

STATHIS PSILLOS

## THINKING ABOUT THE ULTIMATE ARGUMENT FOR REALISM

### 1. INTRODUCTION

Alan Musgrave has been one of the most passionate defenders of scientific realism.<sup>1</sup> Most of his papers in this area are, by now, classics. The title of my paper alludes to Musgrave's piece 'The Ultimate Argument for Realism', though the expression is Bas van Fraassen's (1980, p. 39), and the argument is Hilary Putnam's (1975, p. 73): realism 'is the only philosophy of science that does not make the success of science a miracle'. Hence, the code-name 'no-miracles' argument (henceforth, NMA). In fact, NMA has quite a history and a variety of formulations. I have documented all this in my (1999, chapter 4). But, no matter how exactly the argument is formulated, its thrust is that the success of scientific theories lends credence to the following two theses: a) that scientific theories should be interpreted realistically and b) that, so interpreted, these theories are approximately true. The original authors of the argument, however, did not put an extra stress on *novel* predictions, which, as Musgrave (1988) makes plain, is the litmus test for the ability of any approach to science to explain the success of science.

Here is why reference to novel predictions is crucial. Realistically understood, theories entail too many novel claims, most of them about unobservables (e.g., that there are electrons, that light bends near massive bodies, etc.). It is no surprise that some of the novel theoretical facts a theory predicts may give rise to novel *observable* phenomena, or may reveal hitherto unforeseen connections between known phenomena. Indeed, it would be surprising if the causal powers of the entities

<sup>1</sup> I want to dedicate this paper to Alan Musgrave. His exceptional combination of clear-headed and profound philosophical thinking has been a model for me. His commitment to, and defence of, realism have inspired and guided my own work in this area. I hope that our residual disagreements will not obscure our deep agreement. Sections 5 to 8 were inspired by a paper by P. D. Magnus and Craig Callender, titled 'Retail Realism and Base Rate Neglect'. I want to thank Magnus and Callender for many useful comments.

posited by scientific theories were exhausted in the generation of the already *known* empirical phenomena that led to the introduction of the theory. So, on a realist understanding of theories, novel predictions and genuine empirical success is to be expected (given of course that the world co-operates).

The aim of this paper is to rebut two major criticisms of NMA. The first comes from Musgrave (1988). The second comes from Colin Howson (2000). Interestingly enough, these criticisms are the mirror image of each other. Yet, they both point to the conclusion that NMA is *fallacious*. Musgrave's misgiving against NMA is that if it is seen as an *inference to the best explanation*, it is deductively fallacious. Being a deductivist, he tries to correct it by turning it into a *valid deductive argument*. Howson's misgiving against NMA is that if it is seen as an *inference to the best explanation*, it is *inductively* fallacious. Being a subjective Bayesian, he tries to correct it by turning it into a *sound subjective Bayesian argument*. I will argue that both criticisms are unwarranted.

Actually, I would have no problem with Musgrave's version of NMA if deductivism were correct. But, as I will try to argue, the deductivist stance is both descriptively and normatively wrong. To avoid a possible misunderstanding, let me note that I have no problem with deductive logic (how could I?). My problem is with deductivism, that is the view that, as Musgrave (1999a, p. 395) puts it, 'the only valid arguments are deductively valid arguments, and that deductive logic is the only logic that we have or need'. One could cite Bayesianism as a live example of why deductivism is wrong. But, I think, there are important problems with Bayesianism too.<sup>2</sup> Put in a nutshell, the Bayesian critique of NMA is that it commits the base-rate fallacy. Howson tries to rectify this by arguing that a 'sounder' version of NMA should rely explicitly on *subjective* prior probabilities. Against the Bayesian critique of NMA I will primarily argue that we should resist the temptation to cast the no-miracles argument in a subjective Bayesian form. However, I will also explore the possibility of accepting a more objective account of prior probabilities, if one is bent on casting NMA in a Bayesian form.

Here is a brief summary of the menu. Section 2 defines scientific realism and investigates Musgrave's own understanding of it. Section 3 explains, rather briefly, what I take the form and the aim of the no-miracles argument to be. Section 4 criticises Musgrave's deductivism and his attempt to show that NMA is best understood as a deductive enthymeme. Section 5 explains how NMA (as an inductive argument) is supposed to commit the base-rate fallacy. Section 6 argues that there are ways to give a more objective account of the prior probabilities that are supposed to be necessary for NMA to be inductively sound. Section 7 explores some features of the base-rate fallacy and explains why it is reasonable to ignore the base-rates (let's say the prior probabilities, though they are not the same) on certain occasions. Section 8 argues that if we look at case histories we can have strong reasons to be realists about several theories. Section 9 explores two ways to think of NMA that do not involve prior probabilities.

<sup>2</sup> I have tried to explore some of these problems in my (2006).

## 2. WHAT IS SCIENTIFIC REALISM?

I take the following three theses as constitutive of scientific realism (cf. my 1999, xix-xxi; 2000).

*The Metaphysical Thesis:* The world has a definite and mind-independent structure.

*The Semantic Thesis:* Scientific theories are truth-conditioned descriptions of their intended domain. Hence, they are capable of being true or false. The theoretical terms featuring in theories have putative factual reference. So if scientific theories are true, the unobservable entities they posit populate the world.

*The Epistemic Thesis:* Mature and predictively successful scientific theories are well-confirmed and approximately true. So entities posited by them, or, at any rate entities very similar to those posited, inhabit the world.

Musgrave (1996, p. 23) agrees that realism involves the *Semantic Thesis*. He is not very explicit about the *Metaphysical Thesis*. Actually, he is quite critical of the realist view which ‘erects current science into a metaphysic and ties scientific realism too closely to that metaphysic’ (1996, p. 21). As I understand it, the *Metaphysical Thesis* means to make scientific realism distinct from all those anti-realist accounts of science, be they traditional idealist and phenomenalist or the more modern verificationist accounts which, based on epistemic accounts of truth, allow no divergence between what there is in the world and what is issued as existing by a suitable set of epistemic practices and conditions. It implies that if the unobservable natural kinds posited by theories exist at all, they exist independently of our ability to be in a position to know, verify, recognise etc. that they do. Musgrave does accept all this. Throughout his work on realism, he has defended a non-epistemic conception of truth and has argued very persuasively against epistemic conceptions of truth. He has also defended the mind-independent existence of the world (see, for instance his 1989; 1996). So he does, after all, accept a version of the *Metaphysical Thesis* above.

When it comes to the *Epistemic Thesis*, Musgrave seems to distinguish between two versions of it: a weak and a strong one. He does accept the weak version. For, he thinks ‘that some scientific entities do exist and that some of what science tells us about them is true’ (1996, p. 21). He calls ‘ludicrous’ the view that ‘all scientific theories are false’ (1996, p. 22). But he (1996, pp. 19-21) seems to take the strong version of the *Epistemic Thesis*, which he associates with what he calls ‘mad-dog realism’, to imply commitment to *all* entities posited by current theories and belief in *everything* they say about them. He is quite clear that he denies this strong version. He protests that this view is overly optimistic and unwarranted. I think he is quite right when he says: ‘We should be more confident about atoms and molecules

than we are about electrons, and more confident about electrons than we are about quarks and gluons' (1996, p. 22). He is equally right when he adds: 'Realism about the entities and theories of current science should rather be guarded' (1996, p. 22).

Guarded realism is still realism! Guarded realists need not take current science uncritically. They need not commit themselves to everything that current science asserts. They can have a differentiated attitude towards the theoretical constituents of modern science: some of them are better supported by the evidence than others; some of them play an indispensable explanatory role, while others do not; some contribute to the successes of theories, while others do not. But, I think, we should not lose sight of the general philosophical issue at stake. I take it to be this: is there any strong reason to believe that science cannot achieve theoretical truth? That is, is there any reason to believe that after we have understood the theoretical statements of scientific theories as expressing genuine propositions, we can never be in a warranted position to claim that they are true (or at least, more likely to be true than false)? What the *Epistemic Thesis* means to assert is that theoretical truth is achievable (and knowable) no less than is observational truth. So, the *Epistemic Thesis* is meant to be optimistic: science has succeeded in tracking truth. To be sure, this requires a certain *epistemic luck*: it's not a priori true that science has been, or has to be, successful in truth-tracking. If science does succeed in truth-tracking, this is a *radically contingent fact* about the way the world is and the way scientific method and theories have managed to 'latch onto' it.

The debate about the *Epistemic Thesis* has brought to focus one central issue: are the ampliative-abductive methods of science reliable and can they confer justification on theoretical assertions? The defence of the *Epistemic Thesis* requires a positive answer to this question. For, it is part of the realist thesis that the ampliative-abductive methods employed by scientists to arrive at their theoretical beliefs are reliable: they tend to generate approximately true beliefs and theories. The *no-miracles argument* (NMA) has played a pivotal role in this defence.

### 3. THE NO-MIRACLES ARGUMENT

How does NMA support the *Epistemic Thesis*? As I have argued elsewhere (cf. Psillos 1999, chapter 4), the structure and role of NMA in the realism debate is quite complex. To a good approximation, it should be seen as a grand Inference to the Best Explanation (IBE). The way I read it, NMA is a philosophical argument which aims to defend the reliability of scientific methodology in producing approximately true theories and hypotheses. I don't want to repeat here the exact formulation of the argument (see Psillos 1999, pp. 78-81). However, I want to emphasise that its conclusion has two parts. The *first* part is that we should accept as (approximately) true the theories that are implicated in the (best) explanation of the *instrumental* reliability of first-order scientific methodology. The *second* part is that since, typically, these theories have been arrived at by means of IBE, IBE is reliable. Both parts are necessary for my version of NMA.

The main strength of NMA rests on the first part of the conclusion. Following more concrete types of explanatory reasoning which occur all the time in science, it

suggests that it is reasonable to accept certain theories as approximately true, at least in the respects relevant to their theory-led predictions. So, it is successful instances of explanatory reasoning in science which provide the *basis* for the grand abductive argument. However, NMA is not *just* a generalisation over the scientists' abductive inferences. Although itself an *instance* of the method that scientists employ, it aims at a much broader target: to defend the thesis that Inference to the Best Explanation, (that is, a *type* of inferential method) is reliable. This relates to the second part of its conclusion. What, I think, makes NMA distinctive as an argument for realism is that it defends the achievability of theoretical truth. The second part of the conclusion is supposed to secure this. The background scientific theories, which are deemed approximately true by the first part of the conclusion, have themselves been arrived at by abductive reasoning. Hence, it is reasonable to believe that abductive reasoning is reliable: it tends to generate approximately true theories. This conclusion is not meant to state an a priori truth. The reliability of abductive reasoning is an empirical claim, and if true, it is contingently so.

It should be noted that, as I conceive of it, NMA needs a qualification. Although most realists would acknowledge that there is an explanatory connection between a theory's being empirically successful and its being, in some respects, right about the unobservable world, it is far too optimistic—if defensible at all—to claim that *everything* that the theory asserts about the world is thereby vindicated. So, realists should *refine* the explanatory connection between empirical and predictive success, on the one hand, and truthlikeness, on the other. They should assert that these successes are best explained by the fact that the theories which enjoyed them have had *truthlike theoretical constituents* (i.e., truthlike descriptions of causal mechanisms, entities and laws). The theoretical constituents whose truthlikeness can best explain empirical successes are precisely those that are essentially and ineliminably involved in the generation of predictions and the design of the methodology which brought these predictions about. From the fact that not every theoretical constituent of a successful theory does and should get credit from the successes of the theory, it certainly does *not* follow that none do (or should) get some credit.

There are a number of objections to this explanationist version of NMA. One of them has also been pressed by Musgrave (1988, p. 249; 1999, pp. 289-90), and this is particularly hurtful. The objection is that NMA is *viciously* circular: it employs a second-order IBE in defence of the reliability of first-order IBEs. As is explained in detail in my (1999, chapter 4), the abductive defence of realism proceeds within a broad naturalistic framework. Within this framework, the charge of circularity loses most of its bite because what is sought is not justification of inferential methods and practices (at least in the neo-Cartesian internalist sense) but their explanation and defence (in the epistemological externalist sense). In any case, I (1999, pp. 81-90) argued that (a) there is a difference between premise-circularity and rule-circularity (a premise-circular argument employs its conclusion as one of its premises; a rule-circular argument conforms to the *rule* which is vindicated in its conclusion); (b) rule-circularity is *not* vicious; and (c) the circularity involved in the defence of basic rules of inference is rule-circularity. Though these points had already been made with regard to basic deductive and inductive rules, I showed how the above defence of IBE is rule-circular. So, the employment of IBE in an abductive defence of the

reliability of IBE is *not* viciously circular. As a support of all this consider the following case. Many (if not all) use modus ponens unreflectively as an inferential rule and yet the establishment of the *soundness* of modus ponens proceeds with an argument which effectively uses modus ponens. This procedure can still explain to modus ponens-users why and in virtue of what features deductive reasoning is sound.

Being a deductivist, Musgrave thinks that the *only* kind of validity is deductive validity. He denies that there are such things as non-deductive cogent arguments (cf. 1999a). He takes it that rule-circular arguments in favour of inferential rules may have only some *psychological* force (cf. 1999, pp. 289-90). But he (1999, p. 295) is aware of the point that the proof of the soundness of modus ponens requires the use of modus ponens. How does he react to this? It seems that he has wavered between two thoughts. The *first* is that ‘there is little future in the project of ‘justifying deduction’’ (1999, p. 296). As he acknowledges, ‘Any “justification” which is non-psychologistic will itself be a deductive argument of some kind, whose premises will be more problematic than the conclusion they are meant to justify’ (*ibid.*) To be sure, he immediately adds that there is a difference between deductive rules and non-deductive (ampliative) ones in that, even if neither of them can be ‘justified’, non-deductive rules can be *criticised*. But how much pause should this give us? Let us grant, as we should, that none of our basic inferential rules (both deductive and non-deductive) can be ‘justified’ without rule-circular arguments. The fact that the non-deductive rules can be criticised more severely than the deductive ones may make us be much more cautious when we employ the former. That’s all there is to it. The *second* thought that Musgrave has (cf. 1980, pp. 93-5; 1999, pp. 96-7) is that there is a sense in which deduction *can* be ‘justified’, but this requires an appeal to ‘deductive intuitions’. As he (1980, p. 95) graphically puts it: ‘In learning logic we pull ourselves up by our bootstraps, exploit the intuitive logical knowledge we already possess. Somebody who lacks bootstraps (‘deductive intuition’) cannot get off the ground’. This is, I think, exactly right. But, as I have argued in some detail in my (1999, pp. 87-9), exactly the same response can be given to calls for ‘justifying’ non-deductive rules. When it comes to issues concerning the vindication of inference to the best explanation, if one lacks ‘abductive’ intuitions, one lacks the necessary bootstraps to pull oneself up.

#### 4. DEDUCTIVISM

To realists, it might come as a surprise that Musgrave (1996, p. 19) takes realism to be, ‘first and foremost a thesis about the aim of science. It says that the aim of a scientific inquiry is to discover the truth about the matter inquired into’. So he takes realism to be an ‘axiological thesis’: ‘science aims for true theories’.<sup>3</sup> There is clear motivation for this view: even if *all* theories we ever came up with were false, realism wouldn’t be threatened (cf. 1996, p. 21). As we have seen, Musgrave does

<sup>3</sup> This axiological thesis has been a constant pillar of his realism. For some early formulation of it, see his (1977).

not think that all our theories have been, or will be, outright false. But he does take this issue (whatever its outcome may be) to have no bearing on whether realism is a correct attitude to science. There are, however, inevitable philosophical worries about the axiological characterisation of realism. *First*, it seems rather vacuous. Realism is rendered immune to any serious criticism which stems from the empirical claim that the science we all love has a poor record in truth-tracking (cf. Laudan 1984). *Second*, aiming at a goal (truth) whose achievability by the scientific method is left unspecified makes its supposed regulative role totally mysterious. Finally, all the excitement of the realist claim that science engages in a cognitive activity which pushes back the frontiers of ignorance and error is lost.

Though Musgrave does not address these worries explicitly, he does so implicitly. For, he does try to defend the prime realist argument for epistemic optimism, viz., the no-miracles argument. He (1988, p. 237; 1999, p. 60) takes NMA to be an inference to the best explanation. Besides, he (1988, p. 232; 1999, p. 119) has been one of the first to stress that what needs to be explained is *novel* success (that is, the ability of theories to yield successful novel predictions). And he has been one of the first to note that NMA should focus on the novel success of *particular* theories (cf. 1988, p. 249). He has also produced some powerful arguments to the effect that non-realists' explanations of the success of science are less satisfactory than the realist one. Most of them appear in his (1988). In fact, he (1988, p. 249) concludes that the realist explanation is the *best*. The issue then is this: Does Musgrave endorse NMA? The answer to this question is *not* straightforward.

Precisely because Musgrave takes NMA to be an inference to the best explanation, he takes it to be deductively invalid, and hence fallacious. Being a deductivist, he takes it that the only arguments worth their salt are deductive arguments. So he cannot endorse NMA, at least as it stands. Musgrave takes all *prima facie* non-deductive arguments to be *enthymemes*. An enthymematic argument is an argument with a missing or suppressed premise. After the premise is supplied (or made explicit), the argument becomes deductively valid. But it may or may not be sound (cf. his 1999, pp. 87 & 281ff). According to Musgrave, non-deductive arguments are really deductive enthymemes, with 'inductive principles' as their missing premises.

As it is typically presented, IBE has the following form (cf. Musgrave 1988, p. 239; 1999, p. 285):

(IBE)

1. F is the fact to be explained.
2. Hypothesis H explains F.
3. Hypothesis H satisfactorily explains F.
4. No available competing hypothesis explains F as well as H does.
5. Therefore, it is reasonable to accept H as true.

Given that this argument-pattern is invalid, Musgrave proposes that it should be taken to be enthymematic. The missing premise is the following epistemic principle (cf. *ibid.*):

(*missing premise*) 'It is reasonable to accept a satisfactory explanation of any fact, which is also the best explanation of that fact, as true'.

Add to (*IBE*) the missing premise, and you get a valid argument. Briefly put, the deductive version of *IBE* is this:

(*D-IBE*)

1. If hypothesis *H* is the best explanation of the fact to be explained<sup>4</sup>, then it is reasonable to accept *H* as true.
2. *H* is the best explanation of the evidence.
3. Therefore, it is reasonable to accept *H* as true.

This is a valid argument. Besides, Musgrave (1999, p. 285) thinks that 'instances of the scheme might be sound as well'. In any case, he thinks that the missing premise 'is an epistemic principle which is not obviously absurd' (1999, p. 285). In light of this, it's no surprise that Musgrave reconstructs *NMA* as an enthymeme. That's how he (1988, p. 239) puts it:

The fact to be explained is the (novel) predictive success of science. And the claim is that realism (more precisely, the conjecture that the realist aim for science has actually been achieved) *explains* this fact, explains it *satisfactorily*, and explains it *better* than any non-realist philosophy of science. And the conclusion is that it is reasonable to accept scientific realism (more precisely, the conjecture that the realist aim for science has actually been achieved) as true.

This is a deductive enthymeme, whose suppressed premise is the aforementioned epistemic principle (*missing premise*). What is worth stressing is that Musgrave takes *NMA* to aim to tell in favour of the *Epistemic Thesis* (see section 2). Though he formulates the argument in terms of his own axiological thesis, he takes it that, if successful, *NMA* makes it reasonable to accept that truth has been achieved.

I would have no problem with (*D-IBE*) if deductivism were correct. But, I think, the deductivist stance is so radically at odds with the practice of science (as well as of everyday life) that it would have to give even the most dedicated deductivist pause. Human reasoning is much broader than deductivists allow. It is *defeasible*, while deductive reasoning is not. That is, it is sensitive to new information, evidence and reasons in a way that is not captured by deductive arguments. The latter are monotonic: when further premises are added to a valid deductive argument, the original conclusion still follows. But human reasoning is non-monotonic: when new information, evidence and reasons are added as premises to a non-deductive argument, the warrant there was for the original conclusion may be removed (or enhanced). Human reasoning is also *ampliative*, while deductive reasoning is not.

<sup>4</sup> This, in effect, sums up premises (2) to (4) of (*IBE*).

That is, the conclusions we adopt, given certain premises, have excess content over the premises. Deductive reasoning is not content-increasing. In a (logical) sense, the conclusion of a valid deductive argument is already ‘contained’ in its premises.<sup>5</sup> This is not to belittle deductive reasoning. It’s the only kind of reasoning that is truth-preserving. The importance of truth-preservation can hardly be exaggerated. But we should not forget that, though deductive reasoning preserves truth, it cannot establish truth. In particular, it cannot establish the truth of the premises. If we are not talking about logical (and mathematical and analytical—if there are such things—truths), the premises of deductive arguments will be synthetic propositions, whose own truth can be asserted, if at all, on the basis of ampliative and non-deductive reasoning. So, though deductive reasoning is indispensable, it can hardly exhaust the content and scope of human (and scientific) reasoning.<sup>6</sup> As a *descriptive* thesis, deductivism is simply false.

Is then deductivism to be construed as a *normative* thesis? I am aware of no argument to the effect that deductivism is normatively correct. This is not to imply that deductive logic has no normative force. It does. But recall that deductivism is the thesis that *all* arguments worth their salt *should* be construed as deductive enthymemes. Whence could this thesis derive its supposed normative force? I don’t see a straightforward answer to this question. Musgrave suggests that reconstructing supposed non-deductive arguments as deductive enthymemes ‘conduces to clarity’ (1999, pp. 284-5). That is, it makes their premises explicit. Hence, it also makes explicit what is required for the premises to be true, and for the argument to be sound. I think, however, that this point is problematic. Non-deductive arguments (e.g., simple enumerative induction, or inference to the best explanation) are *not* unclear. If anything, the problem with them is how to justify them. But a similar problem occurs with deduction, as we saw at the end of the previous section. Suppose, however, that we leave this problem to one side. Suppose that we grant that turning a non-deductive argument into a deductively valid one conduces to clarity since it makes its premises explicit. Deductivists still face a problem: what, if anything, justifies the missing premise? To fix our ideas, consider the major premise of (*D-IBE*) above. What justifies the principle ‘If hypothesis *H* is the best explanation of the fact to be explained, then it is reasonable to accept *H* as true’? The sceptic can always object to *this* principle that it is question-begging. How can a deductivist reply to this charge?

Musgrave (1999a, p. 408) does consider this problem. He takes the sceptic to rely on the following idea, which Musgrave calls ‘justificationism’: ‘a reason for

<sup>5</sup> For more on non-deductive reasoning and on the way IBE should be understood as a *genus* of ampliative reasoning, see my (2002).

<sup>6</sup> Musgrave might reply to this by saying that scientists employ ‘demonstrative inductions’, which are really deductions, though not deductions *from* the phenomena, as Newton thought (cf. his 1999, pp. 303 & 306). I don’t want to discuss this issue here, though it certainly needs attention. Briefly put, the thrust of demonstrative induction is that premises of greater generality and premises of lesser generality will yield a conclusion of intermediate generality. But this must be noted: it is wrong to think that demonstrative induction frees us from the need to engage in ampliative inference. As Norton (1994, p. 12) notes: ‘Typically, ampliative inference will be needed to justify “the premises of greater generality”’.

believing P must justify P, show that P is true or at least probably true'. Not surprisingly, he rejects justificationism. So, if justificationism is abandoned, the fact that the reasons which support the major premise of (*D-IBE*) are not conclusive is *not* a reason not to believe in the major premise. I think this is exactly right. But it has a repercussion which Musgrave does not seem to appreciate. Justificationism has also been assumed by the sceptics in their critique of inductive (or non-deductive) reasoning. One way to put their point is that the premises of a non-deductive argument do not establish the truth of its conclusion. If justificationism is to be abandoned, as it should be, it should be abandoned in all contexts. That is, it should be abandoned for deductivism as well as inductivism. It seems, then, that Musgrave himself offers us a strong reason to hold onto inductivism.

Perhaps, deductivism is a fall-back position. It says that arguments can be *reconstructed* as deductively valid arguments. But this thesis is trivial. *Any* argument can be turned into a deductively valid one by adding suitable premises. In particular, any *invalid* argument can be rendered valid by adding suitable premises. Consider the fallacy of affirming the consequent. The argument:

1. If (if a and b) and b, then a
2. If a then b
3. b
4. Therefore, a

is perfectly valid. If all logically invalid arguments were considered enthymemes, there would be no such thing as invalidity. Musgrave is aware of this objection, too. His reply is this: '[Y]ou cannot allow anything whatever to count as a 'missing premise'; what the 'missing premise' is must be clear from the context of the production of the argument in question' (1999a, p. 399; 1999, p. 87, n. 106). But, surely, the context underdetermines the possible 'missing premises'. More importantly, for any 'missing premise', there will be *some* contexts in which it is appropriate.

To sum up, Musgrave's misgivings against NMA were motivated by the thought that if it is seen as an inference to the best explanation, it is deductively fallacious. He tried to correct it, as we have seen, by turning it into a *valid deductive argument*. We found his attempt wanting because we found deductivism wrong. What is interesting is that others, most notably Colin Howson, think that if it is seen as an inference to the best explanation, NMA is *inductively* fallacious. He tries to correct it, by turning it into a sound subjective Bayesian argument. All this will leave Musgrave totally unmoved, since he thinks there is no such thing as inductive logic (cf. 1999a). Still, for those of us who a) think that there is more to reasoning than deduction, b) are critical of subjective Bayesianism, and c) want to defend some form of NMA, it will be important to examine whether the Bayesian criticism of NMA succeeds or fails.

## 5. SUBJECTIVE BAYESIANISM TO THE RESCUE?

Howson (2000, 36) formulates the ‘no-miracles’ argument (NMA) as follows:

(A)

1. If a theory T is not substantially true then its predictive success can only be accidental, a chance occurrence.
2. A chance agreement with the facts predicted by T is very improbable—of the order of a miracle.
3. Since this small chance is so extraordinarily unlikely, the hypothesis that the predictive success of T is accidental should be rejected (especially in light of the fact that there is an alternative explanation—viz., that T is true—which accounts better for the predictive success).
4. Therefore, T is substantially true.<sup>7</sup>

He then argues in some detail that (A) is inductively fallacious. He contests the soundness of all if its premises (cf. 2000, 43). However, the novelty of Howson’s view relates to his criticism of premise (3) and of the inferential move to (4). His prime point is that (A) is wrong because it commits the base-rate fallacy.

Let me introduce the base-rate fallacy with a standard example in the literature, which is known as the Harvard Medical School test.

*(Harvard Medical School test)*

A test for the presence of a disease has two outcomes, ‘positive’ and ‘negative’ (call them + and –). Let a subject (Joan) take the test and let H be the hypothesis that Joan has the disease and –H the hypothesis that Joan doesn’t have the disease. The test is highly reliable: it has zero false negative rate. That is, the likelihood that the subject tested negative given that she does have the disease is zero (i.e.,  $\text{prob}(-/H)=0$ ). Consequently, the true positive rate, i.e., the likelihood of being tested positive given that she has the disease is unity, ( $\text{prob}(+/H)=1$ ). The test also has a very small false positive rate: the likelihood that Joan is tested positive though she doesn’t have the disease is, say, 5% ( $\text{prob}(+/-H) = .05$ ). Joan tests positive. What is the probability that Joan has the disease given that she tested positive? That is, what is the posterior probability  $\text{prob}(H/+)$ ?

When this problem was posed to experimental subjects, they tended, with overwhelming majority, to answer that the probability that Joan has the disease given that she tested positive was very high—very close to 95%.

This answer is wrong. Given only information about the likelihoods  $\text{prob}(+/H)$  and  $\text{prob}(+/-H)$ , the question above—what is the posterior probability  $\text{prob}(H/+)$ ?—is indeterminate. This is so because there is some crucial information *missing*: we are not given the incidence rate (base-rate) of the disease in the population. If this

<sup>7</sup> This formulation does not exactly match the way Howson puts the argument, but it closely resembles it.

incidence rate is very low, e.g., if only 1 person in 1,000 has the disease, then it is very *unlikely* that Joan has the disease even though she tested positive:  $\text{prob}(H/+)$  would be less than .02.<sup>8</sup> For  $\text{prob}(H/+)$  to be high, it must be the case that  $\text{prob}(H)$  be not too small. But if  $\text{prob}(H)$  is low, then it can dominate over a high likelihood of true positives and lead to a very low posterior probability  $\text{prob}(H/+)$ . The lesson that many have drawn from cases such as this is that it is a fallacy to ignore the base-rates because it yields wrong results in probabilistic reasoning. The so-called base-rate fallacy is that experimental subjects who are given problems such as the above tend to neglect base-rate information (that is, the prior probabilities), *even when* they are given this information explicitly.<sup>9</sup>

With this in mind, let us take a look at NMA. To simplify matters, let S stand for predictive success and T for a theory. According to (A) above, the thrust of NMA is the comparison of *two* likelihoods, viz.,  $\text{prob}(S/-T)$  and  $\text{prob}(S/T)$ . The following argument captures the essence of Howson's formulation of NMA (see (A) above).

(B)

1.  $\text{prob}(S/T)$  is high.
2.  $\text{prob}(S/-T)$  is very low.
3. S is the case.
4. Therefore,  $\text{prob}(T/S)$  is high.<sup>10</sup>

What's explicit in (B) is that alternative theories (or the falsity of T) fail(s) to support the evidence. Let us say that the false-positive rate is low and the false-negative rate is zero. That is, the probability of T being successful given that it is false is very small (say,  $\text{prob}(S/-T)=.05$ ) and the probability of T being unsuccessful given that it is true is zero (i.e.,  $\text{prob}(-S/T)=0$ ). Hence, the true-positive rate ( $\text{prob}(S/T)$ ) is 1. Does it follow that  $\text{prob}(T/S)$  is high? NMA is portrayed to answer affirmatively. But if so, it is fallacious: it has neglected the base-rate of truth (that is,  $\text{prob}(T)$ ). Without this information, it is impossible to estimate correctly the required posterior probability. If the base-rate of true theories is low, then  $\text{prob}(T/S)$  will be very low too. Assuming that base-rate of true theories is 1 in 100 (i.e.,  $\text{prob}(T)=.01$ ),  $\text{prob}(T/S)=.17$ . (The calculation mimics the one offered in note 7). The conclusion seems irresistible: as it stands, (B) commits the base-rate fallacy—it has neglected  $\text{prob}(T)$ , or as the jargon goes, the base-rate.

Every cloud has a silver lining, however. So, Howson (2000, pp. 55-9) urges us to think how NMA could become 'sounder' within a Bayesian framework. We are

<sup>8</sup> By Bayes's theorem,  $\text{prob}(H/+) = \text{prob}(+/H)\text{prob}(H)/\text{prob}(+)$ , where  $\text{prob}(+) = \text{prob}(+/H)\text{prob}(H) + \text{prob}(-/H)\text{prob}(-H)$ . Plug in the following values:  $\text{prob}(+/H) = .95$ ,  $\text{prob}(H) = .001$ ,  $\text{prob}(-H) = .999$ ,  $\text{prob}(-/H) = .05$ . Then,  $\text{prob}(H/+)$  is roughly equal to .02.

<sup>9</sup> This problem was first investigated by Tversky and Kahneman (1982). It was dubbed 'the base-rate fallacy' by Bar-Hillel (1980).

<sup>10</sup> To be more precise, we need to state the conclusion thus: Therefore,  $\text{prob}_{\text{new}}(T)$  is high, where  $\text{prob}_{\text{new}}(T) = \text{prob}_{\text{old}}(T/S)$ .

invited to accept that NMA can succeed only if information about base-rates (or prior probabilities) is taken into account. In effect, the idea is this:

(B<sup>1</sup>)

1. prob(S/T) is high.
2. prob(S/-T) is very low.
3. S is the case.
4. prob(T) is not very low.
5. Therefore, prob(T/S) is high.

What has been added is an explicit premise that refers to the *prior probability* of true theories. For (B<sup>1</sup>) to be sound, this probability should not be low. How low prob(T) can be will vary with the values of prob(S/T) and prob(S/-T). But it is noteworthy that, with the values of the likelihoods as above, if prob(T) is only 5%, then prob(T/S) is over 50%. To be sure, (B<sup>1</sup>) is not valid. But, as Howson (2000, p. 57) notes, it is ‘a sound probabilistic argument’. Of course, (B<sup>1</sup>) rests also on the assumption that prob(S/-T) is very low. This can be contested. But, Howson notes, there may be occasions on which this low probability can be justified, e.g., when, for instance, we think of -T as a disjunction of  $n$  theories  $T_i$  ( $i=1, \dots, n$ ) whose own prior probabilities prob( $T_i$ ) are negligible. In any case, his point is that NMA can be a sound argument only when we see that it is based on some substantive assumptions about *prior probabilities*. Being a subjective Bayesian, he takes these prior probabilities to be ‘necessarily subjective and a priori’ (2000, p. 55).

## 6. A WHIFF OF OBJECTIVISM

I will start my criticism of Howson’s argument by resisting the view that one needs to rely on *subjective* prior probabilities in formulating NMA. So for the time being at least, I will assume the foregoing Bayesian reformulation of NMA. Actually, let us reformulate (B<sup>1</sup>), based on what has been called the *Bayes factor*. This is the ratio:

(Bayes factor):  $f = \text{prob}(S/-T)/\text{prob}(S/T)$ .

Recall Bayes’s theorem:

$$\text{prob}(T/S) = \text{prob}(S/T)\text{prob}(T)/\text{prob}(S) \quad (1)$$

where:

$$\text{prob}(S) = \text{prob}(S/T)\text{prob}(T) + \text{prob}(S/-T)\text{prob}(-T).$$

Using this factor, (1) becomes this:

$$\text{prob}(T/S) = \text{prob}(T) / (\text{prob}(T) + f \text{prob}(-T)). \quad (2)$$

(B<sup>1</sup>) can then be written thus:

(B<sup>2</sup>)

1.  $f$  is very small.
2. S is the case.
3.  $\text{prob}(T)$  is not very low.
4. Therefore,  $\text{prob}(T/S)$  is high.

The Bayes factor is small if  $\text{prob}(S/-T) \ll \text{prob}(S/T)$ . Now, whether the conclusion follows from the premises depends on the prior probability  $\text{prob}(T)$ . So, the Bayes factor, on its own, tells us little. But it *does* tell us something of interest. Actually, it tells us something that can take out *some* to the sting of subjectivism in Bayesianism. Two things are relevant here. The *first* is that there is a case in which the prior probability of a theory does not matter. This is when the Bayes factor is *zero*. Then, no matter what the prior  $\text{prob}(T)$  is, the posterior probability  $\text{prob}(T/S)$  is unity. The Bayes factor is zero if  $\text{prob}(S/-T)$  is zero. This happens when just one theory can explain the evidence. Then, we can dispense with the priors. This situation may be unlikely. But it is not a priori impossible. After all, the claim that evidence underdetermines the theory is not a logical truth! Put in a different way, one quick problem that Howson's reconstructions of NMA faces is that it equates, at least implicitly, explanation with deduction. Given this equation, it is trivially true that there cannot be just one theory that explains the evidence, since there will be many (an infinite number of?) theories that entail it. In many places (cf., for instance 2000, pp. 40-1), Howson does make this equation. But this is a Pyrrhic victory over NMA. There is more to explanation than the deduction of (descriptions of) the phenomena from the theory (and deduction is not even necessary for explanation). So, it may well be the case that many theories entail (descriptions of) the relevant phenomena, while only one of them *explains* them. I won't argue for this claim now. Suffice it for the present purposes to note that equating explanation with deduction is question-begging.<sup>11</sup>

Be that as it may, let us grant, for the sake of the argument, that the case in which the Bayes factor is zero is exceptional. There is a *second* thing in relation to Bayes factor that needs to be noted. Assume some kind of indifference (or a flat probability distribution) between  $\text{prob}(T)$  and  $\text{prob}(-T)$ ; that is, assume that  $\text{prob}(T) = \text{prob}(-T) = 1/2$ . Then (2) above becomes:

$$\text{prob}(T/S) = 1 / (1+f) \quad (3)$$

Assuming indifference, the Bayes factor shows that likelihood considerations (especially the fact, if it is fact, that  $f$  is close to zero) can make T much more likely

<sup>11</sup> For more on the realist reply to the argument from the underdetermination of theories by evidence, see my (1999, chapter 8).

to be true. The point here is not that we can altogether dispense with the priors. Rather, the point is that we are not compelled to take a *subjective* view of the prior probabilities. So, there is a version of NMA which, though close to (B<sup>2</sup>) above, does *not* assume anything other than indifference as to the prior probability of T being true.

(B<sup>3</sup>)

1.  $f$  is close to zero.
2. S is the case.
3.  $\text{prob}(T) = \text{prob}(-T) = 1/2$ .
4. Therefore,  $\text{prob}(T/S)$  is high.

(B<sup>3</sup>) strikes me as fine. If one wanted to capture the *thrust* of NMA within a Bayesian framework, one could hold onto (B<sup>3</sup>). This does not commit the base-rate fallacy. Besides, it avoids the excesses of subjective Bayesianism.

So far, I have assumed that prior probabilities and base-rates are one and the same thing. In fact, Howson does assume this too. He (2000, p. 57, n.5) calls the prior probabilities ‘the epistemic analogue of the base-rate’. Normally, base-rates are given by reliable statistics. Hence, they are quite objective. When a subject is asked how probable it is that Jim (a young adult male) suffers from hypothyroidism, given that he has the symptoms, she doesn’t commit a fallacy if she ignores her own prior degree of belief that Jim has hypothyroidism. After all, she might not have any prior degree of belief in this matter. The fallacy consists in her claiming that the probability is high while ignoring some relevant factual information about hypothyroidism, viz., that it is quite rare, even among people who have the relevant symptoms. This is some objective statistical information, e.g., that only 1 in 1,000 young adult male suffers from hypothyroidism. Base-rates of this form can (and should) be the input of a prior probability distribution. But they are *not* the prior subjective degrees of belief that Bayesians are fond of. In incorporating them, Bayesians move away from a purely subjective account of prior probabilities. But what about the converse? If prior probabilities are purely (and necessarily, as Howson says) subjective, then why should an agent rely on base-rates to fix her prior probabilities? That is, why should an agent’s subjective prior probability of an event to occur be equated with the rate of the occurrence of this event in a certain population? Purely subjective priors might be assigned in many ways (and, presumably, there is no fact of the matter as to which way is the correct, or rational, one). An agent might know a relevant-base rate but, being a purely subjective Bayesian, she might decide to disregard it. She won’t be probabilistically incoherent, if she makes suitable adjustments elsewhere in her belief corpus. Or, though the base-rate of hypothyroidism in the population is very low, her subjective prior probability that Jim suffers from hypothyroidism may be quite high, given that she believes that Jim has a family history of hypothyroidism. The point here is that *if* prior probabilities are purely subjective, it seems within the rights of a Bayesian agent to fix her prior probabilities in a way different from the relevant base-rates. So, prior probabilities are not, necessarily, base-rates. Or, more provocatively,  $\text{ba}(y)$ se rates are *not* base-rates.

In light of this, something stronger can be maintained. Subjective Bayesians had better have a more objective account of prior probabilities, if they are to reason correctly (according to their own standards) and to avoid falling victims of the base-rate fallacy. For if prior probabilities are totally up to the agent to specify, then the agent seems entitled to *neglect* the base-rate information, or to adopt a prior probability which is significantly lower or higher than the base-rate. If anything, base-rates should act as an external constraint on Bayesian reasoning, by way of fixing the *right* prior probabilities. The need to take account of base-rates seems to make Bayesianism more prescriptive than it intends to be. The call to rely on the base-rates is a substantive piece of advice, which goes beyond the mere call for synchronic and diachronic coherence.

### 7. IGNORING BASE-RATES

As we have seen, the Bayesian critique of NMA (see argument (B) above) consists in the claim that it *ignores* the base-rates of truth and falsity. But there is a sense in which this is not quite correct. The Bayesian criticism presupposes that *there are* base-rates for truth and falsity. However, it is hard, if not outright impossible, to get the relevant base-rates. The issue is not *really* statistical. That is, it's not really that we don't have a list of true and false theories at our disposal. Nor, of course, is the issue that the advocates of NMA fail to take account of such a list. The issue is philosophical. The very idea of a base-rate of truth and falsity depends on how the *relevant* population of theories is fixed. This is where many philosophical problems loom large. For one, we don't know how exactly we should individuate and count theories. For another, we don't even have, strictly speaking, outright true and false theories. But suppose that we leave all this to one side. A more intractable problem concerns the concept of success. What is it for a theory to be successful? There is no reason here to repeat well-known points (see my 1999, pp. 104-8). But the general idea is clear. By choosing a loose notion of success, the size of the relevant population might increase and a lot of false theories might creep in. True theories won't be left out, but they may be vastly outnumbered by false ones. There will be many more false positives than otherwise. In this population, the probability of a randomly selected theory being true will be low. By choosing a stricter notion of success, e.g., by focusing on *novel* predictions, fewer theories will be admitted into the relevant population. The number of true theories will exceed the number of false theories. The number of false positives will be low, too. In that population, the probability of a randomly selected theory being true will be high. In sum, base-rates are unavailable not because we don't have enough statistics, but because we don't have clear and unambiguous reference classes. And we don't have the latter because our central individuating concepts (theory, success, etc.) are not precise enough.<sup>12</sup>

<sup>12</sup>In connection with the base-rate fallacy, L. J. Cohen (1981) has made the general point that there is no such thing as *the* relevant base-rate.

I want to add one more reason why I think that Howson's reformulation of NMA as a probabilistic argument is *deeply* problematic: it fails to capture the rich structure of theory-change in science. Recall the Pessimistic Induction. Laudan (1984) has invited us to see that if the history of science is the waste-land of aborted 'best theoretical explanations' of the evidence, it might well be that current best explanatory theories might take the route to this waste-land in due course.<sup>13</sup> In response to this argument, realists (cf. Kitcher 1993; Psillos 1999) have argued that theory-change is not as radical and discontinuous as the opponents of scientific realism have suggested. They have aimed to show that there are ways to identify the theoretical constituents of abandoned scientific theories which essentially contributed to their successes, separate them from others that were 'idle'—or as Kitcher has put it, merely 'presuppositional posits'—and demonstrate that those components which made essential contributions to the theory's empirical success were those that were retained in subsequent theories of the same domain. What follows from the relevant realist arguments is this: the fact that our current best theories may well be replaced by others does not, necessarily, undermine scientific realism. All it shows is that a) we cannot get at the truth all at once; and b) our judgements from empirical support to approximate truth should be more refined and cautious in that they should only commit us to the theoretical constituents that do enjoy evidential support and contribute to the empirical successes of the theory. Realists ground their epistemic optimism on the fact that newer theories incorporate many theoretical constituents of their superseded predecessors, especially those constituents that have led to empirical successes. The substantive continuity in theory-change suggests that a rather stable network of theoretical principles and explanatory hypotheses has emerged, which has survived revolutionary changes, and has become part and parcel of our evolving scientific image of the world. I think it is obvious that this rich structure cannot be captured by Howson's reformulations of NMA. In fact, it is not clear at all in what sense we can talk about base-rates of truth and falsity any more. The static picture of some percentages of true and false theories is replaced by a *dynamic* one, according to which theories improve on their predecessors, explain their successes, incorporate their well-supported constituents and lead to a truer description of the deep structure of the world.

These considerations make me very sceptical about the prospects of even starting to formulate the no-miracles argument as a probabilistic argument in the first place. It makes me even more sceptical about the cogency of the Bayesian charge that realists *ignore* base-rate information. But suppose that there are base-rates available. Is it always a bad idea to ignore them?

To address this question, let us go back to the original setting of the base-rate fallacy and take a look at another standard case in which this fallacy is to be committed. This is the Blue Cab/Green Cab case.

<sup>13</sup> It might be ironic that Lewis (2001) argues that the pessimistic induction is fallacious because it commits the base-rate fallacy.

(*Blue cab/Green cab*)

There is a city in which there are two cab companies, the Green cabs and the Blue cabs. Of the total number of cabs in the city, 85% are green and 15% are blue. There was a late-night hit-and-run car accident and the sole eyewitness said that it was a blue cab involved. The eye-witness is very reliable: in test situations involving blue and green objects at night, he made the correct identifications in 80% of the cases and he was mistaken in 20% of cases. What is the probability that the culprit was a blue cab?

When asked the foregoing question, subjects involved in psychological experiments, tended to trust the eyewitness and said, in an overwhelming percentage, that the probability that the culprit was a *blue* cab was very high. This is supposed to be a standard case of the base-rate fallacy, since, given the base-rates for blue and green cabs, the probability that the culprit was a blue cab is low (.41). It's more likely that the culprit was a green cab, since there are many more of those around.

There are two points that need to be noted. *First*, it is one thing to reason correctly probabilistically (the subjects, obviously, didn't). It is quite another thing to get at the truth. For, it may well be that the eyewitness really *saw* a blue cab and that a blue cab was involved in the accident. Unlikely things do happen, and we should be able to identify them no less than we are able to form a belief about what it is likely to happen and what it is not. What is important here is that the base-rate information might have to be ignored, if what we want to get at is the *truth*. There is not, of course, any definite answer to the question: when are the base-rates to be ignored and when are not? But there is an interesting observation to be made. In the case at hand, there is some crucial information to be taken into account, *viz.*, that the situation is *ambiguous*. After all, it was dark and, in the dark, our observations are not very reliable. Actually, as Birnbaum (1983) has noted, if a witness is aware that there are many more green cabs than blue cabs in the city, he is predisposed to see green cabs in ambiguous situations. This, it should be noted, is a piece of information (or background knowledge) that the subjects of the experiment also have. So, the very fact that, despite the prevailing disposition, the witness is reported to have seen a *blue* cab carries *more* weight than the relevant base-rates. So, there is a sense in which the subjects commit a fallacy (since they are asked to reason probabilistically but fail to take account of the base-rates), but there is another sense in which they reason correctly because the salient features of the *case history* can get them closer to the truth.

Transpose all this to the problem of truth and success. If we take the base-rates into account, we may get at the correct probability of a theory's (chosen at *random*) being approximately true, given that it is successful. And this probability may be quite low, if the base-rate of truth is very low. Suppose we conclude from this that this theory is *not* approximately true (because it is very unlikely that it is). But it may well *be* approximately true. The fact that it appears unlikely to be approximately true is not due to the fact that the theory fails to approximately fit with its domain, but rather due to the fact that the very few approximately true theories are swamped by the very many plainly false, but successful. If the theory *is*

approximately true, but—due to the correct probabilistic reasoning—we don't believe so, our beliefs will have been led away from the truth. In fact, we may reason as above. Suppose we grant the prevalence of false theories among the successful ones. Then, one might well be predisposed to say that a theory T is false, given its success. When, then, the eyewitnesses (the scientists, in this case) say that a specific theory T is approximately true (despite that this is unlikely, given the base-rates), they should be trusted—at the expense of the base-rates.

The *second* point can be motivated by a certain modification of the Green cab/Blue cab example. The situation is as above, with the following difference: the subjects are told that 85% of the car accidents are *caused* by blue cabs and 15% by green cabs. In these circumstances, the subjects did use the base-rates in their reasoning concerning the probability that the culprit was a blue cab (see Koehler 1996, p. 10). It is easy to see why they did: they thought that the base-rate information, viz., that blue cabs cause accidents much more often than green cabs, was *causally relevant* to the issue at hand. What needs to be emphasised is that in cases such as these there is an *explanation* as to why the base-rate information is relied upon. It's not just because the subjects want to get the probabilities right. It is also because this causally relevant information has a better chance to lead them to true beliefs.

Transpose this case to the problem of truth and success. Suppose that there is indeed a high base-rate for false theories. This would be relevant information if it were indicative (or explanatory) of success. If falsity did explain success, then, clearly, the small base-rate for truth would undermine belief in a connection between success and approximate truth. But falsity does *not* explain success. What is more, among the false theories some will be successful and some will be unsuccessful. In fact, it is expected that from a population of false theories (shall we say of all possible false theories?), most of them will be *unsuccessful*, while some will be successful. In terms of percentages, it might well be a bit of a fluke that some false theories are successful. The likelihood  $\text{prob}(S/-T)$  will be low. In fact, it can be so low as to dominate over the high base-rate of false theories. So, suppose that  $\text{prob}(S/-T) = .05$ ,  $\text{prob}(-T) = .9$  and  $\text{prob}(S) = .99$ . Then,  $\text{prob}(-T/S)$  is .045. A false theory would get no credit at all from success. Conversely, even if the base-rate of truth is low, there is an explanation as to why true theories are successful.<sup>14</sup> This might well be enough to show why, despite the low base-rate, a certain successful theory may well be deemed approximately true. Its posterior probability may be low,

<sup>14</sup> There is a worry here, voiced by Levin (1984), viz., that the truth of the theory does not explain its success. He asks: '[w]hat kind of *mechanism* is truth? How does the truth of a theory bring about, cause or create, its issuance of successful predictions? Here, I think, we are stumped. Truth (...) has nothing to do with it' (1984, p. 126). Musgrave (1999, pp. 68-9) has answered this worry very effectively. What does the explaining is the theory. But, Musgrave adds: 'Semantic *ascent* being what it is, we do not have *rival* explanations here, but rather equivalent formulations of the *same* explanation. "*H* believed that *G* and *G*" is equivalent to "*H* believed *truly* that *G*" (given the theory of truth that Levin and the realists both accept' (1999, p. 69). He then goes on to claim, correctly I think, that the explanation of the success of an action in terms of the *truth* of the agent's relevant beliefs is a mechanical or causal explanation.

but this will be attributed to the rareness of truth and not to any fault of the individual theory.

Here is another reason why it is, at least occasionally, right to ignore the base-rates. To motivate it, consider again the original Green cab/Blue cab case. As above, 85% of the cabs belong to the Green cab company and 15% to the Blue cab one. Imagine that people involved in car accidents are set on taking the cab companies to court. Suppose that on each occasion of the lawsuit, the court takes account of the base-rates and concludes that the cab was green, despite the fact that the eye-witness testified otherwise. Let's say that the court judges that it is always more likely (given the base-rates) that the cab was green (recall that the probability of the cab being blue is .41) and hence it decides to press charges against the Green cab company.<sup>15</sup> If courts acted like that, then the Green company would pay in 100% of such cases, whereas its cabs were responsible for only 59% of such accidents. Fairness and justice seem to give us *some* reason to ignore the base-rates!<sup>16</sup>

If we transpose this to the problem of truth and success, the moral should be quite clear. If scientists acted as the imagined judges above, they would be unfair and unjust to their own theories. If, as it happened, the base-rate of false theories were much higher than the base-rate of true ones, they would deem false theories that were true. Conversely, if the base-rate of true theories were much higher than the base-rate of false ones, they would deem true theories that were false.<sup>17</sup>

## 8. TAKING ACCOUNT OF CASE HISTORIES

If we leave base-rates behind us, what is left? There are always the case histories to look into. Though, as we saw in section 3, it does make sense to raise the grand question 'why is science successful (as an enterprise) as opposed to paradigmatically unsuccessful?', what really matters is the *particular* successes of individual theories, e.g., the discovering of the structure of the DNA molecules, or the explanation of the anomalous perihelion of Mercury. Now, if we think of it, it does *not* matter for the truth of the double helix model that truth is hard to get. The base-rate of truth (or of falsity)—even if we can make sense of it—is outweighed by the case history. We have lots of detail information about the DNA-molecule case to convince us that the double helix model is approximately true, even if, were we to factor in the base-rate of true theories, the probability of this model being approximately true would be very low. We are right in this case to ignore the base-rate, precisely because we know that this model's being approximately true does *not* depend on how many other true or false theories are around.

This last observation seems to me quite critical. The approximate truth of each and every theory will *not* be affected by the number (or the presence) of other

<sup>15</sup> If probability .59 is too low to capture the court's call that the case should be proven 'beyond reasonable doubt', then we can alter the numbers a bit so that the probability that the cab was green is high enough.

<sup>16</sup> A similar point is made by Windschitl and Wells (1996, 41).

<sup>17</sup> The base-rate fallacy has been subjected to very detailed and informative scrutiny by Jonathan Koehler (1996).

theories (even more so if these are independent of the given theory). Approximate truth, after all, is a relation between the theory and its domain (a relation of approximate fit). This relation is independent of what other (true or false) theories are available. In fact, we can see that there is an ambiguity in the probabilistic formulations of NMA. Though I have hinted at this above, it is now time to make it explicit.

There are *two* ways to think of arguments such as (A) and (B). The first is to apply the argument to a *specific* theory T (say, the electron theory, or Newtonian mechanics or the special theory of relativity). Then we ask the question: how likely is this *specific* theory T to be true, given that it has been successful? The second way is to apply the argument to an *arbitrary* theory T. Then we ask the question: how likely is an *arbitrary* (randomly selected) theory T to be true, given that it has been successful? If the issue is posed according to this second way, then it does follow from Bayes's theorem that the probability of a theory's being approximately true will depend on (and vary with) the base-rate of true theories. But if the issue is raised for a specific theory, then base-rates have no bite at all. Even if we had the base-rates, there are good reasons to neglect them—and scientists do neglect them—when the case history offers abundant information about the approximate truth of a given theory.<sup>18</sup>

## 9. LIKELIHOODISM

We are not done yet. The subjective Bayesian might now come back with a vengeance. He might say: ditch the base-rates, and go for purely subjective estimates of how likely it is that a theory is true. Consider what Howson (2000, p. 58) says: '[F]ar from showing that we can ignore even possibly highly subjective estimates of prior probabilities, the consideration of these quantities is *indispensable* if we are to avoid fallacious reasoning.' So, can we do away with priors altogether? Let us recall the Bayes factor from section 6. As Kevin Korb (forthcoming, p. 4) has argued, this factor reports the 'normative *impact* of the evidence on the posterior probability, rather than the posterior probability itself'. To get the posterior probability, we also need the prior. If the Bayes factor  $f = \text{prob}(S/-T)/\text{prob}(S/T) = 1$ , then  $\text{prob}(S/-T) = \text{prob}(S/T)$ , that is, the success of a theory makes no difference to its truth or falsity. But, the further from unity  $f$  is, the greater is the *impact* of the evidence. If  $f = 0$ , as we saw in section 6, then  $\text{prob}(T/S) = 1$ . And if  $f$  tends to infinity, then, given that  $\text{prob}(T) > 0$ ,  $\text{prob}(T/S)$  tends to 0. Given all this, it seems that we can reformulate Howson's NMA ((B<sup>1</sup>) in section 5) in such a way that it avoids base-rates (prior probabilities). The idea is that NMA need not tell us how probable a

<sup>18</sup> I don't want to deny that high probability is sufficient for warranted belief. But is it necessary? I don't think so. One of the prime messages of the statistical relevance model of explanation is that *increase* in probability does count for warranted belief. Now, empirical success does increase the probability of a theory's being approximately true, even with a low base-rate for truth. This can be easily seen by looking again at the example which preceded argument (B<sup>1</sup>) in section 5. There, the prior probability  $\text{prob}(T)$  of T was 1% but the posterior probability  $\text{prob}(T/S)$  rose to 17%. So, success does make a difference to the probability of theory's being true.

theory is, given the evidence (or its success). Rather, it tells us what the impact of the evidence (or the success) is on the posterior probability of the theory (without assuming that there is need to *specify* this posterior probability, and hence need to rely on a prior probability).

(B<sup>4</sup>)

1.  $f$  is close to zero (i.e.,  $\text{prob}(S/-T)$  is close to zero and  $\text{prob}(S/T)$  is close to 1).
2. S is the case.
3. Therefore, the impact of S on  $\text{prob}(T/S)$  is greater than its impact on  $\text{prob}(-T/S)$ .

(B<sup>4</sup>) can be supplemented with some specification of prior probabilities and hence it can yield a concrete posterior probability. Thus, it can then become either (B<sup>2</sup>) or (B<sup>3</sup>) above. But, even as it stands, it is suitable for *modest* Bayesians, who just want to capture the comparative impact of the evidence on competing hypotheses.

But we should also take a look at what has been called ‘likelihoodism’ (Sober 2002, p. 24). As Sober (2002) understands it, likelihoodism is a modest philosophical view. It does not aim to capture all epistemic concepts. It uses the *likelihood ratio* to capture the strength by which the evidence supports a hypothesis over another, but it does not issue in judgements as to what the probability of a hypothesis in light of the evidence is. In particular, likelihoodism does not require the determination of prior probabilities. So, it does not tell us what to believe or which hypothesis is probably true. Given two hypotheses  $H_1$  and  $H_2$ , and evidence  $e$ , likelihoodism tells us that  $e$  supports  $H_1$  more than  $H_2$  if  $\text{prob}(e/H_1) > \text{prob}(e/H_2)$ . The likelihood ratio  $\text{prob}(e/H_1)/\text{prob}(e/H_2)$  is said to capture the strength of the evidence.

Note that the likelihood ratio  $f^* = \text{prob}(e/H_1)/\text{prob}(e/H_2)$  is the converse of the Bayes factor  $f$ , as defined above. So likelihoodists can adopt a variant of (B<sup>4</sup>):

(B<sup>5</sup>)

1.  $f^*$  is greater than one (i.e.,  $\text{prob}(S/T)$  is close to 1 and  $\text{prob}(S/-T)$  is close to zero).
2. S is the case.
3. Therefore, S supports T over -T.

It is not my aim here to defend either (B<sup>4</sup>) or (B<sup>5</sup>). But it should be stressed that if we have in mind a more modest version of NMA, that is, that success tells more strongly in favour of truth than of falsity, then we can take (B<sup>4</sup>) as a version of NMA suitable for modest Bayesians and (B<sup>5</sup>) as a version of NMA suitable for non-Bayesians.<sup>19</sup>

<sup>19</sup> For a critique of likelihoodism, see Achinstein (2001, pp. 125-131).

## 10. CONCLUDING THOUGHTS

The moral of sections 3 and 4 is that there is no reason to think of the Ultimate Argument for realism as a deductive argument, contrary to what Musgrave suggests. The moral of sections 5 to 8 is that we should also resist the temptation to cast the no-miracles argument in a(n) (immodest) subjective Bayesian form. Once we free ourselves from both deductivism and subjective Bayesianism, there is no reason to think that NMA is either deductively or inductively fallacious. Many will remain unpersuaded. Both deductivism and Bayesianism are all-encompassing (shall I say imperialistic?) approaches to reasoning and they have many attractions (and a number of well-known successes). In fact, they share a common central theme: reasoning has a certain formal structure (given by deductive rules and Bayes's theorem—or better, Bayesian conditionalisation). So the substantive assumptions that are employed in reasoning have to do either with the truth of the premises (in deductivism) or with the prior probabilities (in Bayesianism). But perhaps, the simplicity of both schemes of reasoning is their major weakness. Reasoning is much more complex than either of them admits.

So, what sort of argument is the Ultimate Argument for realism? I know of no more informative answer than this: it is an inference to the best explanation (IBE). And what kind of inference is IBE? I know of no more informative answer than this: it is the kind of inference which authorises the acceptance of a hypothesis H as true, on the basis that it is the best explanation of the evidence. The rationale for IBE is that explanatory considerations should inform (perhaps, determine) what is reasonable to believe. I know all this is too crude to count as an explication. Further explication can be given, as I tried to show in my (2002). In any case, even if the Ultimate Argument for realism were to be found wanting as an *explanatory* argument, it would still be the case that the realist explanation of the success of science remains the best. Musgrave's 'The Ultimate Argument for Realism' is to be credited for making a very compelling—perhaps unparalleled—case for this.

## ACKNOWLEDGEMENT

Earlier versions of this paper were presented in seminars in the Universities of Stockholm, Oslo, Gothenburg and Rotterdam and in a meeting of the Belgian Society for the Philosophy of Science. I want to thank all participants in these events for very useful comments and criticism, and in particular Janneke van Lith (who acted as a commentator in the Rotterdam meeting), Igor Douven, Michel Ghins, Olav Gjelsvik, Leon Horsten, Uskali Maki, Oystein Linnebo, Nils Roll-Hansen, Dag Westerstaal and Asa Wikforss. Research for this paper was funded by the framework EPEAEK II in the programme Pythagoras II.

## REFERENCES

- Achinstein, P. (2001) *The Book of Evidence*. New York: Oxford University Press.  
 Bar-Hillel, M. (1980) 'The Base-Rate Fallacy in Probability Judgements.' *Acta Psychologica* 44: 211-33.

- Birnbaum, M. H. (1983) 'Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem.' *American Journal of Psychology* 96: 85-94.
- Cohen, L. J. (1981) 'Can Human Irrationality be Experimentally Demonstrated?' *Behavioural and Brain Sciences* 4: 317-31.
- Howson, C. (2000) *Hume's Problem*. New York: Oxford University Press.
- Koehler, J. J. (1996) 'The Base Rate Fallacy Reconsidered: Descriptive, Normative and Methodological Challenges.' *Behavioural and Brain Sciences* 19: 1-17.
- Korb, K. (forthcoming) 'Bayesian Informal Logic and Fallacy.' Melbourne: School of Computer Science and Software Engineering, Monash University, Technical Report 2002/120, 22pp.
- Laudan, L. (1984) *Science and Values*. Berkeley: University of California Press.
- Levin, M. (1984) 'What Kind of Explanation is Truth?' in Leplin, J. (ed.) *Scientific Realism*. Berkeley and Los Angeles: University of California Press.
- Lewis, P. (2001) 'Why the Pessimistic Induction is a Fallacy.' *Synthese* 129: 371-80.
- Musgrave, A. (1977) 'Explanation, Description and Scientific Realism.' *Scientia* 112: 99-127.
- Musgrave, A. (1980) 'Wittgensteinian Instrumentalism.' *Theoria* 46: 65-105.
- Musgrave, A. (1988) 'The Ultimate Argument for Scientific Realism.' in Nola, R. (ed.) *Relativism and Realism in Science*. Dordrecht/Boston: Kluwer.
- Musgrave, A. (1989) 'NOA's Ark—Fine for Realism.' *The Philosophical Quarterly* 39: 383-398.
- Musgrave, A. (1996) 'Realism, Truth and Objectivity.' in R. S. Cohen *et al.* (eds) *Realism and Anti-Realism in the Philosophy of Science*. Dordrecht: Kluwer.
- Musgrave, A. (1999) *Essays on Realism and Rationalism*. Amsterdam: Rodopi.
- Musgrave, A. (1999a) 'How to do without Inductive Logic.' *Science & Education* 8: 395-412.
- Norton, J. (1994) 'Science and Certainty.' *Synthese* 99: 3-22.
- Psillos, S. (1999) *Scientific Realism: How Science Tracks Truth*. London: Routledge.
- Psillos, S. (2000) 'The Present State of the Scientific Realism Debate.' *The British Journal for the Philosophy of Science* 51: 705-28—reprinted in Clark, P. & Hawley, K. (eds) (2003) *Philosophy of Science Today*. Oxford: Clarendon Press.
- Psillos, S. (2002) 'Simply the Best: a case for Abduction.' in Sadri, F. & Kakas, A. (eds), *Computational Logic: From Logic Programming into the Future*. LNAI 2408, Berlin-Heidelberg: Springer-Verlag.
- Psillos, S. (2006) 'Putting a Bridle on Irrationality: an Appraisal of van Fraassen's new epistemology.' in B. Monton (ed.) *The Many Faces of Empiricism*, Oxford: Oxford University Press.
- Putnam, H. (1975) *Mathematics, Matter and Method*. Cambridge: Cambridge University Press.
- Sober, E. (2002) 'Bayesianism—its Scope and Limits.' *Proceedings of the British Academy* 113: 21-38.
- Tversky, A. & Kahneman, D. (1982) 'Judgment under Uncertainty: Heuristics and Biases' in D. Kahneman, P. Slovic & A. Tversky (eds) *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- van Fraassen, B. C. (1980) *The Scientific Image*. Oxford: Clarendon Press.
- Windschitl, P. D. & Wells, G. L. (1996) 'Base Rates do not Constrain Nonprobability Judgments.' *Behavioural and Brain Sciences* 19: 40-1.