A minor dissertation submitted in partial fulfillment of the requirements for the award of the degree of

Master of Social Science in Philosophy

The No Miracles Argument: “Capture me if you can!”

Sean Leader

BRKSEA003

Declaration

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: ________________________ **Date:** __________

Faculty of the Humanities

University of Cape Town

2013

The No Miracles Argument: “Capture me if you can!”

Abstract

This paper deals with an issue in the epistemology of science. The issue concerns the debate about *realism* of science. Specifically, I will be investigating the question of how a prominent argument in defence of scientific realism – the *no miracles argument* – ought to be interpreted. The no miracles argument conveys the idea that it would be a perplexing coincidence if vast arrays of phenomena are accurately explained by scientific theories, yet the theories themselves were in fact false. This intuition has attracted much attention, so much so that the no miracles argument has become the argument of choice for scientific realists. It is the aim of this thesis to unpack to the no miracles argument, to discern *what* exactly the argument is – i.e. what *form* the argument ought to assume, and whether the argument is plausible. The paper comprises four chapters. The first three chapters deal with the prominent interpretations of the no miracles argument: *the Bayesian approach*, *Inference to the Best Explanation*, and *the ‘default’ position*. My goal is to elucidate the three accounts and to assess each one’s capacity to capture the essential ingredients of the no miracles intuition, as well as to examine the plausibility of the argument according to the individual formalizations. For each account, I will consider the distinction between *wholesale* arguments for realism – those that concern science in general – and their *retail* counterparts – those that concern individual claims, in order to discern the correct *scope* of the no miracles argument. I support the view that the no miracles argument ought to be interpreted as a wholesale inference, about the truth of science in general. I argue that none of the three prominent accounts succeed in capturing the essential ingredients of the no miracles argument, nor in offering a plausible justification of the

inference from success to truth. In short: the Bayesian account is susceptible to a fallacy of reasoning; Inference to the Best Explanation is subject to the problem of circularity; and the 'default' position is inadequately supported by its premises. In the final chapter, I develop a new formalization, and argue that it is the nature of the activity of frequentist testing in science which explains why the scientific enterprise affords us a significantly larger number of *true* theories than false ones. I argue that this interpretation overcomes the problems which beset the other accounts, and contend that not only does this account succeed in capturing the no miracles intuition, but it provides a plausible justification of the inference from the success of science to the (approximate) truth of science in general – something the other interpretations have left wanting.

Contents

Introduction

> 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)

The philosophy of science can effectively be split into two general domains. *Epistemology* considers what it is to have knowledge of scientific facts; it deals with the question of whether scientific theories are true. *Metaphysics* examines the nature of reality as described by science – it aims to explain what the world is like if our scientific theories are true.[1]

This thesis deals with an issue in the epistemology of science. This issue finds its place within the debate concerning *realism* about science. More specifically, I will be investigating the question of how a prominent argument in defence of scientific realism – the *no miracles argument* – ought to be interpreted. Before outlining the plan of my exposition, it will be useful to make clear the context of the argument under scrutiny, as well as to clarify some definitions.

The realism debate centres on the question of whether we are justified in believing scientific theories, and what view we ought to adopt if we aren't. Scientific realism is the philosophical thesis which says that we have good reason to believe as true,[2] what our best, empirically successful scientific theories tell us about the world. It is a thesis about the very nature of scientific knowledge; one which recommends belief in both the observable and unobservable aspects described by the content of our best theories. To be more precise, scientific realism can be seen as consisting of the conjunction of two theses: (1) an *independence thesis* stipulates that the scientific claims we make are about a mind-independent reality – one which exists irrespective of whether or not we are, or can be aware of it; (2) a *knowledge thesis* which suggests that we already know, or are able to discover, which of these claims are correct.[3]

The contention which threatens scientific realism is the following: if the world exists independent of our awareness of it, how then can we be sure that we can have knowledge of it? The realism debate

[1] Papineau, D. (Editor). (1996), *The Philosophy of Science*. Oxford: Oxford University Press.
[2] As I shall discuss, scientific realists usually, instead of 'true', talk about theories being 'approximately true'.
[3] Papineau, D. (1996).

arises precisely in virtue of this tension; where realists seek to defend both the independence and the knowledge theses at the same time, opponents of realism will look for alternative resolutions. Opponents to realism reject one of the two realism theses. *Idealism* or *verificationism*[4] denies the independence thesis, insisting that the existence of a world outside of the realm within which we perceive it, is unintelligible. *Scepticism*[5] - the most extreme opposition to scientific realism - denies the knowledge thesis and argues that it is not possible to know whether our claims advanced about the world are true or not. Sceptics insists that insofar as the aim of science is to describe a reality out of reach of human perception, we have no basis according to which we can be justified in believing that scientific theories are true; knowledge of these imperceptible entities is impossible to attain.

In contemporary epistemology of science, it is difficult to be convicted about our knowledge of micro-sized, essentially non-observable objects postulated by modern scientific theories – entities such as gravitational waves, neutrinos, or the Higgs Boson. Contemporary philosophers of science generally accept that scientific theories aim to detail, literally, an unobservable reality consisting of infinitesimal particles and immaterial waves. What generally counts as evidence for a theory which incorporates relations between 'theoretical' or, unobservable, entities, is not only evidence for the *truth* of the observable consequences of the theory, but is also evidence that these relations actually *explain* what is going on at the observable level. It is the belief that the causal relations that a theory describes actually function to *generate* the observable phenomena the theory predicts, and that the theoretical entities referred to in this explanation are ontologically, and not merely methodologically, legitimate. Scientific realism, in general, is not of the opinion that the relations portrayed by a theory are merely an instrumental way of simplifying conceptually, talk about observable phenomena the theory predicts, but rather, that these descriptions, and not others which are incompatible with them, are *the reasons why* the regularities in the observable phenomena occur in this mind-independent reality.

Proponents of scientific realism often differ in the way they defend their metaphysical position. We began with a quotation from Putnam, and here we have some more:

> It would be a miracle, a coincidence on a near-cosmic scale, if a theory made as many correct empirical predictions as say, the general theory of relativity of the photon theory of light without what the theory says about the fundamental structure of the universe being correct or 'essentially' or 'basically' correct. But we shouldn't accept miracles, not at any rate if there

[4] Examples include 'phenomenalism' and Dummett style 'anti-realism'.
[5] Examples include 'constructive empiricism', 'fictionalism', 'instrumentalism' and US style 'anti-realism'.

> is a non-miraculous alternative. . . So it is plausible to conclude that presently accepted theories are indeed 'essentially' correct. (Worrall 1989: 140)

> *It cannot be just due to an improbable accident* if a hypothesis is again and again successful when tested in different circumstances, and especially if it is successful in making previously unexpected predictions . . . If a theory h has been well-corroborated, then it is highly probable *that it is truth-like*. (Popper 1983: 346)

What Putnam's and the above so-called 'textbook' formulations have in common, is an appeal to a specific type of argument – what is referred to as *the no miracles argument*. This argument conveys the idea that it would be a perplexing coincidence if vast arrays of phenomena are all explained by science, yet the theories themselves were in fact false. Otherwise said: if we suppose that these theories are far from the truth, it would be nothing short of a miracle if these theories were nonetheless successful. The conclusion of this argument, in light of the widely accepted predictive success of these theories, is that the theories are true - they "latch onto the world" - and from it follows that the claims they advance about the world are correct. In fact, this intuition has attracted much attention, so much so that the no miracles argument has become the argument of choice for scientific realists.

Notice the qualifiers 'approximately', 'essentially' etc. in the conclusions of the particular formulations of the no miracles arguments; any reasonable version of the no miracles argument will have, at most, the conclusion that the theory or theories under consideration are at least approximately true, or that they latch on – "*in some approximate way* – to how things are "beneath the phenomena."[6] Such qualifiers ensure that the conclusion of the no miracles argument is not plainly false, given that, as we shall see, 'current science' today, and what was considered 'current science' in past years, is generally inconsistent. This is precisely the premise of, what we will come to call the pessimistic meta-induction. For example, better, more recent theories tell us that Fresnel's theory, despite its great predictive achievements, is quite simply false.[7] Despite failed attempts to offer an uncontroversial articulation of the notion of 'approximate truth' when it comes to scientific theories,[8] it will be clear that the no miracles argument does not rely on any unequivocal characterization of truthlikeness. Therefore, when it comes to the no miracles argument, a realist

[6] Worrall, J. (2005), *Miracles, Pessimism and Scientific Realism*. Unpublished, revised paper first given at Lunchtime Colloquium at the Center for Philosophy of Science, Pittsburgh, p. 5.
[7] Worrall, J. (2005).
[8] See Niiniluoto, I. (1998) 'Truthlikeness: The Third Phase', *British Journal for the Philosophy of Science*, 49: 1-31.

would want to hold that it would be a miracle if a theory which achieved great success, instead of failing to be *perfectly* true, failed to be at least *approximately* true.

Why, one may question, has the no miracles argument been afforded so much support when it comes to defending scientific realism? Since it does not seem as though anyone has been able to offer a sufficiently plausible formalization of the argument, we should inquire what the impetus is for its widespread appeal. I answer this by way of an analogy, as an expository device to instil in the reader the apparent intuitive appeal for embracing the argument:

> *Suppose you are lost in the woods and you come across a topographical map.*[9] *You hope the map may help guide you to safety, and wonder, thus, whether the map is accurate. After some exploring, you find that everything the map indicates you ought to see - rivers, lakes, roads, peaks and so on – you do in fact see along your journey. Naturally, given the extensive detail of the map, you only see a small proportion of what it depicts, but you conclude that the map must be accurate, generally. Of course, the map may happen to be correct in just those features you have randomly come across, but incorrect in general. In other words, it may just be a fluke that the map was correct in predicting* those *features, and that this was just coincidental. However, you decide to embrace what the map tells you about the terrain in general. Your belief in its accuracy may come as a result of a couple of arguments you may reflect upon: given the numerous ways in which the map could have been incorrect about what you could have seen thus far along your path,* given *the fact that it nonetheless was correct* entails *that it is most probably accurate in general. Alternatively, you may convince yourself of its general accuracy by arguing that it is precisely this general accuracy – about both, parts you have* and *haven't explored – which would* best explain *why you have had success in using it. Or perhaps you may simply have an* intuition *that the map must be accurate; you trust its guidance - you feel it would be a miracle if the map were, for example, a map of another region, yet just* happened to *be correct about everything you have seen* here*, up until now. Whichever way you see it, you feel that you are justified in believing that what the map depicts is true, and trust it to guide you to safety.*

[9] This example is adapted from that found in Lipton, P. (2004), *Inference to the Best Explanation*, Routledge: London, p. 185. Lipton's example is intended to elicit the intuition underlying the no miracles argument as understood as an inference to the best explanation. The articulation here, however, is supposed to accommodate intuitions construed by a *range* of interpretations, each of which will be discussed in their own chapter.

In much the same way, the predictive success of science, or of a scientific theory in particular, does not *entail* that it must be correct – it may just be that the predictions that were verified were some of the true consequences of a generally *false* theory.[10] The predictive success under consideration may just be a fluke, in much the same way that the map may happen to have been correct in *just* those features you randomly came across. But surely such surprising happenings can't merely be attributed to coincidence? Many scientific theories are *impressively* successful; Classical thermodynamics has correctly predicted a great deal about various characteristics of substances for close to two centuries, and quantum electrodynamics accurately predicts the magnetic moment of an electron to something greater than one billionth of a part. It certainly would be a perplexing coincidence if vast arrays of phenomena are all explained by science, yet its constituent theories failed to latch on to the world in some sense – i.e. if they were not at least approximately true. These are the intuitions underlying the no miracles argument; denying that science is at least approximately true would make their success a miracle.

So what do the opponents of scientific realism have to say about all of this? Antirealism, despite granting that science has in fact been astoundingly successful, is in diametrical opposition to the realist thesis. Antirealists assert, instead, one or more of the following: there is insufficient justification for a metaphysical commitment to a mind-independent reality; we cannot take what science says at face value, and that its theories do not establish an epistemological commitment to take what they posit as constituting knowledge (both at the unobservable and the observable level). Consequently, there are numerous ways in which one can be an antirealist about science. Authors like Larry Laudan,[11] for example, subscribe to *the pessimistic meta-induction on the history of science* as their primary argument against scientific realism. This inductive argument conveys the idea that many past theories that have turned out to be false are nonetheless successful, implying that we cannot infer that success necessarily begets *truth* – in other words, just because a theory is successful does not give us sufficient reason to infer that it is (even at least approximately) true. The induction argument invokes a list of scientific theories which, it is suggested, were, at one or another point in time, empirically successful but are now recognised as neither referential nor true.[12] This gambit includes theories like the crystalline spheres of ancient and medieval astronomy, the effluvial theory of static electricity, the phlogiston theory of chemistry, the caloric theory of heat, the theory

[10] Recall that every false statement has many true consequences.

[11] Laudan, Larry. (1981), 'A Confutation of Convergent Realism', *Philosophy of Science,* 48: 19-49. I have adapted, in part, this brief discussion of Laudan and the 'pessimistic meta-induction', from my unpublished Masters coursework essay: Leader, S. (2011), 'Is the Pessimistic Meta-Induction on the History of Science a Decisive Argument in Favour of Scientific Realism?'

[12] Laudan, Larry. (1981), pp. 33-34.

of circular inertia, the theory of optical ether, and numerous others.[13] The idea here is that, insofar as there is enough evidence to indicate that success is not as miraculous as realists would like to believe, and that it is not only true theories which turn out to be successful, there is no reason to infer that our best, successful theories must be (at least approximately) true. Moreover, Laudan argues the following: scientific theories of the past were successful and accepted according to the same kind of evidence that we base adoption of our current theories. In light of our current theories, however, these past theories are now considered fundamentally wrong - their claims regarding the actual structure of the world are misguided, and they employ theoretical terms that don't refer. Therefore, Laudan argues, our current theories are most likely to turn out to be false in much the same way. The inference is from the fact that, successful theories of the past were once considered true but turned out to be false, to the claim that our current theories are most likely to turn out to be false in much the same way; in short, success is no guide to truth.

If what Laudan says is correct, then the history of science compels us to believe that the realist's appeal to the no miracles argument is unfounded, and the realist ought no longer be able to appeal to the *truth* of science in order to explain its success.

It is important to bear in mind that Laudan makes no attempt to preclude the possibility that currently successful theories may *happen* to be (approximately) true – the idea is merely that contrary to what the no miracles argument implies, we are not justified in making the inference from success to truth. This point is subtle; the pessimistic induction argument is not intended to dispel the *possibility* that science is in fact true. Instead, it merely places the burden of proof on realists to *justify* their belief in the success-to-truth connection. In light of this, and given the importance of the realism/antirealism debate, the no miracles argument deserves critical attention, to assess if it can be used as a feasible defence of scientific realism. It is precisely the aim of this thesis to unpack the no miracles argument, to discern *what* exactly the argument is – i.e. what *form* the argument ought to assume – and whether it can afford a plausible justification of the elusive inference from success to truth. Perhaps then, the task of resolving the realism/anti-realism debate will at least become clearer, and headway can be made towards a firmer metaphysical commitment with respect to the truth of science in general.

[13] Laudan, Larry. (1981), p. 121.

The structure of the exposition

The paper comprises four chapters. The first three chapters deal with the prominent interpretations of the no miracles argument: *the Bayesian approach*, *Inference to the Best Explanation*, and *the 'default' position*. Each interpretation aims, independently, to answer the question, "what *is* the no miracles argument?" My goal is to elucidate the three accounts and to assess each account's capacity to capture the essential ingredients of the no miracles intuition, as well as to examine the plausibility of the argument according to the accounts' respective formalizations. For each account, I will consider the distinction between *wholesale* arguments for realism – those that concern science in general – and their *retail* counterparts – those that concern individual claims; whereas the wholesale argument makes an inference about the (approximate) truth of science in general, the retail arguments give rise to an inference about the (approximate) truth of *particular* claims or theories. The goal here will be to discern the correct *scope* of the no miracles argument, and I will champion the view that the argument ought to be interpreted as a wholesale inference, about the truth of science in general.

In Chapter 1, I articulate a Bayesian formalization of the no miracles argument, one formulated in probabilistic terms. I present P. D. Magnus and Craig Callender's argument that this account is subject to the well-known, *base rate fallacy*. The fallacy will be made perspicuous by analogy to a famous diagnostic case, as well as by an assessment of the logical flaw in the context of the no miracles argument itself. I will discuss how the argument advanced by Larry Laudan, against the no miracles intuition, fits into this context, and I will present Stathis Psillos's response to Laudan's challenge, and suggest as well how this ties into the Bayesian formalization of the no miracles argument. I will present Magnus and Callender's own retail argument for scientific realism; one the authors claim avoids the fallacy. We will examine an argument by John Worrall, which claims that Magnus and Callender's retail defence of scientific realism is subject, by their own standards, to the same fallacy they caution against. I will discuss, as well, why the retail formalization fails to capture the no miracles intuition, and indicate that a more detailed explanation, in general, as to why the no miracles argument *ought to be* interpreted as a *wholesale* argument, will be presented in Chapter 2. I will mention, briefly, attempts to salvage a Bayesian conception of the retail version of the no miracles argument, and explain why they offer no resolution. My conclusion will be one which rejects a Bayesian interpretation of the no miracles argument – whether wholesale or retail – since the arguments lack cogency, and also fail to capture the intuition underlying the no miracles sentiment.

Chapter 2 considers the no miracles argument as an Inference to the Best Explanation. I begin with a detailed explanation of this general form of inference, as formulated by Peter Lipton. I will examine, in alignment with Lipton's account, various ways in which something may be an *explanation*, as well as competing views on the ways in which something may be the *best* of competing explanations. I consider once again, the scope of the no miracles argument, and argue here that it is the nature of the no miracles argument, as conveyed by the so-called 'textbook' formulations, and for the purposes for which it is appealed to, that it *ought to be* interpreted necessarily as a *wholesale* argument. I examine a wholesale account of the no miracles argument, as an Inference to the Best Explanation. This wholesale account begins with a description of the individual inferences scientists employ, and the question regarding the *justification* of such inferences is what will be scrutinised; I will examine whether these inferences to the best explanation, that scientists make, really are truth-tropic – whether they reliably take scientists towards the truth. I will present a crucial challenge for this wholesale conception, namely, *the circularity objection*. Although Lipton suggests that the wholesale argument is able to circumvent the problem of circularity, I will argue that Lipton's response comes at the cost of being unfaithful to the no miracles intuition.

Chapter 3 is dedicated to Worrall's understanding of the no miracles argument – one he refers to as *the 'default' position*. This position holds that the no miracles argument is *just an intuition*. I explain why Worrall's position - in part, a response to Magnus and Callender's paper examined in the first chapter - argues that the no miracles argument ought to be construed as a *retail* argument for realism. Worrall claims that arguments that suggest that the no miracles argument is fallacious, are misguided, and that when construed properly, it no longer falls victim to the supposed fallacy. I argue that Worrall's 'default' position suggests *more* than just the claim that the no miracles argument amounts only to an intuition. In doing so, I attempt to formalise this supposed, broader argument in order to assess, first; whether the default position manages to capture the no miracles intuition, and second; whether this account of the no miracles argument can plausibly defend scientific realism. I argue both, that the default position is not promising in capturing the relevant intuition, and that it fails to furnish a plausible argument.

The final chapter deals with my own formalization of the no miracles argument. I argue that not only does this account succeed in capturing the no miracles intuition, but that it provides a plausible justification of the inference from success to (approximate) truth – something the other interpretations have left wanting. I show that, whilst this formalization encapsulates the practices *actually employed* in the scientific enterprise used to develop and corroborate theories – the activity of *frequentist testing* - it nonetheless does not constitute an *instance of* this method. Moreover, this formalization incorporates the *wholesale* perspective I have explained the no miracles argument ought to encompass, and has as its conclusion that the scientific enterprise gives rise to many more true theories than false ones, and therefore, that science in general is (approximately) true. Moreover, I will argue that this interpretation circumvents problems other accounts are susceptible to – namely, circularity and the base rate fallacy – and that a challenge to this type of interpretation – that of 'alternative hypotheses' - can, at least conceivably, be resolved.

1. The No Miracles Argument in *Bayesian* Terms

> *You are lost in the woods and you come across a topographical map. . . After some exploring, you find that everything the map indicates you ought to see - rivers, lakes, roads, peaks and so on – you do in fact see along your journey. . . You decide to embrace what the map tells you about the terrain in general, because you feel it would be a miracle if the map were, for example, a map of another region, yet happened to be correct about everything you have seen here, up until now. Your belief is warranted by the following facts: the probability that any true map would make successful predictions, is significantly high, and the chances of a false map making accurate predictions on the whole, is very slim. Accordingly, you infer that the probability that the map is correct,* given *that it made such accurate predictions, is sufficiently high to warrant accepting that what it portrays is generally true.*

I begin with an articulation of the Bayesian formalization of the no miracles argument. This account aims to capture an effective wholesale argument for scientific realism. P. D. Magnus and Craig Callender suggest that although this account is *the proper* interpretation of the argument, it is subject to the well-known, *base rate fallacy*, and for this reason, it is not an account they, themselves, endorse. The fallacy will be made perspicuous by analogy to a famous diagnostic case, as well as by an assessment of the logical flaw in the context of the no miracles argument itself. I will discuss how the argument advanced by Larry Laudan, against the no miracles intuition, fits into this context, and I will present Stathis Psillos's response to Laudan's challenge, and suggest as well how this ties into the Bayesian formalization of the no miracles argument. I will present Magnus and Callender's own retail argument for scientific realism; one the authors claim avoids the fallacy. We will see examine an argument by John Worrall, which claims that Magnus and Callender's retail defence of scientific realism is subject, by their own standards, to the same fallacy they caution against. I will discuss, as well, why the retail formalization fails to capture the no miracles intuition. I will mention, briefly, attempts to salvage a Bayesian conception of the retail version of the no miracles argument, and explain why they offer no resolution. My conclusion will be one which rejects a Bayesian interpretation of the no miracles argument – whether wholesale or retail – for not only do the arguments lack cogency, they fail to capture the intuition underlying the no miracles sentiment.[14]

The wholesale no miracles argument according to a Bayesian framework

According to Magnus and Callender,[15] the no miracles argument, properly construed,[16] is an argument construed in probabilistic terms. What they claim is the proper account of the argument, takes the following form: "[f]or any theory *x*, let *Sx* stand for the expression '*x* is successful' and let *Tx* stand for the expression '*x* is true.' Let ¬*A* be the negation of *A* and let Pr(*A*|*B*) be the probability of *A* conditional on *B*. We may now gloss the argument for some current theory, as follows: [1] The theory *h* is very likely successful. [2] If *h* were true, it would be very likely to be successful. [3] If *h* were false, it would not be likely to be successful. [4] Therefore, if *h* is successful, there is a high probability that it is true."[17] Formalized, the argument, as it appears in the authors' exposition, goes as follows:

$$\Pr(\mathrm{S}h) \gg 0 \qquad (1)$$
$$\Pr(\mathrm{S}h|Th) \gg 0 \qquad (2)$$
$$\Pr(\mathrm{S}h|\neg Th) \ll 1 \qquad (3)$$
$$\therefore\ \Pr(Th|\mathrm{S}h) \gg 0 \qquad (4)$$

[18]

So that the explanations that follow are more intelligible, I will label the premises (1) to (4) according to the concepts they represent: Let (1) be *the success rate*; let (2) be *the conditional success rate*; let (3) be *the false positives rate*; and let (4) be *the posterior probability* – the conclusion of the no miracles argument. This argument suggests that when we are confronted with a very successful theory – (1), and we believe, almost with certainty, that a true theory would be successful – (2), then insofar as the false positives rate is very low, it will be extremely likely that this theory is in fact true. Invoking the probability calculus of the Bayesian framework is supposed to offer a formalization of

---

[14] Whilst I allude to reasons why the retail Bayesian formalization fails to capture the no miracles intuition, a more detailed explanation, in general, as to why the no miracles argument *ought to be* interpreted as a *wholesale* argument, is presented in Chapter 2.

[15] Magnus, P.D., & Callender, C. (2004), 'Realist Ennui and the Base Rate Fallacy', *Philosophy of Science*, 71: 320–338.

[16] Whilst Magnus and Callender claim that the wholesale account articulated here is the no miracles argument *proper*, it is not an account they defend, in virtue of the fallacy, as we shall discuss, they argue it falls prey to.

[17] Magnus, P.D., & Callender, C. (2004), p. 323.

[18] Magnus, P.D., & Callender, C. (2004), p. 323. The symbolism ">>0" and "<<1" is intended to mean "close to '1'" and "close to '0'" respectively.

this argument. Suppose, for example, we reasonably assume a theory has: a success rate of 0.9; has a conditional success rate of 0.95; and has a false positives rate of 0.05; then the conclusion of the argument is the derived result that the posterior probability equals 0.997. It may bode us well to have a look at the derivation in detail so that we may better grasp the role some of the terms play in leading to the supposedly desired result, particularly since these details are omitted from Magnus and Callender's exposition.

Using the total probability rule, we can calculate first $Pr(Th)$, the probability that the theory is true, from the numbers given, as follows:

$$
\begin{aligned}
Pr(Sh) &= Pr(Sh|Th).Pr(Th) + Pr(Sh|\neg Th).Pr(\neg Th) && \text{(Law of Total Probabilities)}\\
(0.90) &= Pr(Sh|Th).Pr(Th) + Pr(Sh|\neg Th).[1 - Pr(Th)] && \text{(since } Th \text{ and } \neg Th \text{ exhaustive)}\\
(0.90) &= (0.95).Pr(Th) + (0.05).[1 - Pr(Th)]\\
(0.90) &= 0.95.Pr(Th) + 0.05 - 0.05.Pr(Th)\\
0.85 &= 0.90.Pr(Th)\\
Pr(Th) &= 0.85/0.90\\
&= 0.94
\end{aligned}
$$

Substitute this number into Bayes' Theorem, and the value shown in Magnus and Callender is derived:

$$
\begin{aligned}
Pr(Th|Sh) &= Pr(Sh|Th).Pr(Th)/[\,Pr(Sh|Th).Pr(Th) + Pr(Sh|\neg Th).Pr(\neg Th)]\\
&= (0.95).(0.94)/[(0.95).(0.94) + (0.05).(1 - 0.94)]\\
&= 0.997
\end{aligned}
$$

Accordingly, we ought to accept that given the aforementioned premises, it follows that it is extremely likely that our best, successful scientific theories are true. At face value, this is a powerful argument in defence of scientific realism – not only does it enjoy the rigour of deductive inference, its probabilistic elements make it quantitatively illuminating; the conclusion not only tells us that successful theories are likely to be true, but it offers us a *magnitude* of this likelihood, contingent however, on the initial assignment of the relevant probabilities. Given its *supposed* rigour, it certainly deserves a detailed assessment:

Challenges to this account – the base rate fallacy

Some authors, like Larry Laudan,[19] oppose the facts as conveyed by premise (3) – the false positives rate. As noted in the introduction, Laudan suggests that many past theories that have turned out to be false, are nonetheless successful, warranting a higher probability than that identified with Pr(*Sh*|¬*Th*) in the above reconstruction of the argument. Laudan's inductive argument has as its conclusion the claim that our currently accepted scientific theories are most probably false. However, I mention this only in passing - whether Laudan's argument holds is something I do not intend evaluating.[20] I merely allude to one type of response to the no miracles argument in general - one which rejects one of the premises of the argument.

Magnus and Callender take a different approach and claim that when construed as the type of argument it really is – a probabilistic one captured by a Bayesian framework - the no miracles argument falls victim to the base rate fallacy and so it is logically invalid. This fallacy is one which ignores, as we shall see, the prior probability, Pr(T*h*).[21] The argument is made by analogy to a simple diagnostic case, commonly known as the 'Harvard Medical School Test' from the research of Tversky and Kahneman.[22] They argue it is invalid using the diagnostic case to illustrate the fact, and then claim that there are no repairs to the argument which make it sound and non-trivial.

The authors ask us to consider a rare disease which is undetectable by mere appearances, for which there is a reliable test that can identify those infected. "Let *Dx* stand for '*x* has the disease' and let *Px* stand for '*x* tests positive for the disease.' Now suppose that if someone has the disease, then they are sure to test positive; that is, assume Pr(*Px*|*Dx*) = 1. Suppose further that is someone is not infected there is some small chance that they will nonetheless test positive; that is, there is a chance that a positive test result will be a *false positive*. Let that chance of a false positive be five percent,

[19] Laudan, Larry. (1981).
[20] A well-known response to Laudan's attack on realism is found in Psillos, S. (1999), *Scientific Realism: How Science Tracks Truth.* Routledge: London.
[21] Strictly speaking, 'base rate' expresses the relative frequency of the truth of a theory, and 'prior' expresses the *subjective* probability that it is true - something I deal with later in the chapter under the topic of *subjective Bayesianism*. However, as other authors have done as well, I use these labels interchangeably, whilst still maintaining clarity as regards when the term is derived according to a frequency basis, and when it is derived subjectively.
[22] Tversky, A., & Kahneman, D. (1982), 'Evidential Impact of Base Rates', in Daniel Kahneman, Paul Slovic, and Amos Tversky (eds.), *Judgement under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, pp. 153–160.

i.e. $\Pr(Px|\neg Dx) = .05$. Now suppose a patient *a* tests positive for the disease. What is the probability that she actually has it"[23] – that is, what is $\Pr(Da|Pa)$?

"It is tempting to say that $\Pr(Da|Pa) = .95$ or at least to assign a high value to $\Pr(Da|Pa)$. We can construct the inference so as to be formally analogous to the no-miracles argument as we formulated it above: From $\Pr(Pa) = 1$, $\Pr(Pa|Da) = 1$, and $\Pr(Pa|\neg Da) = .05$, infer $\Pr(Da|Pa) >> 0$."[24] At this stage, Magnus and Callender urge us to take cognisance of the sample from which the patient was drawn – something proponents of the above argument are said to neglect. "Suppose, among the people tested, the disease is rare. If only 1 in 1000 people has the disease, then given the assumptions above we should expect about 51 in 100 to test positive. Of those 51 who test positive, only 1 will actually have the disease. Thus, the chance that this patient who tests positive has the disease would be 1 in 51; $\Pr(Da|Pa) = .02$."[25] To be precise: if 1 person in 1000 – *the base rate* - has the disease; and since, then, $\Pr(Da) = 0.001$ (1 in 1000) and the false positives rate, $\Pr(Pa|\neg Da) = 0.05$ (5 in 100, or 50 in 1000 effectively), it follows that given 1000 individuals selected at random, we would expect about 51 of the 1000 (the sum of the above proportions) to test positive for the disease. Now, we are actually concerned *only* with the 51 individuals we expect to have a positive test result, since our question concerns the chances *a* has the disease *given* that he has *already* been identified as having tested positive. As a reminder, these 51 individuals consist of, the 50 we expect to test positive yet do not actually have the disease (the false positives) and the 1 individual who would test positive because he really *has* the disease (the base rate). So of these 51 people who test positive, we would expect that only approximately 1 individual would in fact have the disease. Thus, the probability that individual *a* - who tests positive - will actually have the disease, is approximately 1 in 51 – i.e. $\Pr(Da|Pa) \approx 1/51 \approx 0.02$.[26] The claim that the posterior probability is necessarily high, is an instance of the base rate fallacy.

[23] Magnus, P.D., & Callender, C. (2004), pp. 324-325.
[24] Magnus, P.D., & Callender, C. (2004), p. 325.
[25] Magnus, P.D., & Callender, C. (2004), p. 325.
[26] The demonstration here is of course an approximation, used in place of the Bayesian calculus to make the derivation more accessible. The precise calculation, using, once again, Bayes' Theorem, requires the following: instead of calculating $\Pr(Da)$ by invoking the Law of Total Probability (as was done earlier), we simple make use of the specified base rate - 0.001 - the probability that an individual has the disease - and derive the result as follows:

$$\begin{aligned}\Pr(Da|Ta) &= \Pr(Ta|Da).\Pr(Da)/[\Pr(Ta|Da).\Pr(Da) + \Pr(Ta|\neg Da).\Pr(\neg Da)]\\ &= (0.95).(\mathbf{0.001})/[(0.95).(\mathbf{0.001}) + (0.05).(1 - \mathbf{0.001})]\\ &= 0.018664\end{aligned}$$

This, compared to 1/51 = 0.019608, makes the latter a good approximation.

Of course, the chances that individual *a* has the disease is much higher once we know he definitely tests positive – from 1 in 1000 to 1 in 51 – however, by no means is it *likely*. Thinking that the probability is 0.95, or at least some high value, is a result of ignoring the base rate – a consequence of neglecting the fact that the results of non-diseased people swamp the results for those diseased. Even though the false positives rate is low; when the condition is rare, being confronted with a positive test result should not make us believe that there is a good chance the patient has the disease.

The base rate fallacy besets the no miracles argument for scientific realism in much the same way as it does the diagnostic case. In the context of the former, being ‘true’ is analogous to ‘having the disease’, and ‘making lots of accurate predictions’ is analogous to ‘testing positive for the disease’. Under the reasonable assumption that true theories do not generate false predictions, the false negatives rate – $\Pr(\neg Sh|Th)$ - is zero. Ignoring Laudan’s contention, the false positives rate is low since only a small proportion of false theories end up generating accurate predictions. Suppose that 1 in 1000 theories are true – i.e. $\Pr(Th) = 0.001$, and suppose the false positives rate, $\Pr(Sh|\neg Th) = 0.05$. Then if we were to sample 1000 theories, we would expect 51 in 1000 to be successful (for the same rationale as per the diagnostic case). Now, if we consider the 1000 theories, and ask, “what is the probability that a theory is true *given* that it is successful;” our population under consideration is now the 51 successful theories. But we know that out of the 1000 theories, we would expect only 1 to be true. It follows that $\Pr(Th|Sh) = 1/51 = 0.02$. The posterior probability is in fact *small*, thus severing the explanatory connection between truth and success. To claim that $\Pr(Th|Sh)$ *must* be high is “the false positives fallacy, a form of base-rate neglect.”[27]

What goes wrong in the no miracles argument, just like the diagnostic case, is that, in inferring that theories that enjoy great predictive success are likely to be true, the base rate is ignored. If the base rate tells us that most theories are false, then *despite* the unlikely event that a false theory generates successful predictions, we still need to remember that most successful theories are nevertheless false *anyway*. As Lipton says: “Beware false positives, even if unlikely, because the truth is rare.”[28] Lipton suggests that we are prone to committing the fallacy because we seem to be compelled to focus on comparing, the relatively few false yet successful theories with the many false unsuccessful theories, whilst *failing to recognise*, more importantly, the comparison between the

[27] Magnus, P.D., & Callender, C. (2004), p. 325. This point, in the context of the no miracles argument, was first made in Howson, C. (2000), *Hume’s Problem: Induction and the Justification of Belief*, Oxford: Clarendon Press, pp. 52-54. For interest sake, it has been shown that the same fallacy besets the pessimistic induction argument, in Lewis, P. (2001), ‘’Why the Pessimistic Induction is a Fallacy’, *Synthese*, 129: 371–380.
[28] Lipton, P. (2004), p. 197.

very few true theories and the relatively many more false, successful theories. It may now be perspicuous why Putnam's iconic statement, that it would be a miracle if a false theory generated many true predictions, may elicit the wrong type of thinking when using a probability calculus: the emphasis of his claim sways our attention to the very low proportion of successful theories amongst false theories, whilst diverting our attention away from the crucial fact – that the proportion of *true* theories amongst successful theories is significantly small, precisely because true theories, in general, are rare.

In an attempt to flesh out, even further, the fallacy that besets the no miracles argument, Magnus and Callender propose a reformulation of the argument in the way set out below; a revision they suggest exposes the argument's hidden premise, and as a result, loses some of its probative force:

Suppose *h* is some current theory of a mature science, where $\mathcal{H}$ is the set of present candidate theories, then the reformulated argument is as follows:

$\Pr(Sx|x\epsilon\mathcal{H}) >> 0$ (1*)

$\Pr(Sx|Tx \,\&\, x\epsilon\mathcal{H}) >> 0$ (2*)

$\Pr(Sx|\neg Tx \,\&\, x\epsilon\mathcal{H}) << 1$ (3*)

$\therefore \Pr(Tx|Sx \,\&\, x\epsilon\mathcal{H}) >> 0$ (4*)[29]

Consider the supposed hidden premise, that $\Pr(Tx|\ x\epsilon\mathcal{H})$, the base rate, is low, perhaps because it is impossible that members of a set of inconsistent theories are individually, each likely to be true, or perhaps merely because we cannot be sure that it isn't low! Now, the conclusion of the argument, as per above, is that $\Pr(Tx|Sx \,\&\, x\epsilon\mathcal{H})$, the posterior probability, >> 0; that is, the probability is very high that an arbitrary current theory of a mature science is true given that it is successful. But since $\Pr(Sx|x\epsilon\mathcal{H})$, the success rate, >> 0 , then it follows from the premises and conclusion together, by *modus ponens*, that there is a high probability that any current theory of a mature science is true; that is, $\Pr(Tx|\ x\epsilon\mathcal{H})$, the base rate, is high. This contradicts our hidden premise, that $\Pr(Tx|\ x\epsilon\mathcal{H})$ is in fact low. So the conclusion of the no miracles argument, that $\Pr(Tx|Sx \,\&\, x\epsilon\mathcal{H})$ is high, does not follow.

[29] Magnus, P.D., & Callender, C. (2004), p. 325.

Failed attempts to avoid the fallacy

Evaluating the success rate

As we see, it is the success rate, Pr($Sx|x\in\mathcal{H}$) >> 0, which sets up the contradiction. Proponents of the no miracles argument may protests, and suggest that we can avoid this unwelcome result and evaluate the success rate, by examining the pool of theories, $\mathcal{H}$, in order to discern what proportion of them are really successful. However, such a task seems futile – the concept 'present mature science' is at best a vague one, making it difficult, if not impossible, to tally or even sample the theories we would think we could identify as constituting a mature science.[30] Some may suggest that the candidate theories are precisely those that scientists deem as falling into the category of mature sciences. However, this raises issues: these theories would have been chosen as falling within the mature sciences *because*, in no small part, they *are* successful in the first place; therefore, it would trivially be the case that any theory $x\in\mathcal{H}$ would be successful just in virtue of it being a member of $\mathcal{H}$. Consequently, (2*) would hold by default, and not because it is the truth of the theories that renders them so, as (2*) implies. Moreover, (3*) would be outright false, because almost all the candidate theories would be successful, whether true *or* false, so instead of Pr($Sx|\neg Tx \ \& \ x\in\mathcal{H}$) – the false positive rate - being low, the probability would be very high. So although the argument would circumvent the base rate fallacy that besets it ordinarily, it would do so at the cost of 'sample-selection' bias, whereby the conclusion, instead of exposing real connections between truth and success, would expose ones which merely illuminate a fact about a manufactured situation.[31] Evaluating whether the success rate is correct does not present itself as a way out of this dilemma.

It seems then that proponents of the no miracles argument cannot escape the fallacious reasoning. The lesson to be learnt, according to Magnus and Callender, is that if true theories are rare – i.e. Pr($Tx|\ x\in\mathcal{H}$) is low - then Pr($Tx|Sx$) despite being greater than Pr(Tx), will still be very low, and the conclusion of the no miracles argument fails.

Reducing the false positives rate

Psillos's response to Laudan, as mentioned earlier, can also be seen as an indirect response to the challenge presented by Magnus and Callender. Realists may attempt to circumvent accusations

[30] Worrall, J. (2005).
[31] Magnus, P.D., & Callender, C. (2004), p. 326.

which assert that Pr(*Tx*|*Sx*), instead of being high as proponents of the no miracles argument insist, is in fact low, by decreasing Pr(*Sx*|¬*Tx*), the probability that a theory is successful given that it is false. Besides whittling down Laudan's historical gambit, this move has the mathematical consequence of decreasing the magnitude of the denominator of Bayes' Theorem, thus resulting in a greater posterior probability. To be certain, recall Bayes Theorem:

Pr(*Tx*|*Sx*) = Pr(*Sx*|*Tx*).Pr(*Tx*)/[Pr(*Sx*|*Tx*).Pr(*Tx*) + **Pr(*Sx*|¬*Tx*)**.Pr(¬*Tx*)]

The bolded probability is our aforementioned probability under consideration. Note that this term contributes positively to the denominator. Therefore, all else being equal, as Pr(*Sx*|¬*Tx*) decreases, so too will the magnitude of the denominator. Moreover, as the denominator decreases, it means we are dividing the numerator by a relatively smaller number. This has the equivalent effect of multiplying the numerator by a relatively *larger* number, resulting in a final value greater than before. This move has mathematical rigour, however, a new challenge arises: how does one show that Pr(*Sx*|¬*Tx*) is not only *lower* than made out to be by Magnus and Callender, but *sufficiently low* to effect a value for Pr(*Tx*|*Sx*), high enough to justify a no miracles position?

Psillos's strategy - "success too-easy to-get"[32] - involves tightening the realist's criterion of empirical success, thereby restricting the pool of theories under consideration to meet specific, more stringent conditions, thus reducing Pr(*Sx*|¬*Tx*).[33] According to Laudan, for example, a theory is successful provided it has functioned in a variety of explanatory contexts, has led to several confirmed predictions, and has been of broad explanatory scope. However, Psillos urges that empirical success should be more precise than simply offering a narrative which fits the observable facts, for this, he notes, could turn any theory (no matter how outrageous) into a so-called successful one by doctoring it to have as its consequences the ones actually observed. Psillos suggests that the notion of 'empirical success' that realists ought to embrace is one which involves "the generation of novel predictions which are in principle testable."[34] The no miracles argument then takes the following, refined form: the *genuine* success of scientific theories would be miraculous if those theories were not (at least approximately) true.

[32] Psillos, S. (1999), p. 104. Discussion in the following two pages of Psillos's strategy *in general* has been adapted, in part, from my essay Leader, S. (2011); Its relevance, however, to the Bayesian framework is new.
[33] As alluded to, although I invoke Psillos's strategy as an argument which aims to reduce Pr(*Sx*|¬*Tx*), it is in fact a direct response to Laudan's pessimistic induction argument, and aims to reduce Laudan's historical gambit of genuinely successful yet false theories. In this context, Psillos aims to show that Laudan's list includes theories that did not, under this new criterion, enjoy the success he claims they did.
[34] Psillos, S. (1999), p. 105.

According to Psillos's novel-prediction standard, a successful theory is one which predicts phenomena which display either 'temporal' novelty or 'use' novelty, or both.[35] Respectively, this means that the theory must predict either 'new' phenomena which are not known about when the theory is propounded, or 'new' phenomena which are known whilst no information about the phenomena is used in constructing the theory. Although many past theories seem to be able to 'save the phenomena', this on its own, claims Psillos, is insufficient for the *genuine* empirical success of a theory.

The process of accommodation of the phenomenon within the theory, without having affected its construction, Psillos calls *novel accommodation*. The relevance of this notion comes in its contrast with *ad hoc accommodation*. Psillos tells us that a theory *T* is said to be *ad hoc* as regards phenomenon *E* provided at least one of two conditions are met:

1 A body of background knowledge *B* entails the existence of phenomenon *E*. Information about *E* is used in the construction of theory *T*, and *T* accommodates *E*.
2 A body of background knowledge *B* entails the existence of phenomenon *E*. A certain already available theory *T* does not predict/explain *E*. *T* is modified into *T'* so that *T'* predicts *E* but the *only* reason for this modification is the prediction/explanation of *E*. In particular, *T'* has no other excess theoretical and empirical content over *T*.[36]

In summary, "a prediction *P* of a phenomenon *E* is use-novel with respect to a theory *T* if *E* is known before *T* is proposed, *T* does not satisfy either of the ad hocness conditions and *T* predicts *E*."[37]

Even if we were to presuppose that Psillos's strategy correctly distinguishes between actually successful, and unsuccessful theories, thus reducing Laudan list of "successful" yet false theories, realists will nevertheless *still* struggle to advance the success-to-truth inference they require. Indeed, identifying successful theories may increase the chances that we will be confronted with one that is true, just as restricting the pool of 1000 theories, to 51 theories that test positive, will increase the chances that a theory is true, from 1/1000 to 1/51. But although a strategy like Psillos's will shift $\Pr(Sx|\neg Tx)$ positively by limiting the size of the pool from which we would expect a true theory to be found, the shift is far from sufficient, because we still need to consider base-rates. For

[35] As suggested to me by Worrall, I ought to point out that the original account of use novelty goes back, at least, to Whewell, Lakatos and Zahar. It is not directly attributable to Psillos, as my ommissions of their references may suggest.
[36] Psillos, S. (1999), pp. 106-107.
[37] Psillos, S. (1999), p. 107.

example, suppose we sample 1000 theories and suggest that Psillos's strategy is so discriminating that instead of 0.05, $\Pr(Sx|\neg Tx) = 0.005$ – i.e. we would expect only 5 in 1000 theories to be successful despite being false. Nonetheless, if true theories are rare enough, say, still 1/1000, then we would expect 6 theories in total to be successful, one of which is true. Therefore, $\Pr(Tx|Sx) = 1/6$; certainly not nearly high enough to warrant a success-to-truth inference.

The obvious realist response is to argue that true theories are not as rare as the argument suggests – that is, to appeal to a greater base rate, in which case $\Pr(Tx|Sx)$ can be shifted to something at least greater than 0.5. This move seems to work – why can't we just assume a different base rate – after all, 1/1000 was chosen just as arbitrarily? Appearances, however, are deceptive: as Lipton suggests, proponents of the no miracles argument are faced with a dilemma: either it is possible to know the approximate base rate of truth amongst current theories, or it is not. If it is possible, then it would mean that there exists *independent* justification for believing that it is likely that a theory is true, in which case, the no miracles argument is not required at any rate to defend realism. However, if it is not possible, then the only way of arriving at a satisfactory answer for $\Pr(Tx|Sx)$ is to assume that some significant proportion of our present theories are true, which only begs the question against those who do not endorse the realist's conclusion. If proponents of the no miracles argument hope to defend realism effectively, then using a Bayesian framework is not the way to do it.

Why Bayesian Retail arguments fail to capture the no miracles intuition

Worrall[38] argues that Magnus and Callender's own position is incoherent; these authors embrace the idea of *retail* arguments as concerns scientific theories, after demonstrating that wholesale arguments are subject to the base rate fallacy. However, Worrall argues, insofar as their conception of retail arguments make a surreptitious appeal to the no miracles argument, these retail arguments fall short by Magnus and Callender's own standards.

Worrall believes that wholesale arguments are implausible, not because they are subject to the base rate fallacy, but because there is no such thing as science (in general) and not all sciences elicit the no miracles intuition. He claims that rejecting the wholesale argument and thus embracing a collection of retail arguments does not absolve realists from having to defend the intuitions underlying the no miracles argument.

[38] Worrall, J. (2005).

With the aim of exposing the underlying fallacy, Worrall formalizes the retail argument as concerns a *particular* theory as follows:

$p(e|T) = 1$ (assuming the necessary auxiliaries, e is entailed by T)
$p(e|\neg T) \approx 0$ (it would be a miracle if e occurred were T not true)
Therefore $p(T|e) \approx 1$ and hence, given that e has occurred, $p(T) \approx 1$.[39]

As mentioned, Worrall suggests that the reasoning employed above is fallacious by Magnus and Callender's own standards; any non-extreme probability of T (even one close to zero) is in fact compatible with the truth of the two premises, and the precise value of T given e depends on the value of the prior probability of T.

We may also question whether the intuition of the no miracles argument is properly captured by the probabilistic formalization as construed above. If it is, Worrall suggests, then given the fallaciousness of the argument, the intuition underlying the no miracles argument ought to be surrendered, no matter how appealing. So it is Worrall's aim to show that the formalization does not capture the intuitions, so that the no miracles argument does not necessarily have to be dismissed. Worrall asks how we can possibly interpret the relevant base in, for example, the case of the wave theory of light and Fresnel's white spot or some other theoretical success that elicits the no miracles intuition. To elucidate this concern, Worrall appeals to the diagnostic case, and argues that the case for particular scientific theories is by no means equivalent:

In order to discern probabilities in the diagnostic case, one would simply sample the population and the relative frequency of those who tested positive yet did not have the disease would converge to some proportion as the sample size increased. Similar would be the case for the incidence of the disease itself. However, the case is not as clear cut when it comes to interpreting the probabilities (a) the probability that evidence e would occur if theory T were false, and (b) the 'prior' probability that T is true. How would we sample theories from a population of theories? How would we "draw" a theory and test if it was true? Moreover, what would the 'population' be? And our conception of 'population' is confounded by the fact that, as argued earlier, we are concerned only with *retail* arguments; and so a population consisting of a plethora of theories which are in

39 Worrall, J. (2005), p. 14.

some generic sense "successful" yet unrelated to the predictive success in question, seems implausible. Perhaps, one could argue, the population can consist of the set of all possible alternative, rival theories to the theory at hand. Worrall is doubtful that even such a restriction would be helpful at all; there is no limit to the number of rivals available if we allow for 'audacious' alternatives, particularly in light of Jeffreys-style alternatives in the case of mathematically expressed theories.[40] And if this is the case, any objective interpretation of prior probabilities, given an arguably infinite set of alternatives, is elusive. And as far as p(e|¬T) is concerned, it seems we would need to check the proportion of alternatives to T for which e holds, but our assumption that these alternatives are plausibly realistic (and even if they are, that they hold equal weight) would be difficult if not impossible to ascertain.[41]

Worrall insists that someone who still seeks a population from which T is drawn, would have to restrict the set of alternatives to T, but would do so to no avail: if we restrict the set to T's *active rivals*, then we would characteristically be considering just one theory T', in which case p(e|¬T) is identified with p(e|T') and T' would usually entail ¬e. Here, p(T|e) is trivially equal to 1 (analogous to the logical rule of deductive syllogism, and equivalent in the diagnostic case to there being no false positives). Although the argument here would circumvent the fallacy, it is of no use to the realist as it has nothing to do with the issue addressed by the no miracles argument. The realist is not only concerned with present rivals to T, but future, possible rivals which could entail e, whilst being radically different from T (and thus possibly indicating that T is false despite its predictive success). If T got some prediction correct and was radically false, yet some present, rival theory of T, say T', also entailed e and implied that T was radically false, no one would claim that T's successful prediction would have been miraculous. It could be that T's so-far success only *seems* otherwise miraculous because we are as yet unaware of some unarticulated theory T'' that also has this predictive success, is epistemically more virtuous than T, and which entails that T is radically false. Here, using the no miracles argument to infer that T'' was (approximately) true, means that since T is radically false in light of T'', T's success with e was miraculous. The point is that if in order to address the issue of the no miracles argument, we need to invoke the idea of rival theories which have not yet been articulated and hence ones we are not yet aware of, then it is impossible to discern a population from which we can consider a theory T to be drawn – there is no way to restrict the set of alternatives to T in any feasible way.

[40] I explain in more detail, in Chapter 2, the challenge of 'Jeffreys-style alternatives.'
[41] Worrall, J. (2005).

Doubtful attempts to salvage retail versions of the argument

Some suggest that the aforementioned problems can be avoided if the base rates are seen as prior probabilities according to a *subjective Bayesian* conception. *Subjective, or Personalist Bayesians* suggest that their refined conception offers a way around the problem regarding 'populations of concern', by assigning values to prior probabilities according to the degree of belief, on the part of some individual, that the hypothesis under consideration is true. According to this retail account then, the posterior probability is this person's degree of belief that the given hypothesis is true after certain evidence is accepted. Proponents of this view claim that since, according to this account, the true probability or base rate of the hypothesis is no longer necessary for the calculus, no appeal needs to be made to some elusive *class* of theories in order to determine this term – the subjective assignment will suffice. Although this may eliminate worries about populations of concern, it raises worries of its own: First, anti-realists would not be satisfied with *subjective* assignments of probabilities which result in high posterior probabilities. This, opponents would argue, begs the question of why the priors deserve high assignment values in the first place. Where, after all, do these prior probabilities actually come from, and what constitutes "accepting the evidence"? The problem this account faces is that, given its generality, there are insufficient restrictions on what can warrant belief in hypotheses and evidence such that the posterior probability can be determined satisfactorily. Moreover, any sufficiently high assignment value to priors can be viewed simply as question-begging against the anti-realist – it is precisely the realist's burden to justify *why* it is likely that successful theories will be true; if this justification involves an assumption that *it just is likely* – i.e. that $p(T)$ is greater than say 0.5, then the argument lacks any force. Besides, if $p(T)$ is sufficiently high, then so too is $p(T|e)$; so the latter will be high just in case the theory is likely to be true *anyway*, *whether or not* e has occurred. And adopting, according to this subjective view, a wholesale account of the no miracles argument only make things more problematic; in the wholesale case, where the prospect of realism is entirely dependent upon the priors, it seems impossible to even begin to conceptualise how one could determine a reasonable set of priors.

So there does not appear to be a way to formalize the no miracles argument according to a Bayesian framework– whether wholesale *or* retail - without falling prey to a fallacy, or being confronted with constituent probabilities which cannot be understood in any reasonable way. It seems impossible, for example, to ascertain some objective ratio of true theories amongst all past theories. And even if this ratio does exist, either, there is no way in practice to determine it, or if we could, the no miracles argument would not be required in the first place. Moreover, when viewing the critical probability p(e|¬T) as representing the ratio of, possible alternatives to T which also predict e, to *all the possible alternatives to T*, not only can we not make any sense of this set of theories, we have little reason to believe that all the possible alternatives carry equal weight. And even if we could restrict possible alternatives to those we are aware of (and assume they carry equal weight), we arrive at trivial answers that are not faithful to the no miracles intuition. The Bayesian interpretation, I have argued, neither affords us a plausible defence of scientific realism, nor does it capture the essential ingredients of the no miracles intuition.

2. The No Miracles Argument as an *Inference to the Best Explanation*

You are lost in the woods and you come across a topographical map. . . After some exploring, you find that everything the map indicates you ought to see - rivers, lakes, roads, peaks and so on – you do in fact see along your journey. . . You decide to embrace what the map tells you about the terrain in general, because you feel it would be a miracle if the map were, for example, a map of another region, yet happened to be correct about everything you have seen here, up until now. You feel it would be a miracle if the map were, for example, a map of another region, yet happened to be correct about everything you have seen here, up until now. You believe that it being the true *map of the region, would, if true,* best explain *why it has achieved such astounding predictive success.*

This chapter will consider the no miracles argument as an Inference to the Best Explanation. I will begin with a detailed explanation of this general form of inference, as formulated by Peter Lipton.[42] I will examine, in alignment with Lipton's account, various ways in which something may be an *explanation*, as well as competing views on the ways in which something may be the *best* of competing explanations. I consider once again, the scope of the no miracles argument, and argue that it is the nature of the no miracles argument, as conveyed by the so-called 'textbook' formulations, and for the purposes for which it is appealed to, that it *ought to be* interpreted as a *wholesale* argument. I examine a wholesale account of the no miracles argument, as an Inference to the Best Explanation. Such an account begins with the description that scientists employ Inference to the Best Explanation to infer that a claim is (at least approximately) true, whether it concerns something observed, unobserved but observable, or something unobservable. Granting this descriptive assertion about what scientists do, the question regarding the *justification* of such an inference is what will be scrutinised; I will examine whether these inferences to the best explanation, that scientists make, really are truth-tropic – whether they reliably take scientists towards the truth. I will present a crucial challenge for this wholesale conception, namely, *the circularity objection*. Although Lipton suggests that the wholesale argument is able to circumvent the problem of circularity, I will argue that Lipton's response comes at the cost of being unfaithful to the no miracles intuition.

[42] Lipton, P. (2004).

Inference to the Best Explanation

Some authors have articulated the no miracles argument as an *Inference to the Best Explanation*. Proponents of this interpretation invoke the success of scientific theories, and infer that, the best explanation for this success would be that these theories are (at least approximately) true. Here, the reason for being a realist about science is that it is the best explanation of scientific success.

Inference to the Best Explanation is a form of reasoning which establishes a connection between the practices of explanation and the practices of inference. Lipton begins by offering us a simple characterisation of the relationship between inference and explanation:[43] We begin with our inferences we already have – claims that have already been made about the world. Then, we wish to explain some phenomenon, and we appeal to a subset of our beliefs we hold that are relevant to the situation – a subset consisting mainly of those prior inferences. However, in the case that none of these inferences, nor any conjunction thereof, constitute the explanation we seek, we probe for another uncharted, warranted inference that will do the explaining – an activity that itself may necessitate further investigation. According to this picture, although explanatory considerations guide inference to the extent that they focus our inquiry, inference nonetheless precedes explanation. This seems right given that in order to constitute an explanation, a necessary, basic requirement is that the explanatory information ought to be correct – precisely the type of thing our inferential practices aim to do.

However, as Lipton hastens to add, the aforementioned picture, which views inference as coming before explanation, undermines the importance of explanatory considerations in inference; it is not just the case that explanatory considerations *focus* our inferential inquiry – telling us what we need to look for, they are also indispensible in telling us whether we have *found* it. Suppose you are aware that Bob and George have recently had a heated argument which resulted in the end of their friendship.[44] Today you saw Bob and George walking together along the beach. The best explanation for this, you think, is that they have resolved their differences. Your conclusion is that they must be friends again.

[43] Lipton, P. (2004), p. 55.
[44] The example is adapted from Douven, I. (2011), 'Abduction', *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2011/entries/abduction/, accessed on the 24th September, 2012.

Of course, your conclusion is not *entailed* by the premises – the fact that Bob and George are friends again is not logically necessitated given the premises that, they had a heated argument which ended their relationship and, that you have just seen them walking together along the beach. It wasn't even the case that you had sufficient statistical information – probabilities as per the Bayesian account - about friendships, heated arguments, and people walking together, that may justify making an inference from the information you possess regarding Bob and George, to the conclusion that it is probable that they must be friends once more. Instead, your conclusion could be said to be warranted precisely because Bob and George's being friends again would, if true, *best explain* them having been seen strolling together. This form of inference – often termed *abduction* – is what philosophers refer to as *Inference to the Best Explanation*.

Since Inference to the Best Explanation does not involve a move which is logically necessitated, from premises to conclusion, one may argue that there may be more than one explanation for some fact or phenomenon. How can we be sure, for example, that instead of now being friends again, Bob and George still have not gotten over their row and simply both happened to be walking along the beach unaware that they were walking alongside one another? Or perhaps they were on their way to a secluded part of the beach where they had decided to have a dual. These are competing explanations, albeit, contrived ones. The point of Inference to the Best Explanation is not merely to infer a *possible* explanation, but instead, the *best* of the alternative explanations.

However, "[Inference to the Best Explanation] still remains more of a slogan than an articulated account of induction."[45] Before we begin to assess the cogency of Inference to the Best Explanation as an account of the no miracles argument, we need to examine the supposed appeal of this account, in general, and understand its apparent shortcomings. For now, our description of inference and explanation, and the relationship between the two, will be done under the presupposition of inferential and explanatory realism – a presupposition we will in fact be scrutinizing when we employ this approach for an account of the no miracles argument. Accordingly, in Lipton's discussion, the assumption is that an aim of inference is truth, and that our actual inferential practices are truth-tropic – that is, "that they generally take us towards this goal, and that for something to be an actual explanation, it must be (at least approximately) true."[46]

[45] Lipton, P. (2004), p. 57.
[46] Lipton, P. (2004), p. 57.

The best* EXPLANATION; *and the* BEST *explanation

In order to better understand Inference to the Best Explanation, we need to distinguish, first, between alternative conceptions of *explanation*, and second, between competing conceptions of the *best* explanation. Lipton discusses two potential accounts of 'explanation' - *actual* and *potential* explanations. An actual explanation is one which is necessarily true, whereas a potential explanation is one which simply *may* be true. If we are aiming at articulating an account of our actual inferential practices, Inference to the Best Explanation cannot be understood as Inference to the Best Actual Explanation, since we do not *always* infer truths, and never falsehoods - no matter how reasonable the inferences are. Moreover, it is the nature of Inference to the Best Explanation that we consider a pool of candidate explanations, from which we infer the best one. But if all the candidates are *actual* explanations, it means they are all already true. Of course, this cannot be the case given the possible incompatibility of explanations, as well as the fact that there simply would be no use for Inference to the Best Explanation if all candidate explanations selected were *already* true – we wouldn't require a truth-tropic mechanism to *select* the best explanation.

Lipton's solution to the aforementioned problem is to adopt the approach of Inference to the Best *Potential* Explanation. From our pool of *potential* explanations, he suggests, we infer the best one. The only requirement for constituting a potential explanation is that it consists of a general hypothesis which describes why the phenomenon, contrary to its 'miraculous' explanation, is more likely to occur – there is no requirement that the explanation is true.[47] This way, the pool of candidate explanations may contain incompatible explanations, without concern, and instead of inferring the best actual explanation, we can infer that the best of the potential explanations on offer *is* an actual explanation.

Lipton also distinguishes between two ways in which something may be the *best* of competing explanations: the *likeliest explanation* and the *loveliest* explanation.[48] The likeliest explanation is the one which is the most probable explanation – the one which is backed by the most evidence, and the loveliest explanation is the one which, if correct, would provide the best understanding. Although both standards may pick out the same explanation, of course, this is not necessarily the case (and perhaps it is only seldom the case). Likeliness is concerned with truth, whereas loveliness concerns potential understanding. An explanation can be likely yet unlovely, or unlikely yet lovely. Lipton suggests that it is not a sufficient condition for constituting the best of competing

[47] Lipton, P. (2004).
[48] Lipton, P. (2004), p. 59.

explanations that it be the likeliest explanation; if we opt for likeliness, we can't be certain that this takes us beyond mere triviality. "It is extremely likely that opium puts people to sleep because of its dormative powers (though not quite certain: it might be the oxygen that the smoker inhales with the opium, or even the depressing atmosphere of the opium den)"; despite being the likeliest explanation, it is certainly not the loveliest explanation. In contrast, Lipton suggests that a conspiracy theory, for example, may offer great explanatory power, in the sense that it would help us understand how many, seemingly independent occurrences emanate from, in fact, a single event, and that a number of events we may view as coincidences are actually interconnected. This conspiracy theory would certainly offer a lovely explanation, if true, but concerns about its likelihood would be rather obvious.

It seems appealing to go with the approach which favours strong inductive arguments – one whereby the premises warrant a likely conclusion, but like the example above, the likeliest explanation may serve no purpose beyond triviality. Lipton explains that we require an account of Inference to the Best Explanation which stipulates *criteria* according to which we can identify a more, or most likely inference, amongst others. We need principles in order to spot likeliness – opting for an account which demands, as its basic condition, that we choose the likeliest explanation begs the question of how we are supposed to do this. It is precisely Inference to the Loveliest[49] Explanation, Lipton claims, which is supposed to offer the instructions to do so, using explanatory considerations.

Lipton therefore advocates *Inference to the Loveliest Potential Explanation* as the version of Inference to the Best Explanation that is the most defensible account of scientific inferences, since it holds that the explanation that, if true, would offer the clearest understanding, is that one which is most likely to be true. Lipton concedes, however, that he does not wish to offer a complete account of scientific inferences, let alone scientific practice in general. However, he insists that insofar as explanatory considerations guide us towards truth, then an account which embraces this sentiment, by linking truth and understanding in a fundamental way, must surely be on the right track; "[t]he

[49] Lipton has a lot to say about explanation and what makes something the loveliest explanation – see Ch 4. *Contrastive Inference* in Lipton, P. (2004), pp. 55-71. Lipton's standard of loveliness is one which aims to formulate criteria for the purposes of choosing between competing, scientific hypotheses, and is therefore a standard which is particular to a *retail* conception of the no miracles argument. Of course, it is important to understand this standard in detail if we want to assess properly the feasibility of his retail interpretation of the no miracles argument. However, since I argue further on that the no miracles argument *ought to be* interpreted as a *wholesale* argument for realism, I will confine the discussion to those elements of Inference to the Best Explanation which apply *in general* – for the purposes of an understanding of *wholesale* interpretations - so as not to detract from my ultimate goal of presenting a formalization of *this* scope.

explicit point of explaining is to understand *why* something is the case but, if Inference to the Best Explanation is correct, it is also an important tool for discovering *what* is the case."[50]

Settling on the scope of the no miracles argument – why the argument ought be a **wholesale** *argument*

Before we look at a wholesale interpretation of the no miracles argument, formalized as an Inference to the Best Explanation, I suggest we reflect on the ideas thus far as regards the intended scope of the no miracles argument. After all, in order that we eventually come to some conclusion about the proper formalization of the argument, we will need to cash out the scope the no miracles argument *ought to* incorporate.

Scientific realism is a philosophical thesis - one which aims to say something more than just the conclusions or claims that scientists arrive at. After all, scientists, generally speaking, are invested in corroborating their theories or claims – they seek evidence which supports the propositions encapsulated by their theories. In short, they wish to test their theories to assess whether or not they are *true*. Lest scientific realism is superfluous, its core thesis is to say something over and above merely what scientists do. Retail arguments for scientific realism – ones which claim that particular theories are true do not offer us much beyond those conclusions scientists arrive at. The point, therefore, is not only that retail arguments for realism are redundant, but that the no miracles argument proper is one which is consistently appealed to in order to defend the view that we ought to hold true that our best scientific *theories*, *on the whole*, are (at least) approximately true. Some philosophers may, in the end, wish to be realists only about particular scientific theories, but then they are not practicing philosophy; this is precisely the business of scientists. Granted, inquiring about what it means even for a particular theory to be *true*, is a philosophical question, but the emphasis of this type of inquiry is not what is at stake here. An examination of what it *means* to be 'true' is a question better left to philosophers of language and logic. The realism/anti-realism debate is not a debate, at least not directly, about what it means for scientific theories or a particular theory to be true. What my exposition aims to uncover is the elusive connection between success and truth; how that putative connection, in the context of scientific theories, ought to be interpreted, and how, if possible, it can be justified.

[50] Lipton, P. (2004), p. 66.

Moreover, if we refer back to those statements - the 'textbook' formulations of the no miracles argument - in the introductory chapter, individually conveyed by Putnam, Worrall[51] and Boyd, there is no indication that the connection between success and truth is being made in reference only to *particular* theories. The no miracles argument, *as it has been appealed to*, quite clearly seems to be a wholesale argument, about our best scientific theories in general. One could argue that I have only sampled those formulations which are of the wholesale species, and that if we were to consider, for example, the claim by Lipton, we would see that there is an appeal only to *particular theories*. However, Lipton himself makes it explicit that his retail argument for realism is *not* intended to be an interpretation of the no miracles argument. Lipton, as mentioned, insists that the no miracles argument *ought to be* considered a wholesale argument, but argues that it is ineffective in its defence of realism. For this reason, he advances a retail argument for realism, but a retail argument for realism is all it is; it is not intended to be an argument which captures the no miracles intuition.[52]

If we wish to remain faithful to the no miracles intuition and propound an argument which does not simply *amount to* the practices employed by scientists, but says something more about science on the whole, then I contend that the no miracles argument ought to be interpreted as a *wholesale* argument for scientific realism.

A wholesale *no miracles argument according to Inference to the Best Explanation*

As discussed, Lipton embraces realism for *particular* entities, or theories, and concedes that his formalization is not supposed to account for the no miracles argument proper. However, it is not my intention to defend realism, nor to assess arguments for realism in general; instead, I wish to attempt to answer to the problem of *how the no miracles argument can be captured effectively*? For

[51] Granted, Worrall is no longer of this opinion.

[52] Lipton goes for an account which, instead, appeals to the structure of scientific Inference to the Best explanation, instead of an all-pervasive inference of the same form, as is the case with the no miracles argument. "If I were a scientist, and my theory explained extensive and varied evidence, and there was no alternative explanation that was nearly as lovely, I would find it irresistible to infer that my theory is approximately true. It would seem miraculous that the theory should have these explanatory successes, yet not have something importantly true about it." Lipton believes that the intuition that it would seem miraculous that a theory enjoyed predictive success yet turned out to be false, does not depend on the no miracles argument, and he argues that his account which relies on the structure of *scientific* Inference to the Best Explanation is able to capture the persistent intuition underlying the argument without being subject to criticisms of circularity and that of the underdetermination. He does not think that the truth of the theory is the best explanation of its predictive success, but rather that the theory offers the best explanations of the phenomena that the evidence describes; ". . . the best evidence for scientific realism is the scientific evidence, and the structure of the methods scientists use to draw their inferences from it." Lipton, P. (2004), p. 206.

the purposes of this exposition then, and the chapter under consideration, I will examine what Lipton believes is the *intended* formalization of the no miracles argument – a *wholesale*, Inference to the Best Explanation. I suggest that, despite the fact that Lipton himself does not defend this wholesale view, and argues that this account ultimately fails to defend realism effectively, it nonetheless offers a good representation of *how* the no miracles argument could be interpreted as a wholesale Inference to the Best Explanation.

Earlier, we said that our description of inference and explanation, and the relationship between the two, would be done under the *presupposition* of inferential and explanatory realism. Thus far, it was granted that the aim of inference is truth, and that our actual inferential practices are truth-tropic in the sense that to count as an actual explanation, it needs to be (at least approximately) true. The aim now is to examine this presupposition – to assess whether Inference to the Best Explanation can be employed to account for the no miracles argument.

Lipton explains that scientists employ Inference to the Best Explanation to infer that a claim is (at least approximately) true, whether it concerns something observed, unobserved but observable, or something unobservable. Granting this descriptive assertion about what scientists do, the question regarding the *justification* of such an inference is now what is under scrutiny. Are these inferences to the best explanation, that scientists make, really truth-tropic – do they reliably take scientists towards the truth? Lipton takes a 'scientific realist' to be someone who grants both, the descriptive fact that, scientists employ Inference to the Best explanation for their inductive practices, as well as the fact that these practices are truth-tropic. Under this conception, the no miracles argument is an argument which itself employs Inference to the Best Explanation, and says that we ought to infer that scientific theories which enjoy predictive success are (approximately) true, since the fact that they are true would be the best explanation of their success. But, since we ought to infer, or so this account suggests, that predictively successful theories are true, then it follows that the inferences scientists make, *which constitute these theories*, must be truth-tropic. In short, the no miracles argument is therefore an argument from predictive success to truth, and from truth to truth-tropism; it gives us reason to believe that the inferences scientists make are truth-tropic.

Proponents of the no miracles argument for scientific realism hold that Inference to the Best Explanation is an acceptable type of inductive inference, and therefore, endorse the inferences, which are of precisely the same form, that scientists make. But, as Lipton tells us, the no miracles argument itself is not a scientific inference, but rather a philosophical argument, which takes the same form of the inferences scientists make, with the conclusion that these scientific inferences are truth-tropic. This argument has the following form:

(1) Scientists employ Inference to the Best Explanation to infer the truth of a particular claim.

(2) Some people endorse Inference to the Best Explanation as an acceptable method of inferential reasoning.

(3) The no miracles argument itself employs Inference to the Best Explanation to infer that the inferences scientists make (their claims in general) are truth-tropic.

(4) So the no miracles argument, some will agree, employs an accepted method of inferential reasoning.

(5) Therefore, the no miracles argument, for those who already accept Inference to the Best explanation, effectively justifies the truth of scientists' claims in general.

To elaborate on the above argument; authors such as Boyd and Psillos – those who advance a wholesale account of the no miracles argument, construed as an Inference to the Best Explanation – embrace a distinctively naturalistic *epistemology*, employing only those methods scientists themselves use. Accordingly, this epistemological account makes no appeal to any philosophical method, which takes priority over scientific method or which can be utilised to evaluate first-order scientific problems. Psillos articulates what he and those before him believe is the type of explanatory reasoning which scientists employ: "Suppose that a background theory T asserts that method M employs causal processes $C_1, \ldots, C_n$ which, according to T, bring about X. Suppose, also, that we follow T and other established auxiliary theories to shield the experimental set-up from factors which, if present, would interfere with some or all of the causal processes $C_1, \ldots, C_n$, thereby preventing the occurrence of effect X. Suppose finally, that one follows M and X obtains. What else can better explain the fact that the expected (or predicted) effect X was brought about than the theory T – which asserted the causal connections between $C_1, \ldots, C_n$ and X – has got these causal connections right, or nearly right? If this reasoning to the best explanation is cogent, then it is reasonable to accept T as approximately true, at least in those respects relevant to the theory-led prediction of X. To be more precise, more is needed for the acceptance of T as relevantly approximately true. For instance, *T* is to be contrasted with available alternative hypotheses, and should emerge as *the* best explanation. *T* should also offer a 'good enough' explanation in its own right, e.g. an explanation which can adequately account for all salient features of the experimental facts. But such considerations are part and parcel of these more concrete applications of explanatory reasoning in science. And although we may not always be in position to choose a hypothesis as clearly the best explanation, that does not entail that we never are."[53]

According to a wholesale perspective Lipton believes the no miracles argument is intended to convey, the no miracles argument is distinctively a *philosophical* one, which aims to defend the reliability of scientific methodology in generating (approximately) true theories and hypotheses. However, its force, argues Psillos, relies on the concrete type of explanatory, naturalistic, scientific reasoning captured above. The import, for this account, of this concrete type of explanatory reasoning in science, for the no miracles argument, is obvious: successful cases of this reasoning constitute the foundation for the general abductive inference employed by the no miracles argument. Accordingly, the no miracles argument is an *instance* of the method used by scientists, however, its aim has wider appeal: to support the claim that Inference to the Best Explanation is reliable. To be precise, here, the no miracles argument consists of two layers of inference: there is the layer of first-order instances of reasoning that scientists actually use, namely Inference to the Best Explanation. These first-order instances are employed by the argument to defend the view that *particular* theories are (approximately) true. Then, there is the second-order reasoning – also Inference to the Best Explanation – that relies on those first-order instances in order to defend the view that science is truth-tropic – i.e. it can offer us true theories or claims. Thus, the no miracles argument is a philosophical argument - employing Inference to the Best Explanation - which says that the best explanation for the predictive success commensurate with scientific methodology (which itself involves Inference to the Best Explanation) is that the theories which are implicated in this practice are (approximately) true. Otherwise said, the no miracles argument uses Inference to the Best Explanation to defend the reliability of the abductive reasoning – precisely Inference to the Best Explanation – which scientists employ.

[53] Psillos, S. (1999), p. 79.

Challenges to the wholesale interpretations

Lipton discusses some standard objections to the wholesale no miracles argument, interpreted as an Inference to the Best Explanation. The most common charge against such inductive justification of induction is that it is viciously circular and begs the question against opponents of realism:[54] if the purpose of the argument is to defend the reliability of Inference to the Best Explanation, but the argument *presupposes* that Inference to the Best Explanation is a reliable inferential method, then it follows that this reading of the no miracles argument assumes true precisely what it needs to show. Fine summarises this line of thought, when he suggests that the realist is "not free to assume the validity of a principle whose validity is itself under debate."[55] Moreover, since critics of realism reject (or are not convinced by) the reliability of Inference to the Best Explanation, and the no miracles argument presupposes it, it seems as though the argument "preaches to the choir" and cannot convince someone who is not already a realist.

Who, after all, is supposed to be convinced by the no miracles argument? Those who don't accept Inference to the Best Explanation will not accept the no miracles argument, either because they plainly do not accept any form of inductive inferences, or because, whilst they accept some forms of induction, they don't accept Inference to the Best Explanation. Moreover, even if someone does accept Inference to the Best Explanation, it is not necessarily the case that they will accept the no miracles argument; ". . . one might claim that the best potential explanation is a guide to inference, yet that what we ought to infer is not that the explanation is true but only, for example, that its observable consequences are true. Such a person would not be moved by the miracle argument, at least not in the direction its proponents intended."[56]

Lipton suggests, therefore, that the only individuals that will be moved by the no miracles argument are those who, both, accept Inference to the Best Explanation *and* accept that this form of inference is truth-tropic. And since this is precisely, for Lipton, what constitutes a scientific realist, it follows that the no miracles argument will be compelling only for a scientific realist. But if the purpose of the no miracles argument was to *justify* the position of scientific realism, then, it seems, it has done no work at all. "In short, the miracle argument is an attempt to show that Inference to the Best

[54] Laudan, Larry. (1981).

[55] Fine, A. (1986), 'Unnatural Attitudes: Realist and Instrumental Attachments to Science', *Mind* 95: 149-179, p. 161.

[56] Lipton, P. (2004), p. 186.

Explanation is truth-tropic by presupposing that Inference to the Best Explanation is truth-tropic, so it begs the question."[57]

The circularity problem for the no miracles argument, Lipton claims, is similar to any attempt to give an inductive justification of induction itself. But he says that he is not convinced that the inductive justification of induction *nor* the no miracles argument, cannot overcome the circularity problem; this, in part by virtue of the fact that despite the intuitive appeal of the circularity objection, circularity is difficult to characterise.

Lipton suggests that there are additional, positive reasons that justify using inductive arguments to support inductive inferences. The first reason is based on the 'charting method'. Lipton conveys his experience as a clerk on the London Metal Exchange, taking an interest in the methods dealers established their views as regards the future movements of the markets in the different metals they traded: "Chartists construct a graph of the prices of the metal over the previous few months and use various strange rules for projecting the curve into the future. . . The dealers provided no explanation for the success of this method, but many of them put stock in it."[58] This is clearly an example of an inductive method the reliability of which could be evaluated using an inductive argument regarding past performance. Whilst the method is not convincing, Lipton suggests that successful past performance would warrant affording it some merit. The second reason relates to 'persistence forecasting'. Lipton relays a conversation he had with a meteorologist, about predicting the weather. The meteorologist told him that "there had been studies comparing the reliability of various predictive methods and that the technique of persistence forecasting still compared favourably with other methods. Persistence forecasting, it turns out, is the technique of predicting that the weather will be the same tomorrow as it was today."[59] Likewise, whether or not the meteorologist was sincere about the effectiveness of persistence forecasting, its reliability could be discerned by the extent to which it had worked in the past. Lipton's suggestion here, is that we can genuinely appraise inductive methods *by arguing inductively*. Therefore, having worked in the past would provide a reason for believing that persistence forecasting or charting would be reliable in the future as well, because if they were unlikely to be reliable in the future, it would be unlikely that they would have been reliable in the past.

[57] Lipton, P. (2004), p. 186.
[58] Lipton, P. (2004), p. 188.
[59] Lipton, P. (2004), p. 188.

However, if charting and persistence forecasting had an impressive history of success, it would still not convince an inductive skeptic – it would only confirm the beliefs of those who already endorse induction. As Lipton suggests, it appears that circularity is audience relative; whilst the argument may be feasible for some specific audience, it will simply beg the question for another. The obvious question is whether methods like the aforementioned can serve any purpose if they are only relevant for those who don't need to be convinced? Lipton suggests that perhaps it may be useful for convincing those who may not initially trust these methods, yet by their own rational commitments, ought to endorse the arguments from their track record. But for the case of the inductive justification of *induction*, it is difficult to imagine how the argument could add any value to someone who does not already embrace its conclusion – this person, after all, *is* an inductive skeptic.

Lipton aims to show that, even amongst individuals who *are agreement* about a particular claim, an argument can still be useful when it can be utilised to settle a disagreement about the *extent to which* the claim is true. Lipton suggests that the track record of induction could be used to settle dispute, amongst proponents of induction, over the *degree* of reliability of induction, giving these people an additional reason for holding their belief, or a legitimate reason to re-assess it. "Consider someone who is much too optimistic about his inductive powers, supposing that they will almost always lead him to true predictions. Sober reflection on his past performance ought to convince him to revise his views. If he admitted that he had not done well in the past, yet claimed for no special reason that he will be virtually infallible in the future, he would be inductively incoherent. Similarly, someone who is excessively modest about his inductive powers, though he gives them some credit, ought to improve his assessment when he is shown how successful he has been in the past."[60] If the intention is to convince the inductive skeptic that his view is incorrect, then the inductive justification of induction serves no purpose; but this does not imply that it may not serve any purpose at all – it may move an inductive endorsee who is not yet convinced of the degree of reliability of the conclusion of the argument. This result, claims Lipton, shows us that, whilst Hume's argument against the inductive justification of induction is warranted, the fact that his conclusion is that the argument is circular does not preclude the fact that it nonetheless has appeal.

The examination of the legitimacy of the inductive justification of induction is important to discern for whom the no miracles argument is cogent. So given Lipton's conclusion as regards salvaging the inductive justification of induction for at least *someone*, he aims to show that the case is analogous to the no miracles argument – that it is at least not circular for someone who accepts Inference to

[60] Lipton, P. (2004), p. 190.

the Best Explanation, insofar it can be used to settle disputes. As Lipton has shown us, the inductive justification of induction can be defended when it can be used to settle disputes over degrees of reliability. Lipton suggests the no miracles argument could be used to this end as well. According to Lipton's account, the no miracles argument is just one type of argument which employs the inductive justification of induction. Recall that we are dealing with the standard objection to the no miracles argument, interpreted as an Inference to the Best Explanation, as construed by Lipton. The objection for this account of the no miracles argument was one concerning the problem of circularity. If Lipton's result is correct - that the inductive justification may still have probative force for those involved in settling a dispute regarding the degree of reliability – then the no miracles argument may be able to circumvent the putative problem of circularity. In what sense, we ought to ask, can some sense of dispute over the degree of reliability be captured in the case of the no miracles argument? Lipton tells us that, since even successful theories can make false predictions, it follows therefore that the theory, in this case, cannot literally be true. So for the case of the no miracles argument, the dispute can be considered one over degree of approximation to the truth. Scientific realists may diverge on their beliefs about how effectively or quickly scientists move towards the truth, and the no miracles argument could be employed to deal with this type of disagreement.

Whilst Lipton's response to the problem of circularity may perhaps be feasible, it comes at the cost of failing to remain faithful to the no miracles intuition. There is nothing in the 'textbook' formulations of the no miracles argument, which suggests that its purpose is to settle disputes about the degree of approximation to truth amongst those who are already committed to realism. The no miracles argument, instead, is quite simply one which endorses the belief that success begets truth, *irrespective* of one's original metaphysical commitments. The argument is quite consistently employed to defend realism, against anti-realist contentions, so that the dispute that may be settled is one over the truth of science, *not* over the degree of approximation to truth amongst those who already believe that science is truth-tropic. Inasmuch as we insist on interpreting the no miracles argument as an inductive justification of induction – where both methods of induction are inferences to the best explanation – *and* wish to remain faithful to the no miracles intuition, then the problem of circularity persists. [61]

[61] Instead of *addressing* the problem of circularity, Boyd's wholesale account, for example, *embraces* it as part and parcel of his epistemological holism. Whilst Boyd would suggest that circularity is not a problem for his account, those who do not share the same epistemological views would not agree. In fact, it has been argued that Boyd's account begs the question; Boyd places sufficiently stringent conditions on 'being an explanation' that it turns out that it is only *current science* that can offer explanations for what we observe. One could argue that this is just the type of circularity discussed above: just as presupposing the reliability of Inference to

I have argued that Inference to the best Explanation not only fails to capture the essential ingredients of the no miracles intuition, but it cannot afford us an effective wholesale interpretation I have suggested is incumbent on the no miracles argument. Either (i) we accept problems of circularity, in which case the account can be said to be fallacious or question-begging; or (ii) we can deny circularity, in which case the argument may be valid, but fails to serve the purpose the no miracles argument is *intended* to serve – to convince those who are not *already* realists.

Whilst Inference to the Best Explanation fails to capture the no miracles argument effectively, accounts like Boyd's and Psillos's, I suggest, do have some merit. Their appealing aspects have to do with an appeal to what scientists *actually do*. Since it is the success of science in general which evokes the no miracles intuition, it seems favourable to hang on to those elements of the examined accounts which appeal to the very practices in science which are integrally bound up with its impressive success. By sticking to this rationale, perhaps we need not appeal to some *other*, additional method of inference – one which requires independent justification - in order to argue for the (approximate) truth of science in general. Instead, the hope is that we can restrict ourselves to those methods *actually* employed by scientists, whilst at the same time, ensuring that the interpretation developed infuses scientific realism with something *more* than just the normal scientific practices actually employed. Thus, if we wish to find a suitable formalization of the no miracles argument, a good place to begin would be with the activities scientists employ which *distinguish* plausible hypotheses from implausible ones. This is the impetus for my formalization I develop in the final chapter. However, before I develop my formalization along these lines, it will be useful, in light of the thus-far-failed-attempts to capture the no miracles argument, to consider an argument which suggests that the no miracles argument *cannot* be captured – that it is nothing more than an intuition. This idea is examined in the following chapter.

the Best Explanation, in order to justify the reliability of Inference to the Best Explanation, is circular, so too is presupposing the approximate truth of current science, in order to justify that current science is approximately true. The point is that, wholesale accounts of the no miracles argument, viewed as an Inference to the Best Explanation, do not seem able, in general, to circumvent problems of circularity.

3. The No Miracles Argument as *the 'Default' Position*

You are lost in the woods and you come across a topographical map. . . After some exploring, you find that everything the map indicates you ought to see - rivers, lakes, roads, peaks and so on – you do in fact see along your journey. . . You decide to embrace what the map tells you about the terrain in general, because you feel it would be a miracle if the map were, for example, a map of another region, yet happened to be correct about everything you have seen here, up until now. You simply have an intuition that the map must be accurate; you trust its guidance - you feel it would be a miracle if the map were, for example, a map of another region, yet happened to be correct about everything you have seen here, up until now. There is nothing more that can be said – no additional, positive argument justifying why you ought to accept the truth of the matter that the map is really a map of the region within which you find yourself being lost. This intuition alone, you believe, is sufficient to establish, as the 'default' position, the correctness of the map.

In his paper,[62] Worrall presents his understanding of the no miracles argument – one he refers to as *the 'default' position*. This position holds that the no miracles argument is just an intuition. I explain first why Worrall believes the no miracles argument ought to be construed as a *retail* argument for realism. This argument is, in part, a response to Magnus and Callender's paper examined in the first chapter. Worrall claims that arguments suggesting that the no miracles argument is fallacious, are misguided, and that when construed properly, it no longer falls victim to the supposed fallacy.[63] I argue that Worrall's 'default' position suggests *more* than just the claim that the no miracles argument amounts only to an intuition. In doing so, I attempt to formalise this supposed, broader argument in order to assess, first; whether the default position manages to capture the no miracles intuition, and second; whether this account of the no miracles argument can plausibly defend scientific realism. I argue, both, that the default position is not promising in capturing the requisite intuition, and that it fails to defend realism effectively.

[62] Worrall, J. (2005).
[63] Worrall shows as well that arguments suggesting that the pessimistic meta-induction is fallacious, are also misguided.

Revisiting the scope of the no miracles argument

Worrall asks what the scope of the no miracles argument should be. He suggests it should not speak for *all* scientific theories; "It seems that there is instead a wide variety of science*s*, not all of them 'surprisingly successful', certainly not in any sense that elicits the 'no miracles intuition'."[64] For example, theories from sociology, some from psychology and parts of dietetics, are not mature. He concedes that Magnus and Callender do suggest only that the success of *mature* sciences is best explained by the fact that all (or perhaps most) mature scientific theories are true. But he (and others, including Laudan) suggests that 'maturity' is too vague. Realists often claim that success is to be identified with *predictive* success, and mature theories are those, simply, that *have* achieved predictive success, by way of predicting 'new' phenomena. But if this is the case, then the inference being made, anyway, does not seem to be one from some loosely defined property, the 'success of science', as some suggest wholesale realists would have it.[65]

Also, Worrall suggests that attempts to articulate the no miracles argument as one about 'most' successful theories, are also misguided. If even a small number of theories achieved predictive success but turned out now to be false, then the realist could no longer appeal to the predictive success of most of our sciences, as a *miracle*, since miracles, surely, ought to be rarer than the actual, observed frequency of successful yet false theories.[66]

He suggests, therefore, that it is misguided to consider any wholesale version of the no miracles argument. It is particular scientific theories which elicit the no miracles intuition. "Belief in 'atoms' really translates into beliefs that various theories that use the term are at least 'on the right lines' because those theories have had striking empirical successes (e.g. in Perrin's work on Brownian motion), to an extent that seems entirely implausible if they are not on the right lines."[67] So Worrall

[64] Worrall, J. (2005), p. 8.
[65] It is not my intention, here, to challenge Worrall's reasons for believing that the no miracles argument ought to be interpreted as a retail argument – I have already offered arguments in defense of a wholesale conception. However, I will grant this concession for argument sake, in part so that we can suss out any plausibility the 'default' position *may* nonetheless have, and in part because I believe the 'default' position fails for other reasons. I claim that the formalization I present in the following chapter *does* answer to the problems Worrall claims a wholesale view is susceptible to, since my formalization makes no appeal to a notion of the 'success of science', as I will make clear when we get there.
[66] Recall that this is part of the reasoning behind Laudan's pessimistic induction argument, to illustrate that a significant number of theories that achieved predictive success (and were considered true), later turned out to be false.
[67] Worrall, J. (2005), p. 10, footnote 5: Worrall admits, that as was pointed out to him by Callender, their view is not incoherent insofar as the no miracles argument is seen exclusively as a *wholesale* argument for realism, involving 'proper' probabilities – in this case, despite the apparent appeal to the no miracles intuition, their

defends a retail version of the no miracles argument and suggests that any wholesale argument simply amounts to the union of a number of retail arguments for realism about *particular* theories that have achieved maturity by displaying predictive success. Such a wholesale conception would be trivial – an empty truism; *no one* who accepts a number of individual retail arguments for realism would deny a wholesale version which *merely* consisted of the conjunction of such retail arguments. This type of wholesale argument would offer no illuminating elements over and above what the retail arguments offer us.

Although both Worrall and, Magnus and Callender, champion retail arguments for realism, their respective justifications diverge; Worrall argues for retail arguments for the reasons mentioned above, whereas Magnus and Callender do so, as we discussed previously, because they argue that the wholesale version is subject to the base rate fallacy. Moreover, Worrall contends, as discussed already, that Magnus and Callender's position is incoherent because their retail argument makes an implicit appeal to the no miracles intuition, and therefore, by their own rationale, is *also* subject to this fallacy of reasoning.

Worrall goes on to argue that scientific realism is not a scientific view because it is not testable. He suggests that an explanation is a scientific one provided it is subject to independent testability. We cannot test whether the prediction, that the next scientific theory we are confronted with is true. There is no way to test even its approximate truth. "The whole point about scientific realism, as Magnus and Callender. . . fail to grasp firmly, is that it attempts to defend a link between the effective, decidable notion of success and the essentially undecidable, if you like transcendental, notion of truth (or approximate truth)."[68] Worrall, in comparison, suggests that ascribing 'truth' to a particular scientific theory or claim is simply making a *judgment* that it is true (or false when it is in conflict with ones which are currently accepted) – something we shall soon see, is for him, merely an intuition that it is true, given its predictive success.

Hence, Worrall's retail interpretation of the no miracles argument has as its conclusion that a *particular* theory is (at least approximately) true. He adds that the argument will then not contain some sweeping notion of 'success' commensurate with wholesale arguments, but instead, the genuine *predictive* success of the theory under consideration: the theory must predict some general

retail arguments do not appeal to 'the' no miracles argument (intuition) proper. It is only incoherent inasmuch as the no miracles intuition underpins both the wholesale argument that they denounce as well as the retail arguments that they accept, which, they claim, it doesn't.

[68] Worrall, J. (2005), p. 11.

type of empirical result – one that actually manifests itself when the experiment is carried out. By 'predictive success', Worrall means the ability to predict a phenomenon *provided* this phenomenon was not 'used in the construction' of the theory. Here, as others too have done, Worrall alludes to Psillos's notion of *ad hoc* construction of a theory, which, as a reminder, constitutes the idea of having built the phenomenon into the theory, thus making it trivially the case that it would "predict" it.

Accordingly for Worrall, when a theory T has exhibited a great predictive success, it would be miraculous if T were able to be so precise about some phenomenon yet were not (at least approximately) true. And since we should not accept that a miracle has occurred whilst there exists a better explanation – that T predicted the phenomenon successfully precisely because T is at least approximately true – it follows that ought to infer that T is approximately correct.

Worrall warns of the logical possibility that, as we are shown by Jeffreys-style transformations, a theory could demonstrate "surprising" predictive success, yet, at the same time, actually be false. These transformations tell us that we can construct infinitely many rivals to a theory T, which are inconsistent with T, yet all entail the 'novel' phenomenon we applauded T for predicting: When T is a mathematical theory relating X to Y by the function $y = f(x)$, and suppose, according to the theory, $f(x_0) = y_0$ and that when $X = x_0$ we actually witness this supposedly surprising result y_0, precisely as the theory predicts. However, if we consider T' the theory the function of which is $y = f(x) + (x - x_0)g(x)$, for any arbitrary $g(x)$, then it follows that at $X = x_0$, T' will simply be $f(x)$, thus predicting precisely the same "surprising" result as T. Suppose now that T' *were* (approximately) true, then far from latching onto reality, T could be outright false *except* when $X = x_0$. But Worrall, along with others, suggest that these theories involve *ad hoc* accommodations – precisely the type of "built-in" predictive capacity that fails to satisfy the criterion of 'novel' prediction to warrant *genuine* success. Worrall suggests, therefore, that such efforts to undermine the success-to-truth connection are not legitimate. Instead of showing that we *cannot* infer truth from success, the most Jeffreys-style transformations may tell us is that just because a theory is successful, does not logically *necessitate* that it is true. Besides, I doubt any realists would deny this, for it is not only compatible with realism, but it is probably widely accepted amongst realists.

Persisting intuitions

Worrall suggests that the no miracles argument ought not to be interpreted probabilistically, and that just because it cannot be formalized as such, does not imply that the no miracles argument loses its probative force. Worrall suggests that the psychological research Magnus and Callender appeal to – that even educated people have a tendency to commit this fallacy of reasoning – is insufficient to dismiss the no miracles argument. What this research of Tversky and Kahneman actually demonstrates, suggests Worrall, is a matter of dispute.[69] Furthermore, he suggests that many people are immune from "suffering" the base-rate fallacy and, that despite conscious consideration of the fallacy as exposed by Bayesian considerations, the pull of the no miracles intuition *still* persists. Accordingly, Worrall suggests that instead of giving up on the intuitions, we should accept instead, as we saw in Chapter 1, that the failure exists in the attempts to capture the intuition using a probabilistic argument.

Worrall considers the possibility for an alternative, more precise formulation of the no miracles argument; perhaps, as he suggests, as an Inference to the Best Explanation: "The idea, after all, is that there seem to be only two *explanations* for why some theory T has scored some striking predictive success: (a) that T itself (including those parts of it talking about 'deep structure') is at least approximately correct and (b) that T, although perhaps radically off-beam at the noumenal level, nonetheless just happens to get this particular prediction correct by chance. (a) is much the better explanation (if indeed (b) counts as any sort of explanation at all). Hence we should accept (a)."[70] If we presuppose that Inference to the Best Explanation is the best formalization of the no miracles argument, then it seems as though the probabilistic arguments considered by Magnus and Callender cannot capture such an inference – the probabilistic account would have it that T simply *entails* some datum e, but there is no requirement that it also *explains* it. Unlike the Jeffreys-style constructions which merely entail the data from which they were constructed, Inference to the Best Explanation, as we have discussed in Chapter 2, necessitates that T possess other epistemic virtues over and above mere entailment. However, the importance of "explanation" as a primary epistemic virtue seems unfounded. Additionally, Worrall makes it clear that a Bayesian interpretation, whether on its own terms or as an Inference to the Best Explanation using a probabilistic account, fails to do justice to the no miracles intuition. Likewise, as we have already seen, Inference to the Best Explanation, as an account by itself, fails to offer any significant improvement to the core intuition we began with.

[69] Worrall, J. (2005), p. 18, footnote 9.
[70] Worrall, J. (2005), p. 23.

So Worrall insists that, in light of the fact that neither the Bayesian approach, nor Inference to the Best Explanation, can offer a feasible account, we are left with essentially *just* the intuitions. Worrall has argued that Magnus and Callender are misguided in supposing that the intuitions hold no probative force. Bringing these arguments together, Worrall claims that we should not be concerned that we are left with "only" the intuitions; he suggests that perhaps we should even expect that there cannot be an argument *proving* that a theory is true just because of its striking predictive success. "Proofs and objective probabilities are not what 'the NMA' is about. The impact of predictive success, together with the notion of 'approximate truth',[71] is intuitive – ineliminably so. It is of course possible that our current theories are radically false despite their predictive success, but it seems so downright implausible. Implausible enough, I suggest, to set realism as the default position. It is surely not surprising, at least on reflection, that the implausibility here cannot be captured by any sensible analysis in terms of proofs or of objective probabilities. (Some) realists may wish for something stronger from 'the' NMA, but nothing stronger is defensible."[72]

Yet, Worrall concludes, we do not doubt the seemingly reasonable view that what our mature sciences say about what we can expect to observe generally are in fact correct. Likewise, as regards theories which work with unobservable entities, we do not doubt what those theories prescribe, despite not having proof, nor an argument for high objective probability. The default position here, Worrall suggests, is also that these theories are (at least approximately) true. This intuition, as it stands, is all we have, and as such, is the impetus for setting realism as the default position: there is nothing stronger one can say about the no miracles argument.

Assessing the 'default' position

Worrall's 'default' position is not easy to criticize. After all, insofar as the two most prominent interpretations of the no miracles argument fail to capture the core intuition and, besides, are either fallacious or unacceptably imprecise, it seems as though *the most we can say* about the argument is that it amounts simply to the intuition itself. However, Worrall's conclusion is by no means

[71] It seems key to Worrall's understanding of what is distinctive about the no miracles argument, that the inference is only to *approximate* truth. This understanding concurs with my explanation of approximate truth in my introduction. Moreover, it is interesting to note that, it seems as though this looser notion of truth cannot be captured by the Bayesian account, offering more justification for why the probabilistic calculus fails to capture the necessary intuitions.

[72] Worrall, J. (2005), p. 26.

perspicuous; for starters, it is not obvious whether the default position is *just* the claim that all there is to the no miracles argument, is an intuition, or whether his default position is *also* a normative claim, over and above the former, which argues that we *ought to* adopt realism in no small part because *all we have* is this intuition. In other words, we need to ask whether Worrall is simply presenting a descriptive claim about what the no miracles argument *is*, or whether he is advancing a normative claim about what metaphysical position we *ought to* adopt given, in part, what we can say *about* the no miracles argument. It is important to distinguish between these two possibilities, because the answer will determine the relative feasibility of his account.

Let us begin with the possibility that the default position is just the descriptive claim – that all there is to the no miracles argument is the intuition itself. According to this interpretation, it is simply a matter of fact that when appeal is made to the no miracles argument, all that is going on is that when we consider that the predictive success of a theory may just be a fluke, in much the same way that the map may happen to have been correct in *just* those features you randomly came across, we believe, only intuitively, that such surprising happenings can't merely be attributed to coincidence. The intuition, more specifically then, is that it certainly would be a perplexing coincidence if some surprising phenomena are explained by a scientific theory, yet the theory failed to latch on to the world in some sense. Denying that this scientific theory is at least approximately true would make its success a miracle, but there is no positive argument, or proof, which can capture this intuition.

Fair enough, we can't argue with the fact that most of us do feel the probative force of this intuition – besides, I do not wish to challenge the claim that we do in fact carry this intuition. Instead, I would like to ask what it could *mean*, in light of having only the intuition, to "set realism as the default position," when all that is being presented is a descriptive claim? Granted, if the no miracles argument cannot possibly be formalized as a proof, then Worrall may be correct in saying that all there is to the no miracles argument is the intuition. But it seems as though Worrall is saying something more (whether he intends to do so or not) than the fact that this intuition merely *exists* and that nothing more can be said. Instead of the solitary claim that the content of the intuition is *that of* realism of particular theories; it seems as though Worrall is inferring from the fact that the content of the intuition is that of realism, together with the fact that other formalizations fail, that the position we *ought to adopt* is that which is in alignment with all that we have – the intuition.

It is no coincidence that Worrall employs language like ". . . Implausible enough, I suggest, to **set realism** as the default position."[73] It does not appear as though Worrall is suggesting *only* that the intuition is all we have; to "set" *realism* as the default position, I suggest, is too emphatic to express just the descriptive claim that we have nothing more than the intuition itself. Rather, the instructional, or even inferential force of the phrase, seems to hint of a surreptitious appeal to some form of *argument* which aims to *convince* the reader that realism ought to be adopted. Although this may not be Worrall's intention, it certainly *seems* to be the case. To be precise, I suggest that Worrall is presenting a positive, normative argument in favour of scientific realism (retail in nature), which, I believe, could be formalised as follows:

(a) **The intuition:** A lot of people are inclined to hold the underlying, no miracles intuition; "The impact of predictive success, together with the notion of 'approximate truth', is intuitive – ineliminably so. . ."[74]

(b) **Attempts at formalization:** formalizations of the no miracles argument, according to a Bayesian account and according to Inference to the Best Explanation, have failed, both in terms of being able to capture the core intuition, and in the sense of being able to offer a valid argument for realism.

(c) **Implicit assumption:** there are no alternative formalizations of the argument which may be effective.

(d) **Sub-conclusion:** Therefore, all we are left with is the intuition alone – the no miracles argument is *just* an intuition.

(e) **Implicit assumption:** if all we have to support a position is an intuition that it is correct, and a lot of people are inclined to hold this intuition, then we ought to embrace the content of this intuition.

(f) **Conclusion:** Therefore, we ought to embrace the content of this intuition – that scientific realism for particular theories, is correct.

As mentioned, (a) is assumed true – the no miracles argument is the most prominent argument in defense of scientific realism, and common amongst all interpretations of this argument is the core intuition captured by (a). It is an accepted fact that many people hold this intuition.

[73] Worrall, J. (2005), p. 26.
[74] Worrall, J. (2005), p. 26.

In Chapters 1 and 2, I have argued that, by themselves, both the Bayesian approach as well as Inference to the Best Explanation, fail to capture the no miracles argument effectively. (b), therefore, it seems, is true.

Just because no formalization of the no miracles argument has, *as of yet*, been propounded effectively, does not imply that it there *cannot* be an effective formalization. Any move from the fact that a formalization has not yet been effected, to the fact that it *cannot be* effected, is of course not valid. Perhaps, at most this suggests that it is less likely than not that someone will come up with an effective formalization, but even this claim is by no means obvious and would require independent justification. Besides, Chapter 4 in this exposition is dedicated to an alternative formalization of the no miracles argument; one I argue *does* capture the no miracles intuition effectively and defends (wholesale) realism in a feasible way. For these reasons, (c), I claim, is therefore apparently false.

We said that (a) and (b) seem to be true. Suppose for argument sake, we grant Worrall the truth of (c); then, given the validity of the sub-argument (a) – (d), the truth of these premises guarantees the truth of sub-conclusion (d). And this conclusion – that all there is to the no miracles argument is the intuition – is perhaps not too difficult to stomach. However, we should bear in mind that this conclusion relies on the truth of (c), and if my claim, that (c) is false, is in fact correct – because there still could be some undiscovered formalization, or in light of the possible success of an alternative account I shall advance - then (d) will turn out to be false.

However, most importantly, even if we grant Worrall the truth of (d), I have argued that he is proposing something more than just this, and so this broader claim, may *still* fail. The implicit assumption I have argued is evident in Worrall's writing - captured by (e) - is just far too contentious. Inasmuch as Worrall's exposition turns out, as I claim, to be a normative argument in defense of realism, *based on* his interpretation of the no miracles argument, he would owe us a convincing explanation of why we ought to believe that our intuition under consideration is actually reliable; just because it is something a lot of people are inclined to believe, does not give us an obvious, or perhaps even *any* reason to believe it.[75] And I am not convinced that an appeal to the general reliability of our intuitions would be warranted; besides, any such argument would simply beg the question: *in virtue of the past successes our intuitions have had, we ought to infer that they are (at*

[75] Worrall himself has said similar things in discussions about naturalism. See Worrall, J. (1999), 'Two Cheers for Naturalised Philosophy of Science or: Why Naturalised Philosophy of Science is Not the Cat's Whiskers', *Science and Education*, 8: 339-361.

least approximately) true. This is precisely the type of inference we are trying to formalise from the get go. Substituting the task of justifying the inference from predictive success to truth of a scientific theory, to the task of justifying the inference from the past successes of our intuitions to the truth thereof, gets us nowhere. Perhaps, instead, one may argue that Worrall could defend the reliability of this intuition *in particular*. But in order to do this, we would need to be able to show that realism about particular scientific theories is true. This move would be absurd because recall that the reason why something like the no miracles argument is advanced at all is *in order to* defend this sort of realism. If there was an independent way of showing that realism is true, we wouldn't need to appeal to arguments like the no miracles argument in the first place.

It appears as though the problems besetting (e) are insurmountable. And since the conclusion in (f) – one that supports realism as the 'default' position – depends on the truth of (e), this putative argument is no good. Of course, realism may be true, but certainly not (exclusively) for the reasons outlined in this argument.

We considered the possibility that Worrall's 'default' position is merely a descriptive claim about what in fact the no miracles argument is. To his credit, I concede that insofar as his claims regarding the failure to formalize the "argument" (both to date and in the future) are correct, then the no miracles argument *will* turn out to be just an intuition. However, notwithstanding the fact that this conclusion makes some arguably unwarranted assumptions, I maintain that Worrall's exposition regarding the default position has all the ingredients of an appeal to a broader, positive argument in favour of realism. However, when we examined the soundness of this putative argument, even when we granted for argument's sake the truth of some of the premises, the conclusion – that we ought to be realists – is inadequately supported by the premises. Most notably, the mere existence of the no miracles intuition as a reliable indicator of the truth of realism, is unfounded, and considerations to resolve this seem futile. Worrall's (sub-)claim, that the no miracles argument is just an intuition, is by no means convincing if it relies on the contingent fact that no one *as of yet* has been able to formalise the argument effectively. And if he is suggesting anything more – namely that we ought to adopt realism about particular theories because, in part, we have nothing else to go by *but* the intuition, then Worrall's interpretation of the no miracles argument, I conclude, does not afford us a plausible defense of scientific realism.

4. A New Formalization

Hitherto, we have been unsuccessful in presenting an adequate interpretation of the no miracles argument – one which both captures the no miracles intuition and which justifies the inference being made from the success of science to the (approximate) truth of science in general. I present what I refer to as "an unbiased summary of the 'textbook' formulations" of the no miracles argument, to illuminate a particular formalisation I argue, not only succeeds in capturing the no miracles intuition, but plausibly justifies the essential inference from success to (approximate) truth. I show that, whilst this formalization encapsulates the practices *actually employed* in the scientific enterprise, which are used to develop and corroborate theories – the activity of *frequentist testing* - it nonetheless it nonetheless does not *identify* the no miracles argument *as* a scientific test. Moreover, this formalization incorporates the *wholesale* perspective I have explained the no miracles argument ought to encompass, and has as its conclusion that the scientific enterprise gives rise to many more true theories than false ones, and therefore, that science in general is true.[76] Moreover, I will argue that this interpretation circumvents problems other accounts are susceptible to – namely, circularity and the base rate fallacy – and that challenges to this interpretation – both ones accounted for already as well as ones anticipated here – can, at least conceivably, be resolved.

An unbiased summary of the 'textbook' formulations

I would like to present, first, a general formulation of the no miracles intuition. This formulation, I suggest, can be viewed as an unbiased summary of the so-called 'textbook' formulations we have looked at thus far:[77]

(i) When our scientific theories predict in a non-*ad-hoc* manner particular observational data, yet these theories are not even approximately true, then the fact that the

[76] My use of 'science' here refers only to *science that employs frequentist testing*, and my use of science 'in general' refers to the fact that, as I will argue, science, *in the most part,* affords us (approximately) true theories. Henceforth, For the sake of brevity then, any use of the phrases 'science', and science 'in general', in my own account, is actually a reference for the aforementioned concepts.

[77] I have adapted the formulation presented in Howson, C. (2003). Where Howson formulates the intuition in the context of an *particular* test for theory T, I present the intuition in the context of scientific theories *in general*.

predictions are in alignment with what was observed in the data must be a mere coincidence – a mere chance occurrence.

(ii) The nature of the data observed makes it extremely unlikely to coincide, per chance, with the theories' predictions.

(iii) Chances as small as these are so exceptionally unlikely that, the hypotheses which say that these are *merely* chance happenings can reasonably be rejected with confidence, especially if there exists a better explanation for their account.

(iv) Therefore we can infer with confidence that these theories are at least approximately true, where the extent to which the chance in (iii) is small, is an indication of the degree of this confidence.[78]

I believe that the strategy statistician have established to infer, or draw conclusions from data, possess all the essential ingredients of the no miracles intuition above – thus far, something all the formalizations we have considered have left wanting. Before I begin to argue that what has come to be known as *frequentist testing* can maintain the essential form of the no miracles intuition, it will be useful to offer a brief summary and rationale of how such tests, of accepting and rejecting hypotheses, work in practice.[79]

An example of frequentist testing and the general procedure

Suppose a manufacturer insists that, on average, his batteries work for 100 hours until they run out.[80] If we know that the population standard deviation of the life of a battery σ is 12 hours, and we sample 50 batteries and find the sample mean $\bar{X}$ is 95.5 hours; what can we conclude about the truth of the manufacturer's claim?

Let us assume that the manufacturer's claim is true, and that the true mean μ is in fact 100 hours. We then ask: "What is the probability that $\bar{X}$ < 95.5, given that μ = 100?" We know that

[78] I explain below what this "degree of confidence" is, according to a frequentist framework.
[79] An example where frequentist testing is applied to an actual scientific investigation can be found in Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*, Chicago: The University of Chicago Press, Chapter 7 – 'The Experimental Basis from Which to Test Hypotheses: Brownian Motion', p. 214-250. It should be noted that whilst my example involves relatively low level phenomena, like testing light bulbs, Mayo, for example, offers discussion of how this sort of testing can be employed to discriminate even between very high-level theories, like the General Theory of Relativity.
[80] Adapted from Underhill, L. & Bradfield, D. (2007), *INTROSTAT*, Cape Town: Juta, pp. 202-205.

$\bar{X} \sim N\,(\mu, \sigma^2/\text{n})$

and

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N\,(0, 1)$$

For our example $Z = \dfrac{\bar{X} - 100}{12/\sqrt{50}} \sim N\,(0, 1)$.

Since $\bar{X}$ = 95.5, the associated *z*-value is therefore

$$z = \frac{\bar{X} - 100}{12/\sqrt{50}} = -\,2.65.$$

Thus

Pr[$\bar{X}$ < 95.5] = Pr[$Z < -\,2.65$] = 0.0040,   from the standard normal tables.

In other words, presupposing the manufacturers claim is true (that is, μ = 100), the probability of observing a sample mean which is 95.5 or less, is 0.004;[81] a very small chance.

At this point we have to decide: Either

(a) the manufacturer is correct, and a very unlikely occurrence has taken place – one which would be expected to occur 4 times in every 1000 samples tested, or
(b) the manufacturer's claim is not true, and the true population mean is less than 100, implying that a population mean of 95.5 or less a more likely event.

The statistician would go for option (b). He would argue that alternative (a) is so improbable that he can reject it with confidence, and he would conclude that the manufacturer is exaggerating.

This straight-forward example is intended to introduce the concept of *statistical inference* – how scientists, based on the work of statisticians, infer or draw conclusions from data. *Tests of*

[81] A p-value of 0.004 is the complement of the degree of confidence, or the confidence level of the test. It means if the trial is repeated a number of times, we can expect to have drawn the correct conclusion from the tests (1-0.004)$\times$ 100% of the time – that is, 99.6% of the time. When the test employs a *significance level* $\propto$,it means that we can expect to have drawn the correct conclusion (1-$\propto$)$\times$ 100% of the time – our 'degree of confidence'.

hypotheses, otherwise known as *frequentist testing* or *significance tests*, form the basis of statistical inference. When a claim about some quantifiable state of affairs in the world needs to be assessed, a frequentist test, as outlined below, can be employed using the step-by-step procedure. I will use the above problem to illustrate the steps, and thus, the rationale involved:

1. Introduce the *null hypothesis*. Here a statement is made, almost always, about the value of a population parameter. For our example, the claim is that the true mean μ is equal to 100. Usually, the null hypothesis involves an acceptance of the claim being conveyed (and we hope to reject it).

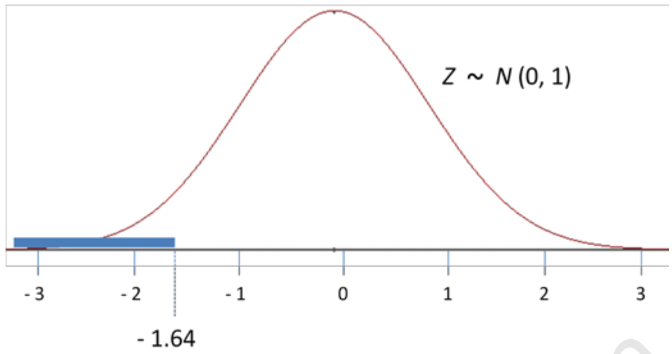
   The null hypothesis is formulated as follows:
   $H_0 : \mu = 100$

2. The *alternative hypothesis* H_1 is deployed. H_1 is effective when our test suggests that we ought to reject H_0. For our example, the alternative hypothesis is $H_1 : \mu < 100$. This we call a "one-sided alternative" and results in a "one-tailed significance test".

3. Choose a *significance level*. The significance level represents the probability of rejecting the null hypothesis when it is actually true, if the trial were to be repeated a large number of times. For example, using a 5% or 0.05 significance level means that on average, or, in the long run, one will mistakenly reject a true null hypothesis 5% of the time, i.e. once every twenty tests. If the repercussions of mistakenly rejecting the null hypothesis are serious, for whatever reason, a higher level – for example 1% - may be utilized. For illustrative purposes, a 5% significance level will suffice for our example.

4. Using the computational statistical tables, we identify the set of values which entail having to reject the null hypothesis. This is called the *rejection region*. In our example, since the test statistic will have a standard normal distribution, we make use of the normal tables (variables in other contexts may be distributed differently, which would necessitate the use of different tables, but the rationale for each step, and the procedure in general, remains the same). Since H_1 is one-sided and consists of a "<" sign, the rejection region is found on the lower (left hand) end of the standard normal distribution. What happens is that we reject H_0 if the sample mean $\bar{X}$ is sufficiently smaller than the hypothesized population

mean μ, i.e., if $\bar{X} - \mu$ is sufficiently negative. Since we used a significance level of 5% we therefore identify the lower 5% mark of the standard normal distribution; which is $-$ 1.64.



The rejection region is integrally related to: (a) the particular distribution of the test statistic; (b) the type of inequality associated with the alternative hypothesis, and (c) the magnitude of the significance level. The value we identify in the table, which demarcates the "edge" of the rejection region, is referred to as the *critical value* of the test statistic.

5. Calculating the *test statistic*. We know that $\bar{X} \sim N(\mu, \sigma^2/n)$. If H_0 is true then

   $\bar{X} \sim N(100, 12^2/50)$ and

   $$Z = \frac{\bar{X} - 100}{12/\sqrt{50}} \sim N(0, 1).$$

   For our example, $\bar{X}$ = 95.5 and the observed value of the test statistic z is

   $z = (95.5 - 100)/(12/\sqrt{50}) = -2.65.$

6. At this point we make our conclusion. We compare the observed value against the critical value: if the observed value is more extreme than the critical value – i.e. the observed value lies within the rejection region - we reject the null hypothesis H_0 and accept the alternative H_1. In the case of rejecting, the result is referred to as *statistically significant*. In our example, since our test statistic $-$ 2.65 is less than the critical value – 1.64 we reject H_0 at the 5% significance level.

Additional constraints – randomization

Before I articulate the meaning of randomization in the context of frequentist testing, it is important that we grasp the basic rationale of frequentist testing. The chance distribution associated with each individual null hypothesis is the probability function which maps the random variable being measured to the distinct probability formula, used to determine probabilities, in the long run. Each individual hypothesis will warrant its own particular chance distribution, depending on the characteristics of the random variable under consideration[82] So to say that there is a connection between the null hypothesis and a determinate chance distribution, means that the entire domain of possible results for a given experiment can be afforded associated probabilities – chances of occurring. The chance distribution is a function of the parameter about which the null hypothesis makes a claim, and thus a given null hypothesis warrants a particular chance distribution such that we are able to answer the question, "what is the probability of observing the result we did actually observe, assuming the null hypothesis is true." If we can answer this question, we can decide, based on how unlikely the result would be if the null hypothesis were true, whether there is significant evidence to suggest that we should reject the null hypothesis. After all, if for example it is so unlikely, say 1 in 10000, that we observe a result at least as extreme as the one we did observe, then the very fact that we nonetheless *did* observe it suggests that we can infer that the chance distribution used to derive this result is misguided as regards the value of the parameter of which it is a function. And since we have decided to reject this value – the one suggested under the null hypothesis – of the parameter, the alternative option we allowed for is to embrace the numerical property of this parameter dictated by the alternative hypothesis.

Fisher believed that an additional constraint is needed in order to justify inferring a determinate chance distribution from the null hypothesis.[83] This constraint involves a systematic *randomization* of the sample selected (in my example, of the batteries tested). Randomization, in this sense, is satisfied by repeating the experiment a number of times, whilst on each occasion, randomly (quite literally) selecting individual batteries to constitute the (new) sample. It is true that after any one of the randomized trials, there may be present alternative explanations of why the sample mean may

[82] For example, if we have a situation where we are given a period of time during which events occur at random and the average rate at which events occur is λ events per time period, then our random variable has the Poisson distribution with parameter λ and has probability mass function $p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$

[83] Fisher elucidates his solution to the problem of induction through his famous thought experiment based on a chance model dealing with a simple guessing game of a lady who correctly predicts whether a cup of tea had milk added before or after the cup contained the tea. See Fisher, R. A. (1925), *statistical Methods for Research* Workers, 12th edn., London: Oliver and Boyd.

have been lower than the hypothesized population mean of 100 hours, all of which are consistent with the truth of the null hypothesis that the true mean life of the batteries is 100 hours. In the context of the no miracles argument, this is analogous to saying that there exist alternative explanations of why the theory (seemed to have) made a successful prediction - explanations which are compatible with the fact that the theory is false: In our battery example, an unrepresentative sample of particularly short-lived batteries may have coincidently been chosen; the trial may have taken place during the coldest week of winter, thus affecting battery performance; the experimental apparatus may have been set up incorrectly; and so on.[84] However, whilst all this may be true, Fisher believed that this poses no problem when the experiment is properly randomized. In a properly randomized experiment, we are not concerned with a particular case, but instead, the *long run*.

According to Fisher, chance is a phenomenon that is instantiated only in the long run. This view is referred to as a *frequency theory* of chance.[85] Based on this theory, chances are features of ordinary events, such as 'a head occurs' for flipped coins, and quantified as the relative frequency (the proportion of successes out of the total number of repetitions) of these events when the trial is repeated indefinitely – i.e. in the long run. The evidence conveyed by such repeated experiments is indicative of the fact that these long-run frequencies are fixed, natural properties of these random phenomena. This law-like property of the frequency theory of chance has been labelled *the Empirical Law of Large Numbers*, which dictates that as the number of trials approaches infinity, the observed relative frequency of an event converges towards its theoretical probability. For example, it would not be completely surprising if a coin, tossed 6 times, revealed 'heads up' 4 times – a relative frequency of 7/10 – or 0.7, but tossed 10000 times, the same relative frequency of 7000/10000 would be far more surprising – instead something *closer to* 5000/10000 would be expected. As the number of repetitions increases, the variation in the proportion of heads tends to diminishes. So, whist apparently random results occur for each individual trial, very uniform relative frequencies emerge for sufficiently large, collective outcomes - a phenomenon which is confirmed in practice. Thus, as the number of repetitions becomes significantly large, the relative frequency of 'heads up' is expected to be closer to the theoretical probability of 0.5, and the 'probability' of 'heads' is identified with the limit, as the number of repetitions tends towards infinity, of the relative frequency of 'heads'.

[84] Whilst randomization certainly provides safeguards against certain kinds of bias, admittedly, it is difficult to see how it can remove a *systematic* difficulty (that would simply persist in the long run), like setting up the apparatus incorrectly, because, for example, the experimenter believes that this incorrect way *is* the correct way.

[85] Howson, C. (2003).

So frequentists argue that randomization, through "the actual manipulation of the physical apparatus used in games of chances, cards, dice, roulette etc., or, more expeditiously, from a published collection of random sampling numbers"[86] justifies inferring a determinate chance distribution from the null hypothesis. Accordingly, randomization, some believe, guarantees that the relative frequency with which particular results occur in the long run, is approximately the same as the calculated probability of that result occurring using the chance distribution associated with the null hypothesis.

Of course, what still requires explanation, is *why* randomization ensures the connection between the null hypothesis and a determinate chance distribution – it is not enough to simply contend that *it does*. Consider a repeatable test, and a particular outcome O of it. A series of repetition of this test is said to be *i.i.d. – 'independent and identically distributed'* – if the following two conditions hold: (a) the probability of O occurring for any individual trial is fixed and equal say, to p; and (b) the probability of observing some specific string of Os and non-Os for a trial repeated any number of times, is merely the product of their individual probabilities p and (1-p).[87] Bernoulli presented a proof which shows that when a series of repeated trials satisfies i.i.d., then as the number of repetitions tends towards infinity, the probability that the proportions of Os and p differ by a number as small as one wishes, converges to 1[88] - in other words, the relative frequency of Os tends towards the theoretical probability p of getting an O. Now, von Mises[89] explains that outcomes which manifest in a random sequence of repeated trials *are* i.i.d. and therefore, as the number of repetitions increases, the chances that any particular sequence of outcomes occurs is approximately equal to the result derived from the chance distribution associated with the nature of the given experiment. Whilst I do not intend to articulate this account in any detail, it is worth noting that von Mises achieves this result by appealing to two principles: (1) The principle of *convergence*, which says that relative frequencies for repeated trials should converge towards some distinctive value, and (2) outcomes within the sequence of repeated trials must occur *randomly*, where randomness, according to von Mises, is satisfied when there cannot be any algorithm which can use information regarding past outcomes to identify an infinite subsequence within which the relative frequencies of the different outcomes are *different* to the relative frequencies in the sequences of outcomes actually observed. It follows deductively from von Mises' two principles, that the outcomes which constitute a random sequence of trials, are i.i.d., and thus afford a determinate chance distribution.

---

[86] Fisher, R. A. (1935), *The Design of Experiment*, London: Oliver and Boyd, p. 11.
[87] Howson, C. (2000).
[88] Bernoulli, J. (1715), *Ars Conjectandi*, Basle.
[89] von Mises, R. (1964), *The Mathematical Theory of Probability and Statistics*, New York: Academic Press.

So Fisher believed that it is a necessary and a sufficient condition that an experiment be properly randomized in order for the null hypothesis to be identified with a determinate chance distribution. Randomization "is the only point in the experimental procedure in which the laws of chance, which are to be in exclusive control of our frequency distribution, have been explicitly introduced."[90]

Before I examine the cogency of this argument, I wish to articulate exactly why I suggest that frequency testing, as construed above, captures an effective wholesale version of the no miracles intuition.

How frequentist testing captures the no miracles intuition

I suggest now that frequentist testing, by way of the chance model discussed, captures effectively the no miracles intuition encapsulated in what I referred to as "a summary of the so-called 'textbook' formulations we have looked at thus far." Whilst, as we shall see, my account of the no miracles argument is certainly a wholesale interpretation, I contend that it is the nature of an *individual* test which accounts for the 'miracle' ingredient of the no miracles argument.

Steps (i) and (ii) say that if the scientific claims are not good approximations to the actual features of the data (e.g. their population parameters), then whenever we observe values sufficiently close to the ones hypothesized, it would just be very lucky – a mere chance occurrence.

The null hypotheses H_0 of the frequentist tests suggest that the scientific claims are not even approximately true. The chance distribution associated with each individual null hypothesis is the probability function which maps the random variable being measured to the distinct probability formula, used to determine the chances, in the long run, of observing particular values the variable may instantiate. As mentioned, each type of scientific claim will warrant its own particular chance distribution, depending on the characteristics of the random variable under consideration. This link, between the null hypothesis and a determinate chance distribution, thus enables us, under the assumption that the null hypothesis is true, to calculate 'how miraculous' the measurements are that were actually observed, so that we can infer, when the results are sufficiently 'miraculous', that the null hypotheses must be false. The test is 'significant' when, assuming the null hypothesis is not

[90] Fisher, R. A. (1935), p. 19.

even approximately true, the chance of observing a result like the one we did, is sufficiently small to warrant rejecting the null hypothesis. Significant results are therefore analogous to the 'miraculous event' incorporated in the no miracles intuition, which says that, it would be a miracle if the tested claims were false, yet made accurate predictions, such that we ought to reject the assumptions that they are false, and thus infer that they are at least approximately true.[91]

Clarifying the distinction between an individual hypothesis test and the long run activity of frequentist testing in science

It is crucial to understand that I am *not* presenting the *individual* hypothesis test as *the* argument which captures the no miracles intuition. As discussed above, it is *randomization* which guarantees the necessary connection between the null hypothesis and a determinate chance distribution. Only when the random trial is repeated a number of times, can a determinate chance distribution be linked to the null hypothesis. This ensures that we can plausibly calculate probabilities associated with observing the results we actually do, and decide accordingly whether to accept or reject the null hypothesis. But the probabilities *do not* express the chances of rejecting the null hypothesis (when it is in fact true), *particular* to the *individual* hypothesis test. For example, for an individual test, when we set our significance level at 5%, this does not mean that there is a 5% chance of rejecting the given null hypothesis when it is false. It is only by virtue of *repeating* the test a number of times that we are able to calculate probabilities using the determinate chance distribution. For an *individual* test, we cannot quantify our degree of certainty. *There is no inference that can be made, for any particular test, about the truth of the individual scientific claim being made*.

Instead, an individual hypothesis test is merely an *instruction* for the long run activity of frequentist testing. The instruction is that, if we repeatedly reject null hypotheses when the test statistic falls within the rejection region (or the p-value associated with the test statistic is significantly small), then *in the long run* we can expect only to have incorrectly rejected the null hypothesis 5% of the time (or correctly rejected it 95% of the time – our 'degree of confidence' in our decisions). Only in the long run can we quantify our degree of confidence in having made the right decision. The frequency theory of chance tells us that as the number of times repetition of the randomized trial increases, the relative frequency of correct decisions to total decisions made converges towards the theoretical probability afforded to us by the determinate chance distribution.

[91] Although Fisher did not use this methodological tool for the purposes of the no miracles argument, he is credited with being the first to suggest that this chance model could be used to draw conclusions from data.

Whilst repetition of trials, on the frequentist account, ensures that scientists embrace the correct hypothesis approximately 95% of the time (or whatever percentage is identified with the confidence level), then *in sum*, roughly 95% of claims taken on board in the development of scientific theories will be (approximately) true. Of course, I am not suggesting that we necessarily need to be able to *quantify* the percentage of scientific claims or theories that are true; the point is that a *large* percentage of them will be (approximately) true insofar as each sequence of repetitions for an individual test work according to a low significance level, like 5% (or equivalently a high confidence level of say 95%). There may be some additive procedure for finding a weighted average[92] of the (often different) confidence levels afforded by each of the various claims taken on board, so that we can quantify approximately the percentage of these claims that are (approximately) true. But as I suggested, such a formula is not essential to my argument; my point, quite simply, is that, from the fact that scientists are very confident about the decision they have made to take *each* of the individual claims or theories on board, then it follows that we can be very confident about the belief that *most* of the claims or theories we have are (approximately) true; the scientific enterprise affords us a significantly larger number of *true* theories than false ones.

Recall that I agreed with Worrall, that a *wholesale* interpretation of the no miracles argument which amounts simply to the union of a number of individual retail arguments for realism, is trivial and does not offer anything illuminating about science in general. I concur with this sentiment – although my conception *involves* the practice of frequentist testing, and thus an accumulation of *individual* hypothesis tests, it does not simply *amount to just this*. I am not merely arguing that we have *a lot of* (approximately) true theories in science, in the sense that the sum of all the individual claims or theories accepted by scientists constitutes *many* of them. Such a wholesale argument, I agree, *would be* trivial, amounting merely to the union of individual retail arguments. However, my account appeals to the nature of frequentist testing in the long, which *justifies* the inference being made from a high degree of confidence in the truth of individual theories to a (similar) high degree of confidence in science in general. There is no trivial union of individual theories here, but instead an *inference* which is made about science in general. I contend that this interpretation affords precisely what I have previously argued the scope of the no miracles argument ought to take on – a (non-trivial) *wholesale* argument for scientific realism.

[92] Perhaps the weighting of the confidence levels for each individual test would be the relative frequency of the number of repeated trials that test affords out of the total number trials conducted for *all* tests the claims of which are embraced. This is just a suggestion, and admittedly, the best estimate for the weighting may require some, more sophisticated, mathematical procedure.

Why this wholesale inference adds more than just the normal scientific practice of accepting and rejecting hypotheses

I have argued previously that the no miracles argument ought to be interpreted as a wholesale, argument for realism – an inference which ought to add more than just the normal scientific practice of hypothesis testing. There is no *testable* hypothesis which could discern the (approximate) truth of science in general. Such a procedure is inconceivable since there are no new possible phenomena that could confirm the "theory" of scientific realism. The only type of phenomenon scientific realism aims to explain is the success of science. Besides, if there was such a test, it would render the wholesale no miracles argument superfluous – the no miracles argument is appealed to precisely *because* it aims to justify the truth of science in general; but if there were some independent way to determine this, there would be no need for the miracles argument in the first place.

My interpretation, I claim, avoids any contentious appeal to, what we saw Worrall refer to as, the vague notion of 'present mature science' that I discuss in Chapter 1 of my paper. Instead, a successful claim is identified with the rejection of the null hypothesis at the given significance level, such that in the long run, consequently, this *gives rise to* the class of successful theories. Moreover, there is no need to *count* what percentage of these successful theories are (approximately) true – the wholesale inference we are afforded, justifies by its very nature, the fact that science gives rise to many more true theories than false ones.

Recall that Lipton, as we saw previously, suggests that it is clear that the no miracles argument is an *instance* of the method used by scientists, and that, however, its aim has wider appeal: to support the claim that Inference to the Best Explanation is reliable. However, I showed that an argument which presupposes Inference to the Best Explanation in order to justify the reliability of Inference to the Best Explanation, is viciously circular and begs the question against opponents of realism, let alone fails to capture the essential ingredients of the no miracles intuition. Since the formalization presented in this chapter does *not* consider the no miracles argument "an instance of the method used by scientists", as argued above, it also avoids problems of circularity commensurate with formalizations which do.

So, whilst my interpretation does appeal to the actual activities scientists employ to draw conclusion, I have argued that the formalization does not amount just to this - it is not *a test for* the (approximate) truth of science in general. Instead of being an *instance* of the methods used by scientists, this argument involves the *long run* scientific activity of frequentist testing - the collective tests of hypotheses, carried out on a continual basis, which *begets the inference* that the scientific enterprise gives rise to a significantly larger number of *true* theories than false ones. And here we have the wholesale inference we sought, which says that the science in general is (approximately) true.

And finally, the formalization

1) Frequentist testing[93] is the most widely used method in science, used in the practice of accepting and rejecting hypotheses.
2) According to frequentist testing, hypotheses which make significantly unlikely, non-*ad hoc* yet successful predictions, are retained to constitute our scientific theories.
3) According to the 'frequency theory of chance', as the number of randomized 'trials' of the individual tests increases, the chances of embracing the hypothesised claim (rejecting H_0 when it is actually true) converges towards the significance level of the individual tests – a small percentage (otherwise said: the confidence level converges towards the confidence level of the tests).
4) If we have a high degree of confidence in the truth of individual theories, then we can infer that scientists will accumulate significantly more true theories than false theories.
5) Therefore, it can reasonably be inferred that science in general is true.

This, I claim, captures what has been left wanting - the *wholesale* intuition of the no miracles argument, inferring the (approximate) truth of science in general.

Additionally, this formalization, unlike the wholesale versions we have already looked at, has the advantage of avoiding problems of circularity. Since this account does not view the no miracles argument *as* a scientific test - an *instance* of the methods used by scientists – it makes no presupposition of the reliability of frequentist testing in order to justify the reliability of this very method. Whilst the methods scientists actually employ are appealed to, the conclusion concerns the (approximate) truth of science in general, and *not* the reliability of the methods used. Moreover, over and above just avoiding the problem of circularity, I suggest that this formalization has far

[93] In fact, I will argue that it is frequentist testing *in combination with* 'severe tests', which actually constitute this premise.

wider appeal than those wholesale version we have looked at which fall victim to the challenge of circularity: whereas the latter has appeal exclusively to those who *already* endorse scientific realism yet wish merely to settle disputes about the degree of approximation to truth (what I have argued does not even accord with the no miracles intuition), the former appeals to anyone, whether realist *or* anti-realist, contingent only on the fact that they already endorse frequentist testing - the methodology most widely accepted in science. Recall that I asked, in Chapter 2, for whom the no miracles argument is supposed to have appeal; according to this no miracles formalization, the argument *can* in theory be accepted by someone who is not yet convinced of scientific realism – and this I have argued previously, is in part, the purpose of the no miracles argument.

I contend that my formalization affords us the justification of inferring, from the success of science, to the (approximate) truth of science in general. This formalization, I suggest, therefore avoids problems of circularity, *and* also serves the proper purpose of the no miracles argument – to appeal to those who do not yet subscribe to scientific realism, whilst ensuring that all the essential ingredients of the wholesale no miracles intuition are captured effectively.

Challenges

'Alternative explanations' and my response

As I have mentioned previously, whenever a theory makes a successful prediction, it is not *necessarily* by virtue of the fact that the theory is true, just as we said that whenever a theory *is* true, it does not *necessarily* follow that a specific prediction will be successful. Recall, however, that authors such as Boyd and Psillos contend that, if a successful prediction had been made on some occasion, the *best explanation* is that the theory that made the prediction is true - an alternative explanation would involve less likely hypotheses. According to their approaches, a scientific success could be by virtue of different reasons, such as chance, divine providence, etc., but it is still consistent with realism to hold that *on the whole*, the best explanation of the success of some theory is because this theory is true, by and large.

However, instead of appealing to Inference to the Best Explanation, as these authors have in order to justify the (approximate) truth of a theory or hypothesis, it has been my goal to argue that we should appeal, instead, to the normal scientific practice of accepting and rejecting hypothesis, using frequentist (and severe tests, as we shall see). It is precisely such an appeal, *instead of* one to Inference to the Best Explanation, which is required, as I have already argued, to avoid the critical problem of circularity that besets the wholesale no miracles argument construed as an Inference to the Best Explanation.

So the challenge of alternative hypotheses suggests that frequentist testing fails to take into account the possibility of these alternative explanations. For instance, regarding my example, the manufacturer may have rigged the testing apparatus to record unrealistically high results; he may have doctored his batch of batteries that he made available especially for the experiment; and so on. If these situations are plausible, then the chances of observing, for example, a value of 95.5 or less if the null hypothesis is true, is the combined chance of *all* possible alternatives. Of course, one might say that these possibilities are outlandishly unlikely such that they need not to be accounted for insofar as their chances of occurrence are negligible. However, recall that it is precisely the *point* of such significance tests *to determine* plausibility of supposed facts – it seems impossible to distinguish between plausible and implausible situations when these situations *themselves* would require experimental testing for plausibility. In other words, if in order to successfully infer a determinate chance distribution which happens to be a function of the chances of a host of events which cannot themselves be tested for plausibility, then we would be attempting to calculate something which cannot be calculated. Without a *determinate* chance distribution then, we are unable to calculate probabilities of events, and thus, unable to assess whether the results observed are significant or not – we are left with no standard against which we can draw conclusions about the truth of scientific claims being advanced. If this is correct, then steps (i) and (ii) of the formulation of the no miracles argument presented at the beginning of this chapter, now seem false: For step (i), our scientific theories may predict sets of data; the theories may be false; yet the alignment of what the theories predict and the observational data may be in virtue of something *other than* chance. For step (ii), once we have accounted for the alternative explanations, the data observed may be less unlikely than we originally thought, so much so that we would no longer count it as sufficiently significant so as to reject the null hypotheses.[94]

I believe that there are reasons to suggest that the problem of alternative explanations is not as worrisome for frequentist testing as it is made out to be. Advocates of the challenge fail to recognise the importance of 'local testing' in science. In practice, scientists (at least the good ones) do not, without further enquiry, commit themselves to the conclusions afforded to them by frequentist

[94] Howson, C. (2000).

testing. Scientists are in the business of testing which hypotheses and auxiliaries are *better* tested and more complete, and which afford greater explanatory power. Scientific methodology employs *additional* tests to determine which hypotheses have passed tests which are "genuine, independent, reliable, and non-*ad hoc*. . ."[95]

Mayo, in fact, has advanced a sophisticated procedure to determine the plausibility of different hypothesized explanations of an anomalous result, by determining the *severity* of the tests that each explanation passes.[96] Mayo argues, in her *error statistics* account, that contrary to what the challenges suggest, it is not necessary to actually hold constant all factors so that an acceptable explanation for the anomaly can be identified. Anomalous results, she argues, can tell us much about their source if we are able to: "distinguish the pattern of effects of different factors (e.g., a mirror distortion from a deflection effect); learn enough about the extent of an observed effect that could be attributable to a given factor (often by simulations) in order to 'subtract it out' (from the anomaly); distinguish between hypothesized explanations of the cause of an anomaly by distinguishing the severity of the tests each passes."[97]

Mayo's solution to the problem of alternative explanations makes appeal to the following general methodological rule: "Evidence *e* should be taken as good grounds for *H* to the extent that *H* has passed a *severe test* with *e*."[98] Accordingly, the alternative explanations objection would say that for whatever evidence test T manifests which would constitute having passed hypothesis H severely, there will always exist alternative hypotheses that T would pass with equal severity. In order to overcome this challenge, suggests Mayo, the data of the experiment must be *made* to convey something over and above what it would otherwise convey if it were merely observed through ordinary experimentation. The purpose of this active intervention is to mitigate erroneous attributions of experimental outcomes. So whilst the worry is in passing *H* when it is in fact false, passing a *severe test* justifies hypothesis *H* since it counts towards reasonably ruling out particular versions and degrees of this error. Put concisely, "a passing result is a severe test of hypothesis *H* just to the extent that it is very improbable for such a passing result to occur, were *H* false."[99] If *H* were in fact false, then the chance is great that a more discordant outcome ought to have occurred. Determining this probability necessitates focusing on the probability a given procedure produces for

[95] Mayo, D. G. (1997), 'Response to Howson and Laudan', *Philosophy of Science*, 64: 323-333, p. 332.
[96] Mayo, D. G. (1996). Mayo's view is a development of both, the Neyman-Pearson framework, and standard significance tests.
[97] Mayo, D. G. (1997), pp. 329-330.
[98] Mayo, D. G. (1996), p. 177.
[99] Mayo, D. G. (1996), p. 178.

identifying a given type of error. This offers the foundation for discerning the 'well-testedness' of two hypotheses, even though each is as good as the other in fitting the data. "Two hypotheses may accord with the data equally well but nevertheless be tested differently by the data. The data may be a better, more severe, test of one than of the other. The reason is that the procedure from which the data arose may have had a good chance of detecting one type of error and not so good a chance of detecting another. What is ostensibly the same piece of evidence is really not the same at all, at least not to the error theorist."[100]

All this is admittedly a very brief outline of Mayo's strategy – the actual procedure of which I do not articulate - that I appeal to in order to overcome the problem of alternative hypotheses. It is not my intention to elaborate further on this strategy, but merely to allude to the fact that frequentist testing *may* not be doomed to the challenge, as some make out to be the case. Adherents to the problem of alternative hypotheses may insist that I have merely passed the buck to severe testing; instead of effectively responding to the challenge, I have merely replaced the problem of alternatives regarding frequentist testing, to the problem of alternatives regarding severe testing. How, this challenge would go, can we be certain that the explanation for the passing of a hypothesis, say, H_A is not actually in virtue of some *other* second-order alternative hypothesis which is the root cause of H_A having passed, as opposed to our presumption that H_A is actually true? In other words, for every severe test performed, a new can of worms is opened, presenting an additional challenge of alternatives, albeit a challenge of the same nature. And if this challenge is correct, performing perhaps a second set of severe tests to distinguish between plausible and implausible hypotheses, as regards say, hypothesis H_B, which ought to be "blamed" for the anomalous result that some given hypothesis H_A passed the formative severe test, sends us on a path of infinite regress of severe testing and problems of alternative explanations.

I maintain, however, that this problem is only superficial. In order to grasp this, we need to recall the purpose of severe testing. If I am correct, Mayo's strategy is not an attempt to *guarantee* the link between severe tests and a *conclusive* attribution of blame; instead, the strategy is supposed to *justify*, in the sense of affording us *good reason to believe*, that the hypothesis which best passes these tests, is the most likely perpetrator of the anomalous result. The problem of alternative hypotheses is yet to be resolved if we are going to be tenacious about being absolutely certain when it comes to identifying "blame"; but absolute certainty is more than we expect for frequentist testing. Scientists are sufficiently consoled by concluding that the best candidate hypothesis is most

[100] Mayo, D. G. (1996), p. 178.

likely the true one, and Mayo's strategy offers us precisely this. So Mayo's severe tests not only offers a feasible solution to the problem of alternatives, it is an account which illustrates the actual extension of the methodological practice of frequentist testing, shedding light on the complete picture of what scientists actually do – a picture that I suggested in Chapter 2 is indispensible to capturing the no miracles argument itself.

Frequentist testing and the 'base rate fallacy'

Some, including Howson,[101] have argued that frequentist testing is susceptible to the base rate fallacy – the same fallacy that was discussed in the first chapter. These arguments suggest that there are similarities between the diagnostic case spelled out by the Harvard Medical School test, and frequentist testing, such that insofar as the diagnostic case is susceptible to the fallacy, then so too is frequentist testing. If this argument is sound, then it would be problematic for the account I am advancing, in light of the fact that, in Chapter 1, I have rejected interpretations of the no miracles argument *because* they fall victim to this fallacy. Thus for the purpose that I advance a consistent exposition, arguments suggesting that frequentist testing is susceptible to the base rate fallacy ought to be dealt with here.

Aris Spanos, argues against such claims, insisting that instances like the diagnostic case have none of the basic characteristics of a frequentist test - "such as legitimate data, hypotheses, test statistics, and sampling distribution."[102] First, Spanos argues that test data is legitimate only if the sample is randomly selected; however, for the diagnostic case, the given individual who did test positive is *purposefully* selected precisely *because* he tested positive, and not because the person having testing positive, is some kind of description of the incidence of the disease in the relevant population. Second, that the given individual has the disease, in the diagnostic case, does not constitute a legitimate frequentist hypothesis, since here the hypothesis represents an unknown *constant* (either the chosen individual has or does not have the disease), instead of a *random variable* as in the frequentist account. Third, instead of having legitimate test statistics and error probabilities (type I and type II errors), the diagnostic case has only *conditional* probabilities among

[101] Howson, C. (1997), 'A Logic of Induction', *Philosophy of Science*, 64:268–90, and Howson, C. (2000), *Hume's Problem*. Oxford: Oxford University Press.
[102] Spanos, A. (2010), 'Is Frequentist Testing Vulnerable to the Base-Rate Fallacy?', *Philosophy of Science*, 77: 565–583. Admittedly, I do not examine this argument in sufficient detail to warrant a definitive conclusion about its effectiveness. I merely allude to the fact that the charge of the base rate fallacy for frequentist testing *may* be overcome, appealing to responses like Spanos's.

events. Spanos tells us that in frequentist testing, error probabilities are *not* conditional, but instead, are rely integrally, instead, upon the sample size *n*. Moreover, those who support this challenge against frequentist testing appeal to “error probabilities”, in the diagnostic case, as an instrument to make inferences about events; but error probabilities, in frequentist testing, are employed to infer *procedures*. Moreover, the notion of ‘sampling distribution’ is missing from the diagnostic case; it considers only the *single* person chosen (because he has tested positive), and thus $n = 1$. But it is obvious that for $n = 1$ no reliable *test* can be made on a population parameter – no useful estimator, based on a sample size of 1, can be acquired. Spanos concludes that, since examples like the diagnostic case have none of the aforementioned features that epitomise frequentist testing, the analogy between the diagnostic case and frequentist testing is superficial, and thus the reasons the former is undermined by the base rate fallacy are not reasons which beset the latter.

Conclusion

I have argued that there does not appear to be a way to formalize the no miracles argument according to a Bayesian framework – whether wholesale *or* retail - without falling prey to a fallacy, or being confronted with constituent probabilities which cannot be understood in any reasonable way. The Bayesian interpretation fails to capture the relevant intuition, and does not provide a plausible justification of the inference from success to (approximate) truth.

We saw too that Inference to the Best Explanation not only fails to capture the essential ingredients of the no miracles intuition, but it cannot afford us an effective wholesale interpretation I have suggested is required by the no miracles argument. Either (i) we accept problems of circularity, in which case the account is fallacious or question-begging; or (ii) we can deny circularity, in which case the argument may be valid, but fails to serve the purpose the no miracles argument is *intended* to serve – to convince those too who are not *already* realists.

Whilst Inference to the Best Explanation fails to capture the no miracles argument effectively, I suggested that accounts like Boyd's and Psillos's were instrumental in helping us arrive at a more plausible formalization. I claimed that, since it is the *success of science* in general which evokes the no miracles intuition, it seems favourable to hold on to the very practices in science which are integrally bound up with its impressive success. For this reason, I suggested that what is required is an account which focuses on the methods *actually* employed by scientists, whilst at the same time, ensures that the interpretation can be infused with something *more* than just the normal scientific practice of distinguishing plausible hypotheses from implausible ones.

Instead of moving straight into an interpretation based on the aforementioned rationale, I considered Worrall's argument that the no miracles argument is merely an intuition. To his credit, I conceded that insofar as his claims regarding the failure to formalize the "argument" (both to date and in the future) are correct, then the no miracles argument *will* turn out to be just an intuition. However, I argued that, notwithstanding the fact that this conclusion makes some arguably unwarranted assumptions, Worrall's exposition regarding the 'default' position has all the ingredients of an appeal to a broader, positive argument in favour of realism. When we examined the effectiveness of this putative argument, even when the truth of some of the premises was granted, I argued that the conclusion – that we ought to be realists – is inadequately supported by the premises.

Finally, I presented a new formalization of the no miracles argument. This interpretation encapsulates the practices *actually employed* in the scientific enterprise - the activity of *frequentist testing* - used to develop and corroborate theories. We saw that, whilst this formalization does appeal to the actual activities scientists employ to draw conclusions, I argued that instead of being an *instance* of the methods used by scientists, it affords us something more than *just* the normal scientific practice of accepting and rejecting hypotheses. The formalization incorporates a *wholesale* perspective; one I have argued throughout, the no miracles argument ought to encompass.

I suggested that this interpretation can plausibly circumvent problems other accounts are susceptible to – namely, that of circularity and the base rate fallacy – and that the challenge of alternative explanations can, at least conceivably, be resolved. I suggested that this formalization has far wider appeal than those wholesale versions we have looked at which fall victim to the challenge of circularity: whereas the latter has appeal exclusively to those who *already* endorse scientific realism yet wish merely to settle disputes about the degree of approximation to truth, the former appeals to anyone, whether realist *or* anti-realist, contingent only on the fact that they already endorse frequentist testing. This account therefore enables the no miracles argument to perform its intended purposes – to persuade someone who is *not yet* convinced of scientific realism.

I contend therefore that it is the nature of the activity of frequentist testing in science which explains why the scientific enterprise affords us a significantly larger number of *true* theories than false ones. I have argued that not only does this account succeed in capturing the no miracles intuition, but it provides a plausible justification of the inference from the success of science to the (approximate) truth of science in general – something the other interpretations have left wanting.

References

Bernoulli, J. (1715), *Ars Conjectandi*, Basle.

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

Douven, I. (2011), 'Abduction', *The Stanford Encyclopedia of Philosophy, Edward* N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2011/entries/abduction/, accessed on the 24th September, 2012.

Fine, A. (1986), 'Unnatural Attitudes: Realist and Instrumental Attachments to Science', *Mind*, 95: 149-179.

Fisher, R. A. (1925), *Statistical Methods for Research* Workers, 12th edn., London: Oliver and Boyd.

Fisher, R. A. (1935), *The Design of Experiment*, London: Oliver and Boyd.

Howson, C. (1997), 'A Logic of Induction', *Philosophy of Science*, 64: 268–90.

Howson, C. (2000), *Hume's Problem: Induction and the Justification of Belief*, Oxford: Clarendon Press.

Laudan, Larry. (1981), 'A Confutation of Convergent Realism', *Philosophy of Science,* 48: 19-49.

Leader, S. (2011), 'Is the Pessimistic Meta-Induction on the History of Science a Decisive Argument in Favour of Scientific Realism?', (not published).

Lewis, P. (2001), ''Why the Pessimistic Induction is a Fallacy', *Synthese*, 129: 371–380.

Lipton, P. (2004), *Inference to the Best Explanation*, Routledge: London.

Magnus, P.D., & Callender, C. (2004), 'Realist Ennui and the Base Rate Fallacy', *Philosophy of Science*, 71: 320–338.

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*, Chicago: The University of Chicago Press.

Niiniluoto, I. (1998), 'Truthlikeness: The Third Phase', *British Journal for the Philosophy of Science*, 49: 1-31.

Papineau, D. (1996), *The Philosophy of Science*. Oxford: Oxford University Press.

Popper, K. R. (1983), *Realism and the Aim of Science*, London: Rowman and Littlefield.

Psillos, S. (1999), *Scientific Realism: How Science Tracks Truth*, London: Routledge.

Putnam, H. (1975), *Mathematics, Matter and Method,* Vol. I. Cambridge: Cambridge University Press.

Spanos, A. (2010), 'Is Frequentist Testing Vulnerable to the Base-Rate Fallacy?', *Philosophy of Science*, 77: 565–583.

Tversky, A., & Kahneman, D. (1982), 'Evidential Impact of Base Rates', in Daniel Kahneman, Paul Slovic, and Amos Tversky (eds.), *Judgement under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, pp. 153–160.

Underhill, L. & Bradfield, D. (2007), *INTROSTAT*, Cape Town: Juta.

von Mises, R. (1964), *The Mathematical Theory of Probability and Statistics*, New York: Academic Press.

Worrall, J. (1989), 'Structural Realism: The Best of Both Worlds?', *Dialectica* 43: 99-124.

Worrall, J. (2005), *Miracles, Pessimism and Scientific Realism*. Unpublished, revised paper first given at Lunchtime Colloquium at the Center for Philosophy of Science, Pittsburgh.