamental epistemology. I shall argue for the following theses: a) we can hope to resist the relativist challenge only from the standpoint of traditional foundationalist epistemology, b) the 'intuitive' NMA is the only version of the realist argument respecting the guidelines of foundationalism, thus carrying significant normative force, and c) Subjective Bayesianism endorses this normative element.

5.1 The Alleged Link Between Subjective Bayesianism and Relativism

It is widely held that Subjective Bayesianism is *too* subjective to provide a satisfactory endorsement of the NMA. Here is a sample of the ways this complaint has been expressed:

> "[T]he subjective theory of probability ... [is] ... an example of the application of the *relativistic* mentality to such an increasingly important branch of modern mathematics as the probability calculus, and as an essential part of the new vision of science which we want to give in an irrationalist, and, as we shall say, probabilist form" (De Finetti 1931, 172; my emphasis).

> "[The Subjective Bayesian] view of the principles of probabilities makes them like a system of deduction, but it is only like a system of deduction where the choice of premises is a matter of whim ...where we end up depends on where we begin, and where we begin is arbitrary, so far as truth is concerned. So we have no reason whatsoever to suppose that science is taking us towards the truth or even towards empirical adequacy. On [the personalist's] analysis, the price of taking Hume seriously is radical scepticism [perhaps 'relativism' would be a more accurate term], and this is a price he feels obliged to pay" (Lipton 2002, 583).

> "[P]rior probability [assignments] are no less idiosyncratic (from the point of view of the Subjective Bayesian) than specifying [them] by, say, consulting a soothsayer" (Psillos 2004, 87).

> " [*A*]*ny* reason for fixing one's priors counts as legitimate in a Bayesian context. According to standard Bayesian epistemology, priors ... are up for grabs, meaning that one assignment of priors is as good as any other, provided both are coherent ..."[112] (Douven 2005, 340)

Even Worrall, who has explicitly endorsed our subjective Bayesian formulation of his preferred 'intuitive' NMA, has expressed reservations regarding the ability of Subjective Bayesianism to convey the weight of our intuitions:

> "It is not that [the Subjective] Bayesian rendition [of NMA] seems to me incorrect in any way... The assumption that, in cases where the NMA-intuitions kick in most strongly, the 'prior' of the theory concerned is at

[112] Emphasis in the original.

least reasonably high seems just a ... way of reflecting our intuitive judgment about the unity or simplicity of that theory ... (Indeed the intuitions here seem stronger: they characterise, in a way that (personalist) Bayesianism declares impossible, those cases in which it is, and those in which it is not, *reasonable* to assume a reasonably high prior. It is not a subjective matter that a theory claiming that planets move in ellipses is simpler than one that claims that their orbits are some Jeffreys-style monstrosity.)" (Worrall 2007, 146-147)[113].

By declaring the value of the prior probabilities a *subjective* matter, the Bayesian seems to be willing to consider any probability assignment whatsoever as legitimate under the sole condition that the assignment under consideration satisfies the axioms of probability. This means, though, that any hope of providing a rational method for resolving disagreement on substantive issues is lost at the outset. On a personalist construal any coherent view, no matter how intuitively outrageous, will be deemed on an equal footing with any other and in particular on an equal footing with some altogether more sensible view involving 'serious' priors.

This consequence is, of course, no different from a version of relativism regarding non-trivial, substantive beliefs. Bas van Fraassen characterises such a form of relativism thus:

> "Irenic relativism holds: there is (1) no objective criterion or rightness for opinion, and (/or) (2) no non-trivial criterion of rationality – anything goes, there is no truth except truth-for-you" (1989, 176).

Van Fraassen's formulation indicates that he thinks that (1) differs from (2) in some significant way. It is not clear at all, however, that there is any difference between

[113] Emphasis in the original.

them. The emphasis in both (1) and (2) is on the *criterion* for rightness of opinion and rationality, i.e. the means that we have for *deciding* which opinions count as right or rational. Consequently, (1) and (2) seem to be just logically equivalent ways of saying the same thing, namely that irenic relativism holds that we have no means for deciding which substantive beliefs are right/rational. Irrespective of whether (1) and (2) are logically equivalent, however, the fact remains: if it is possible for two different agents, who both respect the constraints set by the axioms of probability, to simultaneously assign quite different degrees of belief to the same set of substantive propositions with equal right, then "justified opinion" becomes agent-relative. Hence the charge that Subjective Bayesianism commits one to a version of relativism, from which it follows trivially that our Bayesian-'intuitive' NMA is of limited value, since it is an argument for just one of the many contradictory, but equally legitimate, views one can have on the issue of the (likely) truth of our current best scientific theories.

In order to answer to this accusation, one must clarify the relationship between Subjective Bayesianism and 'scientific rationality'. One thing we must note at the outset is that some elements in Subjective Bayesianism are as *objective* as it gets. These elements are supplied by the basic requirement that, at any given time, an agent's degrees of belief must satisfy the axioms of probability. As we saw in chapter 1, one can provide powerful reasons for according objective status to those axioms, reasons which, in historical order, originated in the from of *synchronic* Dutch-book arguments and culminated with the formulation of a soundness and completeness theorem for probability theory. It is now arguable that these results vindicate Ramsey's insight that the axioms of probability are best viewed as *logical constraints* directly analogous to those of deductive logic (cf. Ramsey 1926, 82; Howson 2000, 127-134).

It is certainly true that both deductive logic and Subjective Bayesianism *qua* the logic of partial belief will be central elements of any sensible conception of 'scientific rationality'. Indeed, any plausible account of scientific reasoning must contain at least a set of universally applicable standards, which will ensure the internal coherence of all scientific investigations. The crucial issue, then, becomes whether such standards *exhaust* the correct theory of 'scientific rationality' or not.

If it turns out that deductive and probabilistic logic exhaust the correct theory of 'scientific rationality', then there is no doubt that endorsing Subjective Bayesianism entails relativism and 'truth-for-you' regarding substantive issues as the inevitable outcome. Indeed, if all there is to 'scientific rationality' is a set of standards regarding the coherence of one's degrees of beliefs, then nothing precludes the *justifiable* assignment of very different (coherent) degrees of belief to the same set of synthetic propositions. This, however, amounts to succumbing to the relativist thesis that there can be no non-trivial criterion for justified opinion. If it can be maintained, on the other hand, that 'scientific rationality' involves *further* elements than just standards of coherence, then endorsing Subjective Bayesianism is not inconsistent with a more objective account of scientific knowledge and relativism ceases to be a threat.

On this question, I think the reader will agree that the claim that 'probabilistic coherence exhausts the correct theory of 'scientific rationality'' is *not* an integral part of any sensible understanding of the Subjective Bayesian interpretation of the probability calculus. Even if it is true that *in the end* all there is to 'scientific rationality' is standards of probabilistic coherence, this can't solely be the result of endorsing a particular theory like Subjective Bayesianism, which only explicates those standards. On the contrary, whether an acceptable theory of 'scientific rationality' can include non-logical constraints will have to be decided through an independent investigation of our ability to justify certain kinds of content-increasing inferences.

Colin Howson, a Subjective Bayesian, agrees:

> "I have scrupulously avoided discussing scientific rationality, partly because it is a highly contested area, but *mainly because this* [i.e. *Hume's Problem*, in which Howson articulates the logical interpretation of subjective probability I have adopted in this thesis] *is a book about logic, not about rationality.* The rules, if there are any, determining what is rational and what is not to believe or do I am happy to leave to others to fight over" (2000, 239; my emphasis).

Howson correctly distinguishes between a theory of logic and a theory of 'scientific rationality'. Subjective Bayesianism, as expounded by Howson, is a theory of logic, not a theory of rationality. Consequently, endorsing Subjective Bayesianism does not commit one to the view that all there is to 'scientific rationality is standards of coherence', and, *a fortiori*, to relativism regarding substantive questions.

Howson, to be sure, is not very optimistic about the prospects of a substantive theory of rationality. His previous remark continues as follows: "But what I do believe, and I believe that this extended footnote to Hume shows, is that no theory of rationality that is not entirely question-begging can tell us what is rational to believe about the future, whether based on what the past has displayed or not" (ibid. 239). His pessimism, however, does not derive from the fact that Subjective Bayesianism explicates the logic of partial belief. Instead, it derives from Howson's conviction that any non-logical constraint we may propose as a requirement of rationality ends up being question-begging. Surely, though, this is Howson's verdict of an independent investigation, which focuses on the epistemological merits of our content-increasing inferences and the kind of justification we can hope to achieve for them rather than the character of Subjective Bayesianism *per se*. And on this count, there might be some room for disagreement.

I conclude, then, that the decision to employ the Subjective Bayesian framework of analysis in our reconstruction of the realism debate does not by itself commit one to any form of relativism. The relativist threat stems, instead, from our seeming inability to devise independent standards sufficient to vindicate our content-increasing inferences. Consequently, the focus of our investigation should be directed towards the fundamental epistemological assumptions underwriting our content-increasing inferences and to possible ways in which those assumptions might be justified. Should our attempts prove successful, we will have succeeded in vindicating the normative force of the NMA, which, in the end, is just another content-increasing inference. And by doing so, we will have also managed to offer an argument to the effect that some of the possible prior probability assignments, which will inevitably appear in any Bayesian reconstruction of the argument, are more 'reasonable' than others.

5.2 The Epistemological Problem of Relativism

5.2.1 Internalist Foundationalism: The Source (?) of The Problem

One increasingly popular diagnosis of the difficulties facing our attempts to tackle the problem of relativism adequately locates the defect in our preferred method for justifying our substantive beliefs. Recall from the previous section (pp. 200-201) that relativism is understood in van Fraassen's sense, i.e. as the thesis that we have no means for deciding which of our substantive beliefs are right/rational. All problems then, it is said, arise out of our insistence on following an *internalist* variant of *foundationalism* as a method of justification, according to which a belief is justified if and only if *we know* that it is derivable from foundations that are themselves solid (that is, do not need a foundation in their turn). Both internalism and foundationalism trace their origin at least to Descartes' assertion in the 'First Meditation' that

> "I should hold back my assent from opinions which are not completely certain and indubitable just as carefully as I do from those which are patently false .... Once the foundations of a building are undermined, anything built on them collapses of its own accord; so I will go straight for the basic principles on which all my former beliefs rested" (1641, 17).

It should be clear, of course, that by endorsing 'internalist foundationalism' one is not thereby committed to Descartes' own rationalist solution to the problem of justified belief. Indeed, it could very well be the case that the solid foundations of knowledge do not resemble at all Descartes' 'cogito' or his benevolent God, who guarantees that no evil demon is deluding us, but consist instead of some basic observation statements accessible through our senses and in need of no further defence. In other words, 'internalist foundationalism' is consistent with both rationalist and empiricist epistemology, since it makes reference only to the *structure* of justification rather than the potential sources of our fundamental beliefs.

'Internalist foundationalism', then, holds that a belief is justified only if it has been derived by truth-preserving means from some other beliefs which are known to be true and, hence, can function as safe foundations. Suppose, however, that proposition *A* is affirmed on the basis of being derivable from proposition *B*, which is known to be true. But, surely, if proposition *B* were to count as secure basis for A, it ought to have been derived from a further more basic proposition *C*, which is also known to be true and which, in turn, owes its foundational status to an even more basic proposition *D* and so on and so forth. The only way to break the regress in a non-circular way is to find an axiomatic starting point, which is not in need of further support. But of course, the argument continues, there is no principled way to help us decide unequivocally what this starting point should be when it comes to our fundamental substantive beliefs. Different people will take different propositions as their fundamental beliefs; in the absence of a sharp criterion determining which are the right ones, if we are ever justified in believing in any possible fundamental proposition, then we seem to be justified in believing in all of them. This means, however, that different opinions on some one issue will end up being equally justified, having been derived from equally justifiably held foundations. Relativism and 'truth-for-you' on substantive matters seems to be the inevitable outcome.

It is no wonder then, it is argued, that there has been no satisfactory solution to the problem of the justification of our content-increasing inferences so far because the very method of justification employed (i.e. 'internalist foundationalism') already precludes from the outset any satisfactory solution. Our seeming inability to determine unequivocally which beliefs are to serve as the foundation for justified substantive opinion means that our quest for the fundamental level of secure first principles is bound to lead us either to a never-ending process of seeking the ultimate reasons for belief or to outright relativism.

5.2.2 The Externalist Perspective

It follows from this description of the problem that perhaps our difficulties will be resolved to some extent if we switch from the problematic 'internalist foundationalism' to a different method of justification for our beliefs. One way to do this is to abandon the 'internalist' component of 'internalist foundationalism' in favour of *externalism* or *reliabilism* in epistemology. I have already outlined and criticised various aspects of the externalist philosophical outlook in previous chapters, mostly in connection with the explanationist version of the NMA. My primary interest at this point is to meet the challenge of a potential rejoinder to my criticisms there, namely that those criticisms stem from an undue neglect of the subtle details of the externalist theory of justification and the aims it purports to serve. Hence, in this section I defend my earlier criticisms with explicit reference to the viability of the basic epistemological presuppositions underlying externalism.

In outlining the essential elements of the (informal) explanationist NMA in chapter 3, I noted that its most controversial aspect is its allegedly "virtuously" circular nature. To repeat briefly, on the explanationist reading the NMA is a meta-IBE: it asserts that we have reasons to believe that scientific methodology delivers (approximate) theoretical truth on the grounds that it is the best explanation of the empirical success of our best scientific theories (themselves reached through first-order IBEs). The reliability of IBE, however, claim the explanationists, depends crucially on the success of the NMA: it is only its conclusion (i.e. that we have reasons to think science delivers some sort of theoretical truth) that shows IBE to be a reliable rule of inference. In chapter 4 I suggested that the Bayesian framework is flexible enough to accommodate this allegedly harmless type of circularity but also rigorous enough to show off its main defect, namely that in the end alleged "virtuous" circularity is no different from vicious circularity. Both the informal and the formal version of the explanationist NMA, nonetheless, presuppose an externalist theory of the justification of belief, which can (allegedly) resolve these circularity worries.

Externalist theories of justification follow Alvin Goldman's insight that we should want "a theory of justified belief to specify in *non-epistemic terms* when a belief is justified" (Goldman 1979, 105; my emphasis). Goldman offers the following formulation of his version of externalism:

> "[A] belief is justified if and only if it is '*well formed*', i.e. it has an ancestry of reliable and/or conditionally reliable cognitive operations" (ibid. 117)[114].

Goldman understands 'reliability' as "the tendency of a process to produce beliefs that are true rather than false" (ibid. 113), but leaves it an open issue whether these 'tendencies' should be understood as actual long-run frequencies or as propensities to generate those beliefs (cf. ibid. 114). By 'conditionally reliable cognitive operations' he means those processes which yield a sufficient proportion of *true* output-beliefs *given that their input-beliefs are true*. Goldman introduced the weaker requirement of 'conditional reliability' in order to account for the possibility that a process is in fact reliable *but* fails to yield a high proportion of true output-beliefs because it is applied to false premises. In a similar vein, Psillos explicates externalism thus:

> "[O]n [externalist] accounts, if the rule [of inference in question] is reliable, then it thereby confers justification on a conclusion drawn using this rule, insofar as the premises are true. Hence, given externalism, all we

[114] Emphasis in the original.

should require of a rule-circular argument is that the rule of inference employed *be* reliable... ."[115] (Psillos 1999, 84).

On the externalist theory of justification, then, all one has to do is *deny* that one needs to *know* that a rule of inference is reliable in order to use it justifiably. It would be enough if this rule merely *were* reliable as a matter of fact, irrespective of what we think about it. With this assumption in place it follows that there are in fact sound foundations for our knowledge of various aspects of the world. These are simply the conclusions derived through some mode of inference, which, although not known to be reliable, actually *is* reliable as a matter of fact. In Psillos' words, "the correctness of [such conclusions] depends on the rule being reliable, and not on having any reasons to think that the rule is reliable" (ibid. 84). The explanationist NMA becomes then a straightforward illustration of this strategy: from the relatively uncontroversial premise of the predictive success of science and through the (allegedly) objectively reliable IBE one is thereby *justified* to think science delivers (approximate) theoretical truth. Furthermore, we saw in chapter 4 how this process can be expressed in Bayesian terms through our 'Bayesian-explanationist NMA'.

The main problem with this proposed understanding of externalism is that its claim to solve the problem of the justification of beliefs is frankly bizarre. Justification of belief is an *epistemic* affair. On a natural understanding this means that in order for our beliefs to count as justified, we need to be in possession of good reasons, telling us that we are indeed *correct* in continuing to use the methods which generate those beliefs. Now, the externalist wants to conclude that his content-increasing inferences are successful and, thus, that we have, in fact, good

[115] Emphasis in the original.

reasons to believe in the conclusion of those inferences. But what are the *reasons* for the success of those inferences? The externalist will respond: 'the reliability of the rules employed in those inferences, which we use to infer to their conclusions'. The problem with this response, though, is that, by the externalist's assumption, we do *not* have access to this reliability. We do *not* know, that is, whether the rules governing our content-increasing inferences are reliable or not. Still, the externalists insist that we don't need to know this. All we need is that these rules *be* objectively reliable, irrespective of whether we have knowledge of this fact. But, surely, even if we *de facto* arrive at truths by following some method, don't we need to *know* that this is the case in order to be justified in using this method? By answering this last question to the negative, externalism essentially avoids the whole problem of epistemology rather than offers an alternative solution to it.

It seems reasonable to conclude, then, that we are not justified in using a rule of inference without prior knowledge (or reliable evidence) of its reliability because doing so violates the natural understanding of the very notion of 'justification of belief'. In order to preserve the natural interpretation we must insist that knowledge of a rule's reliability is required before we are *justified* in putting it to use. This, however, turns all the allegedly non-vicious circular arguments into vicious ones. If the only reason we can *assert* that a rule of inference is reliable is the conclusion of an argument making use of this rule, then this conclusion must be inserted as a premise in the argument, thus resulting in vicious premise-circularity. In our example of the explanationist NMA, if IBE is to receive any support from it, an assumption regarding the (approximate) truth of our theories must be inserted explicitly among its premises. But, of course, this is precisely the conclusion of the NMA.

The preceding analysis centres on the inability of externalism to engage with the core of the problem of the justification of belief. The obvious response to the above criticisms is to try and bring back the agent's knowledge into the picture by showing that his reasons for belief, though broadly compatible with the reliabilist framework, extend further than merely the objective reliability of the rule of inference in question. Perhaps realising the implausibility of his initial formulation, Psillos incorporates the following *addendum* to his version of externalist epistemology:

> "All that is required is that one should have no reason to doubt the reliability of the rule – that there is nothing currently available which can make one distrust the rule" (ibid. 85).

Clearly this *proviso* falls outside the scope of Psillos' earlier formulation of externalism, since the agent's state of knowledge is an integral part of the picture again. Its main defect, however, is that it seems to amount to no more than shifting the burden of proof to those unconvinced of the rule's reliability. Shifting the onus is a predominantly rhetorical device. It does nothing towards providing positive reasons for the externalist perspective. And, of course, in the absence of those reasons, the onus can be easily shifted back: 'sure', it will be maintained, 'we do not have reasons to doubt the reliability of the rule, but we have no independent reasons to trust the rule either. How can such an absence provide any grounds for justification?'

Peter Lipton has recently attempted to characterise more precisely the externalist's viewpoint by invoking the following 'tracking requirement':

> "[A] strong inductive argument is one that is both an instance of a method that is generally reliable and is also an argument that is counterfactually

reliable in this instance. The premises track the conclusion, in Nozick's sense of the word. . . . Slightly more generally, in a strong inductive inference, if the conclusion had not been true we would not have made the inference" (2000, 184).

Lipton's counterfactual tracking condition aspires to do what 'naïve' externalism fails, i.e. provide positive reasons for the success of particular inductive arguments. The simple externalist requirement that the method followed be objectively reliable is supplemented with the new *proviso* that the argument in question also be *counterfactually reliable*.

Lipton's first explication in terms of 'what *we* would not have done had the conclusion been false' runs the danger of letting psychological considerations influence the evaluation of the truth-value of the relevant counterfactual conditional. It signifies, nonetheless, Lipton's desire to let (some form of) the agent's knowledge of the rule's reliability be part of the process of justification. Hence, when Lipton reveals that he has in mind a more objective notion of counterfactual reliability, understood as the 'circumstance' in which "had the conclusion been false, the evidence would have been different" (ibid. 186), he seems to assume that we are also in a position to affirm the counterfactual. Returning to the NMA, then, Lipton would presumably assert that the inference from success to (approximate) truth through IBE is both an instance of a reliable method of inference and counterfactually reliable. In other words, had our theories been false, they would not have been successful and hence there would be no basis for any inference to any form of truth. In a sense, Lipton intends counterfactual reliability as an *independent signal* for the objective reliability of method. With this signal in place (plus our ability to evaluate the counterfactual conditional) the externalist attempts to show that we can be in possession of strong independent reasons regarding the reliability of certain rules of inference, which justify the use we make of them.

Lipton is explicit, of course, that he has misgivings about the explanationist NMA considered as a *meta-IBE*. His worries stem from the fact that the explanationist NMA "fails to provide any independent or additional evidence for the truth of the theory or the reliability of the methods that lead to its acceptance" (ibid. 194) and, hence, the argument is best understood as "a summary of the first-order scientific evidence for our theory, rather than as a distinct but extremely general inference" (ibid. 197). I certainly agree with Lipton that the most sensible understanding of NMA is one which views it as an inference to the (approximate) truth of *particular* theories. I have also sided with those who think that the only way to make sense of NMA as an over-arching inference is by taking it to be a condensed and economical way to refer to all particular NMAs in one go – a 'summary' of these first-order NMAs as Lipton puts it. Even if the realist views the explanationist NMA as an IBE to the (approximate) truth of *particular* theories, however, the question what justifies these particular inferences retains its force. Consequently, the crucial issue becomes whether there is any *independent* support for Lipton's counterfactual *proviso* that can help the externalist avoid the unnaturalness of his proposed theory of justification.

To answer this, we must consider the circumstances under which we may uncontroversially assert the counterfactual *proviso* Lipton proposes. John Worrall has rightly noted that we can uncontroversially affirm such counterfactuals if "the conditional is a deductive consequence of some general claim that is independently accepted as true" (2000b, 213). If, for example, it were the case that empirical success *implied* (approximate) truth, then we could uncontroversially assert that had the theory not been (approximately) true, it would not have had the success it actually enjoys. The whole debate on realism, though, stems from the very fact that success does *not* imply any sort of theoretical truth. If it did, the NMA would be unnecessary.

Alternatively, the counterfactual would be assertible on the assumption that what bridges the gap between success and truth (in this case IBE) is known to be a reliable mode of inference. For if (approximate) truth were the best explanation of empirical success *and* IBE were known to be reliable, then success plus IBE would entail (approximate) truth and, hence, asserting the counterfactual would be a trivial issue. But, of course, by the externalist's assumption, we do *not* know that IBE is reliable. We merely use IBE in an attempt to vindicate it. Consequently, there appears to be *no independent* source of knowledge about IBE's reliability and, hence, *no independent* source for the assertability of the relevant counterfactual. On the externalist approach only the very conclusion of the NMA that particular theories are (approximately) true can effectively decide the issue of IBE's reliability, resulting once more in viciously circular reasoning.

Hence, far from being a symptom of failing to engage in depth with the basic assumptions underlying our methods of justifying belief, our rejection of the explanationist NMA and the externalist answer to relativism rest on solid epistemological foundations. This in turn means that externalism cannot be used to vindicate the constraints imposed on prior probabilities through our Shimony-inspired Bayesian-explanationist NMA outlined in chapter 4, and that, ultimately, our rejection of it was also justified.

Finally, it should also be noted that another tenet of the explanationist NMA, i.e. naturalism, falls together with externalism. *Naturalism* is the view that "there is no distinctive philosophical method which is either prior to scientific method or can be used to resolve first-order scientific disputes" (Psillos 1999, 78). From an internalist perspective one can easily recognise another *non-sequitur* here: how can it be maintained that NMA makes use of scientific method, when it is this very method that its conclusion aims to vindicate? Clearly, naturalism also presupposes an externalist theory of justification of beliefs, which we have seen to fail.

Externalism, then, fails to substantiate the claim that we can evade relativism by giving up the internalist component of 'internalist foundationalism'. Might it not be, though, that it is foundationalism rather than internalism which has to be abandoned in order to get around the relativist predicament?

5.2.3 Van Fraassen's 'New Epistemology'

Bas van Fraassen has attempted to overcome the relativist challenge by proposing an anti-foundationalist epistemology, which aims to overcome the alleged dead-ends of Cartesianism from a different angle. His alternative proposal, though, differs significantly from the externalist suggestion.

Van Fraassen attempts to motivate his epistemological views through an analogy drawn from jurisprudence regarding different conceptions of law. Here is his often-quoted idea:

> "The difference [between the standard Cartesian conception of rationality and van Fraassen's own conception] is analogous to that between … the Prussian and the English concept of law. In the former, everything is forbidden which

is not explicitly permitted, and in the latter, everything permitted that is not explicitly forbidden" (1989, 171).

The Prussian understanding of law, according to van Fraassen, is akin to the traditional Cartesian conception of rationality, where "what is rational to believe is exactly what one is rationally compelled to believe" (ibid. 171). Van Fraassen's aim is to reject this view of rationality in favour a reversed conception, according to which "what is rational to believe includes anything that one is not rationally compelled to disbelieve" (ibid. 171-172). Since this reversed, 'English' way is much more liberal than its 'Prussian' counterpart, the label 'voluntarism' (which van Fraassen had already introduced in his (1984)) conveys the spirit of his 'new epistemology' (cf. ibid. 175).

The class of sentences we are 'rationally compelled to believe' is determined by what van Fraassen calls "rules of right reason, canons of logic in a wide sense" (ibid. 172), and include deductive and probabilistic logic[116]. However these rules are non-ampliative and, hence, unable to increase our knowledge. All they can do is adjudicate on issues of consistency, be it deductive or probabilistic. The hope, van Fraassen maintains, that some ampliative rules of inference will achieve this kind of rational compulsion is illusory. Yet this is what Cartesian epistemology requires if any non-analytic statement is ever regarded as truly justified, and that is why "traditional epistemology embodies false hopes never to be satisfied" (van Fraassen 2000, 272).

[116] Van Fraassen would also add in the list Bayesian conditionalisation, at least under the assumption that we are following a rule (cf. 2000, 275). I have already argued at length in previous chapters that conditionalisation is not a logical principle even under the assumption that we follow a rule for belief-revision and, hence, that its inclusion in the list of logically compelling methods is erroneous.

The rules of logic also determine what one is 'rationally compelled to disbelieve', namely contradictions and incoherent sets of degrees of belief. Note, though, that van Fraassen's favoured conception of rationality now deems permissible not only tautologies but also synthetic propositions about the world. Hence, what counts as permissible under the traditional conception differs from what counts as permissible under van Fraassen's. In fact, the distinction van Fraassen draws amounts to presenting a *coherentist* theory of justification, according to which "rationality is in some *good* sense empty [and] any truly coherent opinion is truly rational" (ibid. 275)[117]. According to this view, any (degree of) belief which coheres well with the rest of one's beliefs (or degrees of belief in other propositions) is justifiably held by the agent. There is no need to find the ultimate foundations of knowledge any more. The property of coherence with the rest of one's (degrees of) beliefs is enough for justification and logic and mathematics define adequately what the standards of coherence are. Hence van Fraassen's *dictum* that "there is only so much method in science as there is logic and mathematics" (ibid. 275).

Van Fraassen, in effect, then advocates a theory of scientific rationality exhausted by deductive and probabilistic logic. But as we saw a little earlier, maintaining that scientific rationality amounts to no more than standards of coherence (by they deductive or probabilistic) entails relativism on substantive issues as its inevitable consequence[118]. Indeed, on this view different agents can justifiably hold different degrees of belief in the same set of propositions under the sole constraint that those sets of degrees of belief are coherent (i.e. do not violate

[117] Emphasis in the original.

[118] For a similar line of criticism see Psillos (2005), especially sec. 6.

the constraints set by the axioms of probability). For example, on van Fraassen's account both a Darwinian evolutionary biologist and a Creationist would count as rational if only they assigned their degrees of belief to evolutionary biology and Creationism coherently, no matter how different their credences in each were. Now, if we further consider the fact that, given enough ingenuity, *any* degree of belief in any synthetic proposition can be embedded within a coherent system of degrees of beliefs, the relativist position emerges triumphant. The coherentist criterion of rationality that 'new epistemology' puts forth proves too weak to decide between *any* mutually contradictory sets of degrees of belief to certain synthetic propositions. In this way, though, it ends up satisfying van Fraassen's own definition of relativism, i.e. that there is no non-trivial criterion of rationality and that, in the end, there is no substantive (i.e. non-logical) truth except 'truth-for-you'.

Consequently, if 'new epistemology' wants to avoid slipping into relativism, it must provide us with *further criteria* so as to distinguish between those sets of degrees of belief which are 'rational' in the sense of *coherent* and those which are 'reasonable', i.e. more faithful to the actual facts than others. Were some such distinction to be drawn, all 'reasonable' beliefs would be 'rational' but not vice versa.

Van Fraassen comes close to drawing this distinction by rebuffing the relativist threat on the following basis:

> "Logically speaking there may be infinitely many coherent alternatives to our posterior opinion, but almost none of those will be *live* options for us. This is not an accidental feature of our situation, but integral to it, exactly because we have prior opinion and prior understanding, which is not laid out for us on the

side, like a textbook, but which is already our own. … [W]e do not have the wherewithal to respond in more than a limited variety of ways at any given time. Many of the imaginable alternatives to what seems plausible to us cannot be incorporated by us. At least, they cannot be incorporated in a short span of time, through a coherent variation on our prior opinion, while doing justice to our new experience. So those alternatives *do not go*" (2000, 278).

Van Fraassen suggests that our prior opinions are of great epistemic significance. Prior opinions constrain our ways of accommodating new evidence and, hence, preclude serious consideration of the infinity of logically possible alternatives to our posterior opinion. On a Bayesian gloss, this means that some of our opinions are assigned high, while others are assigned low or negligible prior probability. Clearly, the value of the priors will crucially influence that of the posterior probabilities upon the accumulation of new evidence. Hence, even if there is an infinity of coherent alternatives to our posterior opinion, assigning negligible prior probability to most of those alternatives will result in negligible corresponding posterior values. Hence van Fraassen's conclusion that those alternatives simply 'do not go'.

Despite appearances, van Fraassen's solution is unsatisfactory. From a descriptive point of view it is incontestable that we value our prior opinions differentially and that *in practice* this ranking constrains our responsiveness to new evidence. An adequate theory of justification, however, ought to provide us with something more than just a description of which beliefs we assign high and which low prior probabilities. It ought to give us reasons why we *should* assign a high prior probability to certain beliefs and low to others, thus determining standards of correctness according to which we can judge whether our actual prior probability ranking is 'reasonable'. In our previous example, it should give us reasons why we should assign, for instance, a high prior probability to Darwinian theory and a negligible one to Creationism.

Van Fraassen's assertion that 'almost none of the infinitely many coherent alternatives is a live option for us' does *not* provide an adequate answer to the *normative* question it is intended to address. If it is just a report of the fact that most people in a particular community at a particular historical time happen to assign negligible priors to some theories but not to others, then clearly we still need an argument to the effect that they are *right* in doing so. The issue of the justification of belief is, of course, normative and cannot be settled by reference to the way we just happen to assign our prior weights. If, on the other hand, van Fraassen suggests that 'we', i.e. the intelligent and pro-science people, have our own standards of correctness *within our community* regarding the prior weights certain theories deserve, then van Fraassen has simply traded 'individual relativism' for 'community relativism'. Even if it is the case that within our culture it is considered 'correct' to assign, say, a negligible prior to Creationism, we know that there are other communities, in which the converse assignment is considered to be the 'correct' one. In the absence of an argument for the objective superiority of one of the two assignments we can hardly claim that we have evaded the relativist predicament. 'Community relativism' differs from 'individual relativism' only in terms of how numerous those who subscribe to mutually incompatible sets of degrees of belief are. Since both assignments can be held with equal right, however, this difference fails to make 'community relativism' less of a relativism.

Van Fraassen realises, of course, that it is not enough merely to state matters of fact in order to reach normative conclusions. The wording of (part of) his response shows that he intends to make a stronger statement. Hence, in the previous quote he insists that relying on prior opinion 'is not an accidental feature of our situation, but *integral* to it' (my emphasis). He also refers approvingly to Russell's 'colleagues on the Continent' who, in response to the problem of our knowledge of the external world, "were beginning to say: *we are already outside in the external world*!" (1989, 171)[119], presumably implying that our prior belief in the existence of the external world ought to be 1.

Still, it is rather difficult to make unambiguous sense of assertions such as the above. The only way to attribute clear meaning to the first one is to understand it as making the metaphysical, but no less descriptive, claim that it is necessary (i.e. it happens in all possible worlds) that we have to rely on prior opinion when responding to new evidence. Even if we can't help having certain prior beliefs, however, this by itself is not sufficient to show that we are justified in having the prior beliefs we have. What is required is an explanation of what is so special about *our particular* prior opinions which makes them the right ones to have. It is clear that the only standard of rationality 'new epistemology' favours explicitly, i.e. coherence, fails as an explanation. By van Fraassen's own admission, all those infinite alternatives are *coherent* alternatives. Hence, coherence cannot possibly provide the reason why we are right in favouring certain prior beliefs over others. Similar remarks apply to the *dictum* about Russell's 'colleagues from the Continent'. There has got to be some reason for taking the prior opinion that we are already outside in the external world seriously. Again, however, this reason goes beyond coherence, since there are (in principle, infinitely) many coherent ways to

[119] Emphasis in the original.

accommodate our sense-data *without* postulating the existence of a material external world.

The reluctance of 'new epistemology' to offer explicit criteria going beyond coherence for the determination of which prior opinions are the 'reasonable' ones becomes even more conspicuous when we consider some of van Fraassen's central philosophical views; namely, his disbelief in the reality of laws of nature as well as his belief in constructive empiricism. Both realism about laws of nature and scientific realism are, by van Fraassen's own admission[120], internally coherent positions. It follows from this, however, that 'new epistemology' seems unable to underwrite van Fraassen's rejection of them. Indeed, if all that matters to rationality is coherence, how can van Fraassen possibly argue for the superiority of his own views as against those equally coherent alternatives? To be sure, van Fraassen wants to say something more. He thinks that sound philosophical positions ought to belong to intellectual traditions which can be shown to "fare well under critical scrutiny, or else that [those positions provide] us with theoretical innovation of great value and promise" (ibid. 180). To the extent, however, that 'new epistemology' refuses to supplement coherence with further criteria for 'reasonable opinion', those demands fail to mark out as 'more reasonable' any of two (or more) coherent competing alternatives. Surely, if coherence is the sole requirement for rational opinion, then all coherent philosophical positions fare *equally* well under critical scrutiny and provide us with theoretical innovation of *equally* great value and promise.

[120] Especially with respect to realism about laws van Fraassen is adamant that "if some philosophers believe in the reality of laws, they are not *ipso facto* irrational" (1989, 180).

It might be counter-argued at this point that reliance on prior opinion delivers its normative promise only if another aspect of 'voluntarism' is taken properly into account, i.e. that a crucial dimension of epistemic judgements is that they are "most like expressions of intention" (1984, 254). Van Fraassen's idea is that our decision to endorse any non-trivial proposition as our belief entails some sort of cognitive commitment to this belief. By taking a stance on some issue we expect others to rely on our judgements and we also express our commitment to defending these views in the face of criticism, thus making belief 'a matter of the will'.

The alleged importance of this aspect of epistemic judgements with respect to the problem of locating the normative weight of certain prior opinions can, as far as I can see, be summed up in the following argument. It is highly unlikely that we would be willing to commit ourselves in such a strong way by endorsing certain judgements (in this case our prior opinions) unless we thought that these opinions were somehow epistemically privileged. Indeed, if we thought they were not privileged, we would have a very strong incentive to modify, or, at least, seriously qualify, our views. But in this case we would fail to rise to the occasion of our already expressed commitment, and this is too heavy a price to pay. This is because, as van Fraassen puts it, "to express anything but a full commitment to stand behind [our] promises and intentions, is to undermine [our] own status as [persons] of integrity and, hence, [our] entire activity of avowal" (ibid. 254).

I can think of three responses to this argument. First, it is not clear at all that this *de facto* commitment van Fraassen focuses on is the appropriate starting point for establishing the significance of certain of our prior opinions. Indeed, if commitment constitutes a central dimension of the epistemic weight of prior opinion, then the prior opinions which would come out as the 'most reasonable' would presumably be those of true believers or religious fundamentalists, whose commitment to their views appears absolute and nonnegotiable. It is highly doubtful, however, that 'new epistemology' has intended intellectual fanaticism as the most effective remedy to relativism. Consequently, the sort of epistemic commitment van Fraassen alludes to seems to be an insufficient basis for grounding the epistemic significance of certain prior opinions over others.

Suppose for the sake of argument, nonetheless, that epistemic commitment constitutes, in fact, a central dimension of epistemic judgements and, furthermore, that such commitment does presuppose that we *think* our beliefs are epistemically privileged (i.e. that they deserve high prior probability). Still, – and this is the second response- this by itself does *not* imply that these beliefs are indeed epistemically privileged and that we are in some sense *correct* in holding them. Hence, the epistemic significance of our prior opinions must rest on something other than simply our meta-belief that our prior opinions are normatively weighty – something that would *justify* this meta-belief of ours. Coherence clearly fails as a candidate for this 'something else' - for the same reasons as before: there are infinitely many coherent alternatives from the standpoint of which one can respond to new evidence.

Finally, - third response to van Fraassen's argument- even if we assume that we can always perceive correctly the epistemic significance of prior opinion so that no epistemic commitment will (actually) be undertaken on shaky grounds, this fact alone hardly illuminates the reasons behind this success. All we can manage under this assumption is to assert *that* our prior opinions are always epistemically significant; nonetheless, we still cannot explain *why* prior opinion has this significance, as any adequate epistemology surely ought to do. And of course, for the same reasons as above, coherence fails once more to take the place of the missing *explanans*. Consequently, van Fraassen's insistence that his 'new epistemology' avoids relativism by a justified reliance on prior opinion "[without further] constraints on rationality beyond coherence" (2000, 279) is wanting. It seems inescapable that any kind of justification for prior opinion has to involve something more than just the coherence requirement.

The only remaining line of defence for the 'new epistemologist' so far as I can see is that reliance on prior opinion is justified *pragmatically*. If so, then 'new epistemology' amounts to the claim that, although the canons of rationality are identified solely with logic and probability, relativism is defeated on pragmatic grounds. In fact, as well as frequent references to the American Pragmatists (cf. 1984, 235; 1989, 172-173 and 2000, 273-274), there is direct textual evidence that van Fraassen intended 'new epistemology' to be understood in exactly this way:

> "In opting for voluntarism or *pragmatism* in epistemology, one implicitly allows the relevance of just those critical standards that apply to other sorts of enterprise" (1989, 175; my emphasis).

> "The other part [of his assertion that 'it takes courage to rely on the reasoned opinion and skills you have developed, given the contingency of their conditions of applicability'] is the *Pragmatist* theme that epistemology cannot proceed in isolation from value theory – that the epistemic enterprise cannot be adequately conceived without attention to the role of value judgements, intentions, commitments, decisions, and other aspects of praxis" (2000, 273; my emphasis).

Understood under pragmatist lights, 'new epistemology' attempts to utilise elements which are prominent in other areas of human activity, like values and commitments, and even get informed by what count as critical standards for areas of thought outside philosophy. As their widespread employment in everyday life testifies, factors like values and commitments are highly esteemed and, hence, from a pragmatist perspective, legitimately inform our philosophical considerations. Consequently, prior weighting of opinion presumably receives its justification from a blend of such factors and then goes on to constrain the infinity of coherent alternatives to posterior opinion and, thus, defeat relativism.

The main problem with this pragmatist twist is that pragmatism has always seemed a most un-philosophical position. Perhaps the central reason for embarking on the philosophical quest to find adequate standards of justification is the fact that we want to examine whether widely held everyday beliefs are *justifiably* held or not. These beliefs refer to all of values, commitments, decisions and alleged critical standards encountered in various domains of praxis and the aim of the philosophical endeavour is precisely to investigate the extent to which common habits and preconceptions are in fact supported by strong theoretical reasons. It is mysterious how an allusion to those very beliefs we are, ultimately, aiming at vindicating can contribute anything to the process of vindication, unless, of course, viciously circular reasoning is implicitly accepted as legitimate.

It seems, then, that 'new epistemology' and coherentism underwrite a conception of rationality which is too weak adequately to counter the relativist challenge. It follows from this that coherence alone is insufficient to constrain the prior probabilities we assign to our theories. Instead of providing a solution, 'new epistemology' has simply restated the problem. Hence, neither internalism nor foundationalism has been shown to be especially responsible for the apparent force of the relativist position. And this takes us back where we started, i.e. to 'internalist foundationalism'. The inevitable question which arises, then, is: 'does this mean that relativism in the end wins the day?'

5.2.4 Back to 'Internalist Foundationalism'

So far my argument has been mainly defensive. Its aim has been to demonstrate that denying either the internalist or the foundationalist component of 'internalist foundationalism' will not get us anywhere close to overcoming the relativist predicament. At the same time, however, the preceding discussion has brought out the main virtues of the traditional understanding of rationality – these virtues tell strongly in favour of the employment of the traditional understanding to meet the relativist challenge. The first virtue relates to the fact that from an internalist point of view alone can we ever hope to solve the problem of the justification of belief. This is because only the internalist perspective explicates in a natural way the conditions that must be met, if our beliefs are to count as justified. Non-internalist views simply evade the problem, thus rendering internalism the sole appropriate epistemological standpoint from which to approach the issue of justification.

Foundationalism, on the other hand, owes its *a priori* intuitive appeal mainly to logic and this in two ways. First of all, the very process of constructing sound arguments requires that the premises of a valid argument are true. If this condition is satisfied, our argument establishes the truth of its conclusion and, in this sense, it is normatively binding. Furthermore, it is evident that in such a case our true premises are the foundations of our inference. It is a short (and obvious) step to generalise this process and demand that all knowledge is derived from strong foundations (in the guise of true first principles or low-level observation statements). Secondly, and relatedly, this is the form that all axiomatic systems have taken since antiquity, starting with Euclidean geometry. Being considered models of rigorous thinking resting on strong and self-evident foundations (i.e. axioms), it is no wonder that a similar system was hoped to be had for empirical knowledge too.

Consequently, 'internalist foundationalism' is not just the 'second best' option, to be preferred only because of the failure of externalist or anti-foundationalist theories of justification. It is supported by positive arguments on its own right. These are that 1) it only explicates in a natural way the central issues involved in the vexing problem of the justification of belief, and, 2) it is inspired by sound logical reasoning aimed at delivering conclusions with full normative import.

Nonetheless, this intuitively plausible method of justification has long been accused of embodying 'false hopes never to be satisfied', to use van Fraassen's words. The reason for this is, of course, the infinite regress noted earlier. To repeat briefly, suppose proposition *A* is based on another proposition *B*, which serves as its foundation. Surely, then, one needs some justification for this more basic proposition *B*, which can only take the form of an even more basic proposition *C*. Proposition *C* in its turn, though, owes its foundational merit to an even more basic proposition *D* and so on and so forth *ad infinitum*. Clearly, if this regress argument invalidates the traditional conception of rationality and my criticism of the various alternatives in the previous couple of sections is correct, relativism emerges victorious.

There are 2 ways to attempt to free oneself from an infinite regress. The first is to *axiomatically* put an end to it by declaring some stage of the process as our unquestionably secure basis. The second is to claim that *circular* reasoning is entirely acceptable and try to provide reciprocal support to the claims we seek justification for. In our example, this line of reasoning would commit one to the claim that, say, proposition *A* is based on proposition *B*, which in its turn receives support from proposition *A* itself. Crucial to the relativist's argument is the premise that neither option succeeds in salvaging 'internalist foundationalism' and, hence, both options fail to yield adequately justified beliefs. It is arguable nonetheless, that proclaiming certain beliefs as in need of no further defence differs from circularity in some important ways. If it turns out that these differences have epistemic significance, we might be able to come up with a credible case against relativism.

Throughout this essay I have been critical of circular ways of reasoning. The reason for this is not logical but philosophical. A viciously circular argument, though valid, fails to support the truth of its conclusion because it only helps to reassert the truth of what is already *explicitly* accepted as true. I have also been critical of attempts to present a form of circularity that is allegedly non-vicious. Under some natural epistemological assumptions, alleged non-vicious circularity can be shown to be just concealed vicious circularity. Consequently, the relativist is right that circularity cannot be a way out of the regress.

Axiomatically asserting certain fundamental beliefs, on the other hand, does not engage in any kind of *non*-reason-giving form of argumentation since it does not engage in any form of argumentation at all. By taking a stance on some fundamental issue one is not presenting an argument but only making an assertion. It should be stressed, of course, that this 'axiomatic proposal' applies exclusively to the *very fundamental* building-blocks of knowledge. It should not to be understood as encouraging the dogmatic endorsement of any 'high-level' belief; rather, it is a positive suggestion regarding the affirmation of certain initial assumptions, without which the process of rational argumentation about the epistemic status of our 'high-level' opinions cannot even get off the ground. Even so, the relativist argues that merely taking a stance on an issue is *not* sufficient for the justification even of our very fundamental beliefs. Anyone can assert that some fundamental view or another is in need of no further support on purely subjective grounds. It seems to follow from this that the axiomatic way out of the regress can be of no consolation to the rationalist. After all, rationalism is supposed to underwrite the conviction that theoretically informed discourse is the sole appropriate 'rational' means for the resolution of disagreement. How can mere assertion not only serve but, ultimately, lie at the very heart of this conviction?

Questions of this sort have caused widespread dismay at the prospects of an 'axiomatic resolution' of relativist worries, as testified by the recent drive towards alternative theories of justification, such as the externalist and circularist ones I have been criticising. The general impression from the recent literature is that an epistemology resting on axiomatically asserted foundations is just an outmoded relic of dashed Cartesian hopes. Is there any way of countering this?

The axiomatic variant of 'internalist foundationalism' is a coherent position. 'Internalist foundationalism' merely requires that our beliefs be 'derived from some other beliefs which are known to be true and, hence, can function as safe foundations for further inferences'. There is no mention of the way we can come to know the truth of our foundational beliefs. Argumentation suggests itself as a possibility, but surely 'internalist foundationalism' does not commit one to the view that it is the only possible way. In fact, if it did commit one to such a view, then the regress argument would be enough to discredit 'internalist foundationalism' conclusively. The main issue is whether, *within the context of 'internalist foundationalism'*, the axiomatic response to the issue of the justification of belief is a plausible answer. In other words, whether there are circumstances under which we can deem our foundational beliefs justified though they have been asserted without further argument.

If my arguments against the viability of the alternative epistemologies examined are sound, we have to choose between two equally consistent views: an axiomatic variant of 'internalist foundationalism' and a form of relativism regarding substantive issues (henceforth referred to as 'empirical relativism'). 'Empirical relativism' too can be formulated as a consistent position, despite the common assertion that relativism is a self-undermining doctrine. All one needs to do in order to avoid the alleged self-undermining is specify the scope of what one is relativist about, making sure this class excludes the relativist assertion itself. For example, 'empirical relativism' can avoid self-undermining by declaring that contradictory views on empirical issues are equally legitimate but also that 'empirical relativism' itself, i.e. the thesis that 'contradictory views on empirical issues are equally legitimate', is *not* an empirical matter[121]. Thus formulated, there remains no fundamental theoretical incoherence in the relativist ideology.

[121] This solution amounts to nothing more than applying the type-theoretic answer to the various self-referential problems, which arose with particular vigour during the debates in the foundations of logic and mathematics at the beginning of the 20^{th} century, to the issue of relativism. Formulated this way, the relativist challenge acquires its sharpest form and provides a severe test for the axiomatic solution. It also avoids dubious pronouncements, which, if anything, compromise its credentials from the outset. A typical example of this is Bloor (1991), who writes: "[The sociology

It is difficult to deny that *prima facie* this situation presents us with a rather depressing dilemma, given our pre-theoretic intuitions and expectations. Indeed, choosing between 'empirical relativism' and 'axiomatic internalist foundationalism' seems like a choice between two unwelcome options. Still, that they are both unwelcome does *not* necessarily mean that they are *equally* unwelcome. Might it be the case, then, that, despite both options being bad, one of them is sufficiently better than the other to underwrite a satisfactory solution to the problem of justification? Unsurprisingly, I have no telling argument to offer as a resolution. I can only re-emphasize what friends of the axiomatic stance have stressed in defence of an optimistic view of knowledge. Their argument is based on two pillars: the realisation of the fact that, on mature reflection, no more than a 'dogmatic' assertion of the bedrocks of our knowledge can be reasonably expected to provide an answer to the relativist challenge – all other options being worse and/or attempts to disguise the fact that an axiomatic stance has in fact been taken. And, on a more positive note, the assertion that this situation is sufficient to support an optimistic view of knowledge. Despite the inconclusiveness of their reasons, I still think that the plausibility argument to follow amounts to a serious case against 'empirical relativism'.

Just as relativism can take many different forms, the 'axiomatic solution' too has been employed on many different levels. Starting with logic *cum* empiricism, Popper notes:

---

of scientific knowledge] would be reflexive. In principle its patterns of explanation would have to be applicable to sociology itself" (7). One is left wondering about the kind of philosophical gain a relativist can expect from explicitly relativising his own position.

> "The rationalist attitude is characterised by the importance it attaches to argument and experience. But neither logical argument nor experience can establish the rationalist attitude; for only those who are ready to consider argument or experience, and who therefore adopted this attitude already, will be impressed by them... But this means that whoever adopts the rationalist attitude does so because he has adopted, consciously or unconsciously, some proposal, or decision, or belief, or behaviour; an adoption which may be called 'irrational'" (1945, vol. 2, 230-231)[122].

According to Popper, it is impossible to argue persuasively for the rationalist attitude, which in his conception consists in logic and some basic empiricist epistemological assumptions. These assumptions have to be taken at faith, on pain of infinite regress or circularity. True, there has been a lot of discussion lately regarding the way our basic inferential rules get vindicated. In chapter 3 I spent some time reviewing so-called rule-circular attempts at vindicating deductive logic, which also provide the motivation for rule-circular defences of scientific realism. By now it should be obvious that critics of circular arguments are not impressed by such attempts even at the most fundamental level of logic. Our basic rules of inference can't be vindicated by any argument; on the contrary, as Popper surely correctly claims, they are accepted as needing no further defence, thus constituting our most fundamental axioms.

Let's grant for the sake of argument that our relativist is willing to accept deductive logic as a 'given', since investigating the problem of '*deductive*

[122] Worrall (1999, 350), who cites this quote from Popper approvingly, notes that the term 'arational' conveys the essence of Popper's proposal much better than the term 'irrational' Popper originally used.

relativism' would take us far afield. Is the situation different for induction? Lakatos (1968, 1974) suggests that it is not. In his attempt to bridge the gap between Popperian falsificationism and the (desirable) affirmation that as scientific knowledge grows, it takes us closer to the truth, Lakatos explicitly proposes the acceptance of a fallible 'inductive principle' on *a priori* grounds:

> "[T]he truth of an inductive principle must be established *a priori*, because its function is to be the premise in a proof or justification" (1968b, 186) [123].

> "It should be stressed that in my usage 'inductive principle' is ... any principle claimed to be *a priori* true which implies a confirmation function – whether the latter is probabilistic or not" (ibid. 187, fn. 1).

He is also adamant that his motivation is primarily meta-philosophical:

> "By refusing to accept a 'thin' metaphysical principle of induction, Popper fails to separate rationalism from irrationalism, weak light from total darkness. ... With a positive solution of the problem of induction, however thin, methodological theories of demarcation can be turned from arbitrary conventions into rational metaphysics" (1974, 165).

Whether we are Popperians or not, Lakatos's analysis suggests that affirming some fallible 'inductive principle' is our only hope of overcoming the relativist challenge. Lakatos agrees with Hume, of course, that there can be no *demonstration* of such a principle from indubitable axioms on pain of regress. He refuses,

[123] Emphasis in the original. We should note at this point Lakatos' assertion that his solution "should be a cause neither for sceptical resignation nor for *apriorism*" (1974, 163; my emphasis). By 'apriorism', however, he seems to have in mind grand *a priori* philosophising seeking to find indubitable first principles (he cites Kant as an example) rather than an *a priori* affirmation of a weak 'inductive principle' on the basis of our generally reliable 'inductive intuition'.

however, to side with the Humean pessimist, who thinks that "the combination of total lack of proof and strong assent [indicate] mere animal belief" (ibid. 163). 'Inductive intuition' underwrites an 'inductive principle', which is generally reliable *and* can be perceived to be so[124].

One illustration of the implications of Lakatos' 'inductive principle' is provided by the classic philosophical problem regarding the existence of a material external world. It is now a commonplace that no philosopher has ever managed to provide anything like a proof of the existence of a material external world from indubitable first principles. At the same time, though, it is arguable that no one can provide a better case than G. E. Moore's 'proof' in his 'Proof of an External World', at least when his argument is properly analysed. As is well known, Moore's 'proof' starts with the premises a) 'Here is a hand' and b) 'Here is another', to conclude c) 'Two human hands exist'. But (c) is (allegedly) already enough to prove that there are things external to us (since 'human hands' by definition refers to things external to our minds) and, hence, that scepticism regarding the external world is defeated (Moore 1939, 145-146).

The reason why this 'proof' was never taken seriously as a real proof was, of course, that, although a valid inference, Moore's argument seems question-begging (cf. Stroud 1984, 86). Since Moore takes human hands to constitute things by definition external to us, the relativist will respond that Moore's conclusion merely re-asserts what he has implicitly accepted as true in his premises, i.e. that there are

[124] John Worrall agrees that there is "no way of arguing for our basic methodological principles that has any claim to logical priority. Assuming that they do indeed lead to the right division between black and white cases, *we just assert them without argument*" (1989b, 384; my emphasis). He also reports that "Lakatos used to say (only half-jokingly) [that] there comes a point when a rationalist must get out his machine-gun to defend rationality" (ibid. 384)!

things external to our minds. In reality, however, Moore is doing nothing of this sort. His 'proof' essentially amounts to the suggestion that we are entitled to assert without any further argument the existence of things external to our minds on the basis of common sense alone, and that it follows from this 'bare truth' that the relativist is deluded in denying its special status. Indeed, it is doubtful if there is anything else except common-sensical considerations that anyone can cite against the relativist in support of the existence of a material external world[125].

Similar considerations apply when it comes to the truth of scientific theories. The NMA is a reason-giving argument for our belief that scientific theories are approximately true descriptions of the unobservable reality. Its success presupposes, as we have seen, *a priori* assumptions in favour of our theories – assumptions that a Bayesian reconstruction, I argued, makes transparent. These assumptions transcend the evidence supporting our theories because they are logically prior to it. Very much like Moore's case, only an axiomatic assertion of the presumption in favour of our theories can render the NMA an effective argument. Such a presumption, though, is similarly non-evidential: its assertion is dictated by our fundamental 'inductive intuition' and can be based on no further argument on pain of regress.

---

[125] In fact, direct textual evidence that Moore simply asserted the existence of a material external world without further argument can be found in the way he expresses his conviction that he *knows* his premises: "I certainly did at the moment *know* that which I expressed by the combination of certain gestures with saying the words 'There is one hand and here is another'... How absurd it would be to suggest that I did not know it, but only believed it, and that perhaps it was not the case! You might as well suggest that I do not know that I am standing up and talking – that perhaps after all I'm not, and that it's not quite certain that I am!" (1939, 146-147; emphasis in the original).

So, what is the significance of these cases of 'dogmatism'-at-work for the 'intuitive' NMA and its prospects for overcoming the relativist predicament? One thing the foregoing examples establish is this: if one is already convinced of the universal validity of logic and, primarily, the existence of an 'inductive principle' yielding justified beliefs about basic issues, like the existence of a material world, on axiomatic grounds, one is thereby *not* justified in rejecting the 'intuitive' NMA *just because* it also takes certain assumptions as 'given'. Obviously, accepting the validity of logic and the existence of a material external world does not commit one to scientific realism. It does commit one, though, to the view that basic deductive or inductive assumptions, assertible only *a priori*, can function as an effective defence against relativism. Hence, one's reasons for not adhering to scientific realism cannot be that NMA employs unquestionable assumptions *simpliciter* but only that it employs the *wrong* unquestionable assumptions.

Arguably, though, this does not take us very far. It only amounts to the truism that one shouldn't complain about the unquestionable character of certain assumptions of the NMA when one already employs other content-increasing inferences with similarly unquestionable presuppositions. This attitude, however, already assumes that our 'axiomatic solution' can be effective against relativism. What about the relativist who rejects the view that axiomatically asserted beliefs can ever be justified and, thus, believes that the 'axiomatic solution' is generally ineffective against relativism? Can we expect to convince him of the ability of 'dogmatism' to overcome the relativist challenge? Obviously not. *The relativist will never become convinced of the merits of the axiomatic variant of 'internalist foundationalism'.* What is consoling, though, is that this is as it should be. We should never have expected that we would be able to convince the relativist of the invalidity of his view! 'Empirical relativism', as I remarked earlier, is an internally coherent position. We would be able to provide a knock-down argument against the relativist only if we could somehow demonstrate the inconsistency of his views. We have already accepted, though, that his views are consistent. It follows, then, that our hopes for offering a *definitive* refutation of the relativist attitude have been overly optimistic.

Consequently, we now know 1) that the relativist will never get convinced by our arguments, and 2) that this is a result we should have anticipated. That we should have anticipated this situation, however, seems to be epistemically significant in an important way. It points to the idea that convincing the relativist should *not* be our central concern. The value of taking the relativist position seriously and analysing its logical structure cannot be that doing so allows us to eventually refute it once and for all. On the contrary, its significance lies in this: the analysis of relativism serves as a means for the articulation of a sophisticated rival epistemology, which aims at explaining our commonsensical optimistic conclusions about our knowledge of the world. Relativism serves the aim of making this rival elaborate: it exposes possible defects, points to issues that have to be tackled and forces us to construct a logically rigorous conception, so that it is effectively responsive to criticism.

This rival, I have been arguing, takes its most effective form in the guise of the axiomatic variant of 'internalist foundationalism'. 'Internalist foundationalism' asserts that knowledge claims regarding certain propositions are justified only insofar they have been derived from other foundational beliefs known to be secure. The 'axiomatic' element enters the scene when it comes to affirming these ultimate building blocks of our system of knowledge. On pain of regress, argumentation must stop at some point. Those basic beliefs of ours must be taken as 'given', on the basis of our intuition, be it deductive or inductive. Furthermore, this view of knowledge does not violate empiricism in any way. It merely fills the gaps empiricism inevitably leaves behind. Experience is still central on this view; what is acknowledged, nonetheless, is the commonplace that experience alone, though necessary, is not sufficient to settle epistemic questions transcending its boundaries. Our non-empirical help is required before we can draw any credible conclusion about certain matters of fact.

Summing up our plausibility case for the axiomatic variant of 'internalist foundationalism': compared to externalism and anti-foundationalism, 'internalist foundationalism' presents a logically rigorous, transparent and intellectually honest alternative, which brings to light the fundamental presuppositions of empirical knowledge. Instead of trying to hide the inevitability of these presuppositions behind circular reasoning, it affirms this inevitability and accepts it as a fundamental limitation imposed to us by logic. Having revealed those presuppositions, the *axiomatic dimension* of 'internalist foundationalism' allows us to affirm the dictates of common sense on strong intuitive grounds, thus contradicting relativism. As I remarked earlier, the relativist is not expected to concur and convert. But this is not the crucial issue. The crucial issue is whether we have managed to articulate a plausible alternative to the relativist position, to which my answer is affirmative.

These sentiments are further reinforced by the fact that the 'axiomatic response' to relativism seems straightforwardly to explain some of the features that made van Fraassen's pragmatist epistemology so appealing, while at the same time avoiding its problems. All of us in everyday life accept unquestioningly many propositions about various issues, which, intuitively, hardly deserve any further investigation. *Contra* van Fraassen's pragmatism, I have argued that this matter of fact has no epistemic significance by itself, and, hence, that it cannot inform our philosophical considerations. It points, nonetheless, to a crucial feature of the 'axiomatic response' to relativism, i.e. that there are issues about which we feel this way precisely *because* we can affirm, on *a priori* grounds, the *right* answer. Perhaps this is van Fraassen's real thought, at least to the extent he wants to combat relativism on normative grounds. If this is so, however, we have already moved away from the weak criterion of rationality favoured by his 'new epistemology', namely coherence, towards a more robust foundationalist theory of knowledge.

I conclude, then, that the circle which led us back to 'internalist foundationalism' does not condemn us to relativism. I do not claim to have demonstrated that relativism is an untenable doxastic attitude. I believe I have shown, nonetheless, 1) that a more optimistic epistemology can be consistently articulated in the vein of the traditional conception of rationality, and 2) that this epistemology is a plausible rival to the relativist position.

5.2.5 'Axiomatic Constraints' on Prior Probabilities or Why the 'Intuitive' NMA Does Not Surrender to Relativism

I argued earlier that the relativist threat does not arise as a consequence of endorsing Subjective Bayesianism. On the contrary, it centres on the more general problem of vindicating our content-increasing inferences, a problem inevitably reflected in the Bayesian approach but universal to all approaches. In the previous section I defended one such way of vindicating our inductive inferences. I suggested that relying on our 'inductive intuition' is not only a plausible method, it in fact is the *only* method through which we can consistently answer the relativist challenge on normative grounds. Many have felt that something stronger should be said and have revised their conception of epistemology accordingly. Sadly, their revisionist attempts have been unsuccessful. I have tried to argue, nonetheless, that, fortunately, no such revisions were required in the first place. If I am right, we can still claim that our inductive inferences are justified. But if we can do this, we can also claim that certain prior probability distributions are more 'reasonable' than others. This is because prior probabilities encapsulate the inductive assumptions one *must* affirm in order to perform any content-increasing inference. Consequently, if some inductive inferences can be deemed 'more reasonable' than others, then so is the case for the prior probability assignments, which codify the inductive assumptions those 'more reasonable' inferences necessarily presuppose.

The superior status of certain inductive inferences relative to others derives, according to the axiomatic variant of 'internalist foundationalism', from our reliable 'inductive intuition'. It is only with its help that we can reliably infer to certain hypotheses transcending the data. The reliability of our 'inductive intuition', however, can only be asserted on *a priori* grounds on pain of regress or circularity. On the assumption that such an assertion is permissible, it is not hard to find a probabilistic rendering. Our reliable 'inductive intuition' allows us to *correctly* assign higher prior probability to certain hypotheses over others and in this way allows the data to turn the balance decisively in favour of some hypotheses over others. Maybe we would like to say something more about the reasons that this assignment is the correct one. For better or for worse, it seems that nothing more can be said. It is only our intuition that can perform the function of constraining the prior probability of certain hypotheses and, hence, allow the data to favour certain beliefs over others.

The 'intuitive' NMA exemplifies this process. It only explicitly asserts the non-empirical presuppositions, which have to be in place in order for the argument to have any normative import. These presuppositions take the form of an optimistic presumption about our theories on the basis of their *a priori* virtues – essentially simplicity and unificatory power. Such a presumption transcends the evidence in favour of our theories and in a sense goes beyond the dictates of empiricism. Ignoring this presumption, however, results in invalid reasoning – indeed this was the main lesson of the discussion of the 'Base-Rate Fallacy'. Consequently, any logically clear-headed defence of realism *must* affirm this optimistic presumption in favour of simple, unified theories. The grounds of this affirmation, in turn, can be no other than our 'inductive intuition' on pain of circularity or regress. Hence, only the 'intuitive' NMA fully respects our logical situation. It is worth noting that Peter Lipton, though no great friend either of NMA or internalism, reached a strikingly similar conclusion through his criticism of one of van Fraassen's sceptical arguments against the realist thesis[126]:

> "[T]he realist must say that scientists do have the knack of thinking of the truth. This ability is, from a certain point of view, somewhat surprising, but it remains in my view far more plausible than the extreme ignorance,

[126] In his (1993) Lipton is responding to van Fraassen's argument from 'underconsideration', i.e. the thesis that the scientists' ability to rank scientific theories according to their explanatory merit does not suffice for the optimistic inference to the approximate truth of the one ranked the most explanatory.

substantive and methodological, that a coherent critic must embrace" (1993, 106).

The wording of Lipton's remark is revealing. On the realist view, the scientists' 'knack of thinking of the truth' can best be explicated in terms of what a reliable 'inductive intuition' dictates. And the only way to defend the reliability of intuition is through a plausibility argument, precisely as I suggested in the previous section. Perhaps unwittingly, Lipton himself points our attention to (and also embraces) the only defensible version of NMA, i.e. the 'intuitive' version embedded into the traditional context of Cartesian rationality.

In chapter 4 I suggested that the scientists' 'knack of thinking of the truth' is most faithfully captured from within the framework of Subjective Bayesianism. The assignment of a high prior probability to the (approximate) truth of our current (particular) scientific theories expresses this optimistic presumption, which is dictated by our 'inductive intuition'. On the assumption that my defence of the normative import of 'inductive intuition' succeeds, however, we also see that the 'axiomatic' character of these prior assignments is no obstacle to thinking of them as carrying normative force, since they can now be viewed as conveying what our *normatively effective* intuition mandates. Hence, the idea that 'inductive intuition' imposes *reasonable* 'constraints' on prior probabilities, constraints which can neither be given nor need further defence. Though different assignments are logically possible, some of them are the correct ones. I understand that many feel that a more compelling argument should be offered for the preferability of certain prior distributions over others. On mature reflection, however, nothing more can be said in a defensible way.

The 'mild' nature of such intuitively underwritten constraints is also the main reason why I don't think that my proposal takes us outside the spirit of Subjective Bayesianism in particular. Earlier in this chapter I argued that a sensible construal of Subjective Bayesianism is *not* committed to the view that standards of probabilistic coherence exhaust the correct theory of 'scientific rationality'. Since the fundamental constraints that it imposes on degrees of belief are logical in nature, Subjective Bayesianism is not *a priori* inconsistent with further non-logical constraints, provided, of course, that these constraints are argued for in a non question-begging way. All this is respected in my approach; what is added is a 'mild' presumption towards certain beliefs based on a *non-logical*, but still *a priori* assertible, 'inductive principle'. For better or for worse, our Bayesian reconstruction of the debate has made it clear that the only non question-begging way to defend this principle is through resort to intuition. Such a 'mild' (but, I suggested, sufficient) defence of 'inductive rationality' makes no reference to strong, allegedly logical, principles (like the Principle of Indifference) or systematic theories of objective chance (like the frequency interpretation). It only affirms positively certain considerations dictated by reliable inductive common-sense, in perfect accord with the logical constraints on rational belief set by the axioms of the probability calculus.

Finally, it should be noted that the argument in this section is *not* intended as an argument for realism but only as an argument for the claim that the 'intuitive' NMA carries normative weight and, hence, does not surrender the case to the relativist. It may of course happen that full-blown realism loses out on plausibility grounds and that a more modest stance, like structural realism or even constructive empiricism, captures the appropriate doxastic attitude towards our best scientific theories. We saw in chapter 4, for example, that the 'pessimistic induction' might force us to revise our plausibility judgements in the face of the history of science. This is still an open possibility but it is consistent with my purpose to defend the *normative* import of the 'intuitive' NMA, irrespective of whether it ends up defeating all possible counter-arguments. All I wanted to argue for was the thesis that plausibility considerations, after all things are considered, *can* 'constrain' prior probabilities and, thus, avoid relativism.

5.3 Conclusion

In sum: it in no way follows from the decision to reconstruct the realism debate from the Subjective Bayesian perspective that we are somehow falling back on relativism, thus transforming all competing considerations into equally legitimate claims. The problem of relativism is a general epistemological problem and has to be countered on general epistemological grounds. A careful look at the assumptions supporting the relativist challenge, however, reveals that only from the standpoint of traditional Cartesian epistemology can we hope to come up with a normatively effective response to it. Furthermore, the only viable response within the Cartesian context must in the end allude to some 'inductive principle' dictated by our normatively effective 'inductive intuition'. It has been argued that Subjective Bayesianism *is* in a position to capture the normative import of this 'axiomatic' solution to relativism, as evidenced by our Bayesian reconstruction of the 'intuitive' NMA.

Conclusion: The End for Philosophy?

In conclusion, I think that the Bayesian and the realist should be friends. The realist will find the Bayesian framework a very useful tool through which he can express his views and arguments with great clarity. The language of probability allows the realist to supersede the sloppy terminology of 'theory acceptance' and focus accurately on the epistemic dimension of the inference to the (approximate) truth of our best scientific theories. A formal-probabilistic rendition of this inference, however, does *not* reveal major logical flaws in the informal statement of the problem. On the contrary, it helps illuminate the real character of the (defensible version of the) realist No-Miracles argument as a *plausibility claim* and, in this way, resolve a long-standing disagreement within the realist camp. Regrettably, the same *cannot* be claimed for the debate as a whole: Bayesianism will not resolve the question whether realism or some other doxastic attitude is the appropriate one to maintain towards our theories. Still, it contributes a lot towards making clear what the logic of the problem is, and how various considerations can influence one's attitude relative to the question of realism. I would argue that this qualification renders the Bayesian approach neither useless nor redundant. Instead, it points to the essentially open-ended character of all philosophical problems, a feature which also explains the feelings of 'ennui' often expressed with regards to the prospects of an uncontroversial resolution.

Bayesianism, on the other hand, has many dimensions itself, intricately interconnected, which are further illuminated by its application to the problem of realism. More particularly, there are two issues which must be distinguished when one applies Bayesianism to a non-trivial issue like that of realism. The *first* refers to the *probabilistic logic* of the problem. It is uncontroversial that any probabilistic argument (or statement) must respect the constraints set by the axioms of the probability calculus if it is to count as probabilistically valid (or coherent). The *second*, and more controversial, issue relates to the actual value the elements encountered in those axioms (and theorems) should acquire (especially the *prior probabilities* in Bayes' theorem). Unlike the first issue, this second problem is *substantive*. It refers to the question of which is the correct value to be assigned to a certain hypothesis or piece of evidence rather than to the various relations which must be satisfied if various probabilistic assignments are to be integrated into coherent sets of degrees of belief.

A careful appreciation of these points hides important consequences for the assessment of the most controversial version of Bayesian Confirmation Theory, i.e. Subjective Bayesianism. The 'received view' on Subjective Bayesianism, which I outlined sketchily in Chapter 1, allegedly precludes any further constraints on prior probability assignments beyond those dictated by the axioms of probability. Construed in this way, however, Subjective Bayesianism seems to scotch all prospect of an objective account of scientific reasoning, despite our strong intuitions to the contrary. This feature is taken to be the main shortcoming of this interpretation of the calculus.

Through the application of Bayesian Confirmation Theory to the problem of realism I have tried to motivate a different view of Subjective Bayesianism. This alternative picture can be summarized in the following points: 1) Subjective Bayesianism is first and foremost a theory about the *logic* of probabilistic reasoning; 2) as such, it does *not* commit one to the view that logic exhausts the content of 'scientific rationality'; 3) it follows from this that Subjective Bayesianism is not *a priori* inconsistent with further (non-logical) constraints on prior probabilities. My defense of a (moderately) optimistic foundationalist epistemology aimed to show how such constraints may be accepted and why they do not violate the character of Subjective Bayesianism, re-construed in the way I suggested. The conclusion is that there are benefits for both sides. The Bayesian too is expected to use his involvement into the realism debate as an opportunity to elaborate his appreciation of the nature of Bayesian Confirmation Theory in general and Subjective Bayesianism in particular beyond the somewhat simplistic picture conventional wisdom has drawn for him.

Intuition: The End for Philosophy?

My proposed re-evaluation of Subjective Bayesianism claims that Bayesian personalism is consistent with further (non-logical) constraints on prior probabilities. All that it requires is that such constraints be argued for on independent grounds. A central theme of the present essay is that this last problem is in fact much deeper than the problem of choosing one's interpretational stance towards the probability calculus. The real problem we have to face up to is the way we can hope to justify our content-increasing inferences and thus avoid conceptual relativism. Constraining the prior probabilities is only another reflection of this fundamental epistemological predicament.

The problem of relativism and consideration of the ways it might be overcome, however, brings with it one final serious difficulty. It is a central tenet of this essay that the relativist challenge can only be countered from the standpoint of traditional foundationalist epistemology. The problem of the justification of belief, which lies at the heart of the relativist thesis, can be answered only if the demands set by what I called 'internalist foundationalism' are met. This is because only 'internalist foundationalism' explicates in a natural way the conditions that must be satisfied in order for our beliefs to count as justified. Granting that 'internalist foundationalism' offers the proper way to justified belief, however, involves, on pain of regress, admitting the fundamental building-blocks of our edifice of knowledge *axiomatically*, i.e. on the basis of *intuition* rather than argumentation. The 'intuitive' No-Miracles argument, for example, was deemed the only defensible version of the realist claim on the grounds that only it does justice to the fundamental epistemological demands set by 'internalist foundationalism' and, hence, that only it rests on strong epistemological foundations. I argued further that Subjective Bayesianism *is* in a position to accommodate such *axiomatic constraints* on prior probabilities in perfect accord with its fundamental principles.

The natural question which arises at this point is the following: how much philosophical solace can this allusion to *intuition* offer us? If our hopes to answer the relativist challenge ultimately lie in its axiomatic denial on our part, is that not equivalent to saying that we have simply failed to answer to the relativist? Or, at any rate, that philosophy is not adequately equipped to provide that answer? If anything, philosophy has incorporated through the years mankind's hope that all problems can be resolved through rational argumentation. Does not our allusion to intuition mark the end for philosophy as the vehicle for the critical, argumentative resolution of all intellectual disputes?

I think that the answer to the above questions can be naturally extracted from the character of my defence of the 'axiomatic resolution' of the relativist challenge. I have not simply *assumed* that rational argumentation has to be based on unquestionable foundations. To the contrary, this thesis was the result of a detailed analysis of the conditions for justified belief and the logical structure of our argumentative attempts to acquire such belief. It is certainly true that, ultimately, I don't share the view that all intellectual disputes are open to an argumentative resolution. This opinion, however, is not an arbitrary stipulation. Rather, it is the conclusion of an argument, which has also served as a *reductio* of this widespread but untenable presumption.

The hope that rational argumentation can effectively decide all intellectual disagreements is no doubt a noble remnant of the Enlightenment and, perhaps, a spontaneous sentiment of the empirically-oriented modern thinker. Still, clear-headed conceptual and logical analysis shows that this hope inevitably remains unfulfilled. Rational argumentation has limits, which take the form of the basic presumptions one must take for granted even in order to start reasoning. Insisting on finding further arguments for the assertability of those basic propositions will take us nowhere close to fulfilling our goal on pain of regress. At some point argumentation must stop. At this point, only 'dogmatic' affirmation based on reliable intuition can take over, at least so long as we wish to combat relativism in a coherent, logically rigorous and normatively effective way.

There is no doubt that many think that something stronger should be said if our case against the relativist were to count as successful. I, too, wish that something stronger could be said. Unfortunately, nothing can. My defense of the 'axiomatic resolution' of relativism, essentially on the basis of common sense, aimed to show that it still suffices as an alternative. One should be careful, however, when evaluating this claim. The 'axiomatic resolution' of relativism does *not* amount to a *demonstration* of the impossibility of a relativist viewpoint, since we have explicitly admitted that the relativist position can be formulated in a consistent manner. All I am proposing is a less direct answer, which highlights the possibility of the construction of a rigorous rival objectivist epistemology based on the dictates of common sense as its ultimate foundation.

So, is this not simply a re-description of the 'end of philosophy'? Even if philosophical thinking itself has *historically* been motivated by the hope that all disputes admit of an argumentative resolution, there is nothing that guarantees that this hope will be fulfilled. Through the discussion of the problem of relativism I have argued that, in fact, we now know that this hope *cannot* be fulfilled. Need this fact have damning consequences for philosophy? I think not. It does point, however, to an alternative view of philosophy, an alternative way of appreciating its value. Philosophy is not valuable just because it brings our hopes regarding the foundational merits of human knowledge to fruition. In other words, its value does not depend on the condition that, in the end, it must allow us to make the strong claims to knowledge, which have motivated our investigation in the first place. Quite to the contrary, the value of philosophical thinking is independent of its end result. It derives from the fact that philosophical analysis is performed with an eye to conceptual clarity and logical rigor, both necessary elements of any meaningful investigation. In short, philosophy is valuable because it constitutes a rigorous *method* for the treatment of certain foundational questions. Consequently, even if the verdict of this investigation turns out to be different from what we had hoped or expected, this is neither a reason for despair nor a sign that the method of investigation is worthless.

In our case, the pursuit of a probabilistic reconstruction of the realism debate has served as the starting point for a better appreciation of the No-Miracles argument for realism. I have sided with those who think that the argument is ultimately based on our 'inductive intuition' and, on this basis, I also offered an explanation of the feelings of 'ennui' recently expressed in the literature regarding the prospects for a definitive resolution of the controversy. I don't think that these conclusions, with the central role they attribute to intuition, take anything away either from the realist cause or philosophy in general. To the extent that such an analysis respects logical and conceptual rigor, the mature, open-minded and self-conscious researcher should welcome its results as a valuable addition to the long series of attempts aiming to capture the nature and character of human knowledge.

Bibliography

Achinstein, P. (2001) *The Book of Evidence*, Oxford: Oxford University Press.

Bar-Hillel, M. (1980) 'The Base-Rate Fallacy in Probability Judgments', *Acta Psychologica* 44: 211-233.

Bloor, D. (1991) *Knowledge and Social Imagery*, Chicago and London: The University of Chicago Press, 2nd edition.

Boyd, R. (1973) 'Realism, Underdetermination and a Causal Theory of Evidence', *Nous* 7: 1-12.

Boyd, R. (1984) 'The Current Status Scientific Realism' in J. Leplin (ed.) *Scientific Realism*, Berkeley: University of California Press (1984).

Braithwaite, R. B. (1953) *Scientific Explanation*, Cambridge, Cambridge University Press.

Carnap, R. (1950) *The Logical Foundations of Probability*, Chicago: University of Chicago Press.

Carnap, R. (1968a) 'On Rules of Acceptance' in Lakatos (1968a).

Carnap, R. (1968b) 'Inductive Logic and Inductive Intuition' in Lakatos (1968a).

Casscells, W., Arno Schoenberger and Thomas B. Graboys (1978) 'Interpretation by Physicians of Clinical Laboratory Results', *New England Journal of Medicine* 299: 999-1001.

Christensen, D. (1991) 'Clever Bookies and Coherent Beliefs', *The Philosophical Review* 100: 229-247.

Christensen, D. (1996) 'Dutch-Book Arguments Depragmatized: Epistemic Consistency for Partial Believers', *Journal of Philosophy* 93: 450-479.

Day, T. and H. Kincaid (1994) 'Putting Inference to the Best Explanation in its Place', *Synthese* 98: 271-295.

De Finetti, B. (1931) 'Probabilism', English translation in *Erkenntnis* 1989, 31: 169-223.

De Finetti, B. (1937) 'Foresight: its Logical Laws, its Subjective Sources' in H. E. Kyburg, Jr. and H. E. Smokler (eds.) *Studies in Subjective Probability*, New York: John Wiley and Sons, Inc.

Demopoulos, W. and M. Friedman (1985) 'Bertrand Russell's The Analysis of Matter: Its Historical Context and Contemporary Interest', *Philosophy of Science* 52: 621-639.

Descartes, R. (1641) *Meditations on First Philosophy*, John Cottingham (ed.), Cambridge: Cambridge University Press (1996).

Dorling, J. (1992) 'Bayesian Conditionalisation Resolves Positivist/Realist Disputes', *Journal of Philosophy* 89: 362-382.

Douven, I. (1999) 'Inference to the Best Explanation Made Coherent', *Philosophy of Science* 66, Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part I: Contributed Papers: S424-S435.

Douven, I. (2002) 'A New Solution to the Paradoxes of Rational Acceptability', *British Journal for the Philosophy of Science* 53: 391-410.

Douven I. (2005) 'Wouldn't it be Lovely: Explanation and Scientific Realism' (Review Symposium on the second edition of *Inference to the Best Explanation*, with James Ladyman, Igor Douven and Bas van Fraassen), *Metascience* 14: 338-343.

Douven, I. and T. Williamson (2007) 'Generalising the Lottery Paradox', *British Journal for the Philosophy of Science* 57: 755-779.

Duhem, P. (1906) *The Aim and Structure of Physical Theories,* transl. P. Wiener, NJ: Princeton University Press (1982).

Earman, J. (1992) *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.

Feyerabend, P. (1962) 'Explanation, Reduction and Empiricism', in *Realism, Rationalism and Scientific Method. Philosophical Papers, Vol. 1.* Cambridge: Cambridge University Press (1981).

Friedman, M. (1974) 'Explanation and Scientific Understanding', *Journal of Philosophy* 71: 5-19.

Foley, R. (1992) 'The Epistemology of Belief and the Epistemology of Degrees of Belief', *American Philosophical Quarterly*, 29: 111-121.

Galavotti, M. C. (1989) 'Anti-realism in the Philosophy of Probability: Bruno De Finetti's Subjectivism', *Erkenntnis* 31: 239-261.

Gillies, D. (2000) *Philosophical Theories of Probability,* London: Routledge.

Goldman, A. (1979) 'What Is Justified Belief?' in *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, Cambridge, Mass.: The MIT Press (1992).

Good, I.J. (1962) 'Subjective Probability as the Measure of a Nonmeasurable Set' in E. Nagel, P. Suppes, and A. Tarski (eds.) *Logic, Methodology and Philosophy of Science, Proceedings of the 1960 International Congress*, Stanford: Stanford University Press.

Goodman, N. (1954) *Fact, Fiction and Forecast*, Cambridge, MA: Harvard University Press.

Hajek, A. (1998) 'Agnosticism meets Bayesianism', *Analysis* 58: 199-206.

Hardin, C. and A. Rosenberg (1982) 'In Defence of Convergent Realism', *Philosophy of Science* 49: 604-615.

Harman, G. (1965) 'The Inference to the Best Explanation', *The Philosophical Review* 74: 88-95.

Hempel, C. G. (1960) 'Inductive Inconsistencies', *Synthese* 12: 439-469.

Hempel, C. G. (1965) *Aspects of Scientific Explanation*, New York: Free Press.

Howson, C. (1973) 'Must the Logical Probability of Laws be Zero?', *British Journal for the Philosophy of Science* 24: 153-163.

Howson, C. (2000) *Hume's Problem*, Oxford: Oxford University Press.

Howson, C. (unpublished) 'Some Formal Analogies Between Logic and Probability: the Good, the Bad and the Ugly', *BSPS Presidential Address 2005*.

Howson, C. and P. Urbach (1993) *Scientific Reasoning: The Bayesian Approach,* La Salle: Open Court, 2nd edition.

Jaynes, E.T. (1957) 'Information Theory and Statistical Mechanics', *The Physical Review* 106: 620-630.

Jaynes, E.T. (1968) 'Prior Probabilities', *Institute of Electrical and Electronic Engineers Transactions on Systems Science and Cybernetics*, SSC-4: 227 241.

Jaynes, E.T. (1973) 'The Well-Posed Problem', *Foundations of Physics* 3: 477-493.

Jeffrey, R.C. (1965) *The Logic of Decision*, Chicago and London: University of Chicago Press.

Joyce, J.M. (1998) 'A Nonpragmatic Vindication of Probabilism', *Philosophy of Science* 65: 575-603.

Kaplan, M. (1981) 'A Bayesian Theory of Rational Acceptance', *Journal of Philosophy* 78: 305-330.

Ketland, J. (2004) 'Empirical Adequacy and Ramsification', *British Journal for the Philosophy of Science* 55: 287-300.

Keynes, J. M. (1921) *A Treatise on Probability*, London: McMillan.

Kitcher, P. (1993) *The Advancement of Science*, Oxford: Oxford University Press.

Kyburg, H. (1961) *Probability and the Logic of Rational Belief*, Middletown CT: Wesleyan University Press.

Kyburg, H. (1968) 'The Rule of Detachment in Inductive Logic' in Lakatos (1968a).

Kyburg, H. (1970) 'Conjunctivitis', in M. Swain (ed.) *Induction, Acceptance and Rational Belief*, Dordrecht: Reidel.

Lakatos, I. (ed.) (1968a) *The Problem of Inductive Logic Vol. 2*, Amsterdam: North-Holland Publishing Company.

Lakatos, I. (1968b) 'Changes in the Problem of Inductive Logic' in *Mathematics, Science and Epistemology: Philosophical Papers, Vol. 2*, Cambridge: Cambridge University Press.

Lakatos, I. (1974) 'Popper on Demarcation and Induction' in *The Methodology of Scientific Research Programmes: Philosophical Papers, Vol. 1*, Cambridge: Cambridge University Press.

Laudan, L. (1981) 'A Confutation of Convergent Realism' in Papineau (1996).

Levi, I. (1967) *Gambling with Truth*, New York: Knopf.

Levi, I. (1974) 'On Indeterminate Probabilities', *Journal of Philosophy* 71: 391-418.

Lewis, D. (1980) 'A Subjectivist's Guide to Objective Chance' in Lewis (1986b).

Lewis, D. (1986a) 'Postscripts to 'A Subjectivist's Guide to Objective Chance'' in Lewis (1986b).

Lewis, D. (1986b) *Philosophical Papers, Vol. 2*, New York: Oxford University Press.

Lewis, D. (1994) 'Humean Supervenience Debugged', *Mind* 103: 473-490.

Lewis, P. (2001) 'Why the Pessimistic Induction is a Fallacy', *Synthese* 129: 371-380.

Lipton, P. (1993) 'Is the Best Good Enough?' in Papineau (1996).

Lipton, P. (2000) 'Tracking Track Records', *Proceedings of the Aristotelian Society*, Supplementary Volume, 74: 179-206.

Lipton, P. (2002) 'Review of Colin Howson's *Hume's Problem: Induction and the Justification of Belief*', *British Journal for the Philosophy of Science* 53: 579-583.

Lipton, P. (2004) *Inference to the Best Explanation*, London and New York: Routledge, 2nd edition.

Lipton, P. (2005) 'Wouldn't it be Lovely: Explanation and Scientific Realism: Author's Response' (Review Symposium on the second edition of *Inference to the Best Explanation*, with James Ladyman, Igor Douven and Bas van Fraassen), *Metascience* 14: 353-361.

Locke, J. (1690) *An Essay Concerning Human Understanding*, ed. P.H. Nidditch, Oxford: Clarendon Press, 1975.

Lyon, D. and P. Slovic (1976) 'Dominance of Accuracy Information and Neglect of Base Rates in Probability Estimation', *Acta Psychologica* 40: 287-298.

Magnus, P.D. and Callender, C. (2004) 'Realist Ennui and the Base Rate Fallacy', *Philosophy of Science* 71, 320-338.

Maher, P. (1986) 'What is Wrong with Strict Bayesianism?', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1: Contributed Papers, pp. 450-457.

Maher, P. (1990) 'Acceptance without Belief', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1: Contributed Papers, pp. 381-392.

Maher, P. (1992) 'Acceptance in Bayesian Philosophy of Science', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 2: Symposia and Invited Papers, pp. 153-160.

Maher, P. (1993) *Betting on Theories*, Cambridge: Cambridge University Press.

Makinson, D. (1965) 'The Paradox of the Preface', *Analysis* 58: 205-207.

Maxwell, G. (1962) 'The Ontological Status of Theoretical Entities' in H. Feigl and G. Maxwell (eds.) *Foundations of Space-Time Theories*, Minnesota Studies in the Philosophy of Science, vol. 3, Minneapolis, University of Minnesota Press.

Maxwell, G. (1970) 'Theories, Perception and Structural Realism' in R. Colodny (ed.) *The Nature and Function of Scientific Theories*, Pittsburgh: University of Pittsburgh Press.

Miller, D. (1974) 'Popper's Qualitative Theory of Verisimilitude', *British Journal for the Philosophy of Science* 25: 166-177.

Monton, B. (1998) 'Bayesian Agnosticism and Constructive Empiricism', *Analysis* 58: 207-212.

Moore, G. E. (1939) 'Proof of an External World' in *Philosophical Papers*, London: George Allen & Unwin Ltd, 1959.

Moser, P.K. and J. Tlumak (1985) 'Two Paradoxes of Rational Acceptance', *Erkenntnis* 23: 127-141.

Musgrave, A. (1988) 'The Ultimate Argument for Scientific Realism' in *Essays on Realism and Rationalism*, Amsterdam/Atlanta: Rodopi.

Musgrave, A. (2006) 'Responses' in Colin Cheyne and John Worrall (eds.) *Rationality and Reality: Essays In Honour of Alan Musgrave*, Dordrecht: Springer.

Nagel, E. (1961) *The Structure of Science*, London: Routledge & Kegan Paul.

Nickles, T. (1973) 'Two Concepts of Intertheoretic Reduction', *Journal of Philosophy* 70: 181-201.

Okasha, S. (2000) 'Van Fraassen's Critique of Inference to the Best Explanation', *Studies in History and Philosophy of Science* 31: 691-710.

Papineau, D. (1993) *Philosophical Naturalism*, Oxford: Blackwell.

Papineau, D. (ed.) (1996) *The Philosophy of Science,* Oxford: Oxford University Press.

Poincare, H. (1905) *Science and Hypothesis*, repr. New York: Dover, 1952.

Pollock, J. (1995) *Contemporary Theories of Knowledge*, New Jersey: Rowan & Littlefield.

Popper, K. R. (1945) *The Open Society and its Enemies*, 2 vols. London: Routledge.

Popper, K. R. (1959) *The Logic of Scientific Discovery,* London: Hutchinson, 2nd edition, repr. by Routledge (2002).

Popper, K. R. (1963) *Conjectures and Refutations*, 3rd rev. ed., London: RKP (1969).

Popper, K.R. (1983) *Realism and the Aim of Science*, London: Routledge.

Psillos, S. (1999) *Scientific Realism: How Science Tracks Truth*, London: Routledge.

Psillos, S. (2002) 'Simply the Best: A Case for Abduction' in A.C. Kakas and F. Sadri (eds.) *Computational Logic: From Logic Programming into the Future*, LNAI 2408, Berlin-Heidelberg: Springer-Verlag.

Psillos, S. (2004) 'Inference to the Best Explanation and Bayesianism' in F. Stadler (ed.) *Induction and Deduction in the Sciences*, Kluwer Academic Publishers.

Psillos, S. (2005) 'Putting a Bridle on Irrationality: An Appraisal of Van Fraassen's New Epistemology', in Bradley Monton (ed.) *Images of Empiricism*, Oxford: Oxford University Press.

Psillos, S. (2006) 'Thinking About the Ultimate Argument for Realism', in Colin Cheyne and John Worrall (eds.) *Rationality and Reality: Essays In Honour of Alan Musgrave*, Dordrecht: Springer.

Putnam, H. (1975) 'What is Mathematical Truth' in *Philosophical Papers: Vol. 1: Mathematics, Matter and Method*, Cambridge: Cambridge University Press.

Ramsey, F.P. (1922) [1989] 'Mr Keynes on Probability', *British Journal for the Philosophy of Science* 40: 219-222.

Ramsey, F.P. (1926) 'Truth and Probability' in D.H. Mellor (ed.) *F. P. Ramsey: Philosophical Papers*, Cambridge: Cambridge University Press (1990).

Reichenbach, H. (1949) *The Theory of Probability*, Los Angeles: University if California Press.

Rosenkrantz, R. D. (1981) *Foundations and Applications of Inductive Probability*, Atascadero, Calif.: Ridgeview.

Ryan, S. (1996) 'The Epistemic Virtues of Consistency', *Synthese* 109: 121-141.

Salmon, W. (1990) 'Rationality and Objectivity in Science *or* Tom Kuhn Meets Tom Bayes' repr. in Papineau (1996).

Schaffner, K. (1967) 'Approaches to Reduction', *Philosophy of Science* 34: 137-147.

Schick, F. (1986) 'Dutch Bookies and Money Pumps', *Journal of Philosophy* 83: 112-119.

Shimony, A. (1970) 'Scientific Inference' in R.G. Colodny (ed.) *The Nature and Function of Scientific Theories*, Pittsburgh: University of Pittsburgh Press.

Skyrms, B. (1987) 'Dynamic Coherence and Probability Kinematics', *Philosophy of Science* 54: 1-20.

Smart, J. J. C. (1963) *Philosophy and Scientific Realism*, London: RKP.

Stroud, B. (1984) *The Significance of Philosophical Scepticism*, Oxford: Clarendon Press.

Teller, P. (1973) 'Conditionalisation and Observation', *Synthese* 26: 218-258.

Teller, P. (1976) 'Conditionalisation, Observation and Change of Preference' in W. Harper and C. A. Hooker (eds.) *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Dordrecht: D. Reidel Publishing Company.

Tichý, P. (1974) 'On Popper's Definition of Verisimilitude', *British Journal for the Philosophy of Science* 25: 155-160.

Van Cleve, J. (1984) 'Reliability, Justification and the Problem of Induction', *Midwest Studies in Philosophy* 9: 555-567.

Van Fraassen, B. C. (1980) *The Scientific Image*, Oxford: Clarendon Press.

Van Fraassen, B.C. (1983) 'Calibration: A Frequency Justification for Personal Probability' in R. S. Cohen and L. Laudan (eds.) *Physics, Philosophy and Psychoanalysis*, Dordrecht: D. Reidel Publishing Company.

Van Fraassen, B. C. (1984) 'Belief and the Will', *Journal of Philosophy* 81: 235-256.

Van Fraassen, B. C. (1985) 'Empiricism in the Philosophy of Science' in P.M. Churchland and C.A. Hooker (eds.) *Images of Science*, Chicago: University of Chicago Press.

Van Fraassen, B.C. (1989) *Laws and Symmetry*, Oxford: Clarendon Press.

Van Fraassen, B.C. (1998) 'The Agnostic Subtly Probabilified', *Analysis* 58: 212-220.

Van Fraassen, B.C. (2000) 'The False Hopes of Traditional Epistemology', *Philosophy and Phenomenological Research* 60: 253-280.

Van Fraassen, B. C. (2006) 'Structure: Its Shadow and Substance', *British Journal for the Philosophy of Science* 57: 275-307.

Von Mises, R. (1928) *Probability, Statistics and Truth*, New York, Dover Publications Inc., 2nd edition (1961).

Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall.

Williamson, J. (1999) 'Countable Additivity and Subjective Probability', *British Journal for the Philosophy of Science* 50: 401-416.

Williamson, J. (2006) 'Philosophies of Probability: Objective Bayesianism and its Challenges' in A. Irvine (ed.) *Handbook of the Philosophy of Science Volume 9: Philosophy of Mathematics*, North Holland/Elsevier [Also available on-line at http://www.kent.ac.uk/secl/philosophy/jw/2004/ philprob.pdf].

Williamson, J. (2007) 'Motivating Objective Bayesianism: From Empirical Constraints to Objective Probabilities' in William L. Harper and Gregory R. Wheeler (eds.) *Probability and Inference: Essays in Honor of Henry E. Kyburg Jr.*, London: College Publications [Also available on-line at http://www.kent.ac.uk/ secl/philosophy/jw/2005/motivating.pdf].

Worrall, J. (1984) 'An Unreal Image', *British Journal for the Philosophy of Science* 35: 65-80.

Worrall, J. (1985) 'Scientific Discovery and Theory Confirmation' in J. Pitt (ed.) *Change and Progress in Modern Science*, Dordrecht: Reidel.

Worrall, J. (1988) 'The Value of a Fixed Methodology', *British Journal for the Philosophy of Science* 39: 263-275.

Worrall, J. (1989a) 'Structural Realism: The Best of Both Worlds', repr. in Papineau (1996).

Worrall, J. (1989b) 'Fix it and be Damned: A Reply to Laudan', *British Journal for the Philosophy of Science* 40: 376-388.

Worrall, J. (1994) 'How to Remain Reasonably Optimistic: Scientific Realism and the 'Luminiferous Ether'', D. Hull and M. Forbes (eds.), *PSA 1994*, vol. 1, East Lansing, Mich.: Philosophy of Science Association, pp. 334-342.

Worrall, J. (1999) 'Two Cheers for Naturalised Philosophy of Science', *Science and Education* 8: 339-361.

Worrall, J. (2000a) 'Kuhn, Bayes and 'Theory-Choice: How Revolutionary Is Kuhn's Account of Theoretical Change?' in R. Nola and H. Sankey (eds.), *After Popper, Kuhn and Feyerabend,* Kluwer Academic Publishers (2000).

Worrall, J. (2000b) 'Tracking Track Records: Relying on Meta-Induction?', *Proceedings of the Aristotelian Society*, Supplementary Volume, 74: 179-206.

Worrall, J. (2002) 'New Evidence for Old' in P. Gardenfors, J. Wolenski and K. Kijania-Placek (eds.) *In the Scope of Logic, Methodology and Philosophy of Science'*, Vol. 1, Kluwer Academic Publishers (2002).

Worrall, J. (2007) 'Miracles and Models: Why Reports of the Death of Structural Realism May Be Exaggerated' in A. O'Hear (ed.) *Philosophy of Science: Royal Institute of Philosophy Supplement 61*, Cambridge: Cambridge University Press.

Worrall, J. (forthcoming) 'Miracles, Pessimism and Scientific Realism', *British Journal for the Philosophy of Science.*

Worrall, J. and E. Zahar (2001) 'Appendix IV: Ramseyfication and Structural Realism' in E. Zahar, *Poincare's Philosophy: From Conventionalism to Phenomenology*, Chicago: Open Court, 2001.

Zahar, E. (1983) 'Logic of Discovery or Psychology of Invention?', *British Journal for the Philosophy of Science*, 34: 243-261.

Zahar, E. (2001) *Poincare's Philosophy: From Conventionalism to Phenomenology*, Chicago: Open Court.