The Epistemic Advantage of Prediction over Accommodation

Roger White

According to the thesis of Strong Predictionism, we typically have stronger evidence for a theory if it was used to predict certain data, than if it was deliberately constructed to accommodate those same data, even if we fully grasp the theory and all the evidence on which it was based. This thesis faces powerful objections and the existing arguments in support of it are seriously flawed. I offer a new defence of Strong Predictionism which overcomes the objections and provides a deeper understanding of the epistemic importance of prediction. I conclude by applying this account to strategies for defending scientific realism.

Should we have more confidence in a theory if it correctly *predicted* a certain datum, than if it was merely designed to *accommodate* that datum? Many philosophers have thought so, but have had difficulty explaining why, and defending their claim against powerful objections. They have often reasoned roughly as follows. We are rightly impressed by theories which not only fit the existing data, but lead to predictions which are later confirmed. After all, it is not hard to cook up a false theory to account for known facts. But that our theory makes successful novel predictions seems to indicate that it is true, since it is unlikely that a false theory would be so successful.[1]

Recently however, a growing number of philosophers have argued that this alleged epistemic difference is bogus.[2] Their reasons are often roughly as follows. What more could be relevant to assessing the truth of a theory than the content of the theory and the data (and auxiliary assumptions, and background theory), and the relations between them? The *order* in which the theory was constructed and the data discovered, and even the *motivations* of the theorist (whether the theory was constructed with the data in mind) seem beside the point. Indeed it should make no difference whether the theory was constructed at all, or

[1] Philosophers on this side of the debate include Peirce (1931–58), Whewell (1860), Duhem (1954), Geire (1983), Maher (1988), and Worrall (1989).

[2] They include Mill (1843), Keynes (1921), Horwich (1982), Schlesinger (1987), Howson and Franklin (1991), Achinstein (1994), and Collins (1994).

just fell out of the sky. To assess its truth we must simply consider its inherent plausibility and how it fits with all the evidence we have.

The issue is not only important in itself, but is connected to a number of prominent issues in epistemology and philosophy of science. For instance, one central argument for scientific realism claims that the predictive success of scientific theories in general is significant evidence for their truth.[3]

I have much sympathy with the latter camp, who downplay or deny the epistemic significance of prediction versus accommodation. Nevertheless, I will make a case for a version of *predictionism*, the view that in a wide class of circumstances, the fact that a theory predicted, rather than accommodated a certain datum, provides support for the theory, beyond that provided by the datum itself. I will give an explanation of why prediction has an epistemic advantage over accommodation, an explanation which allows us to see which factors govern the degree of this advantage and the circumstances in which it holds. While supporting a version of predictionism, this deeper understanding of when and how prediction matters may actually lead us to downplay its significance.

I will begin by presenting what I take to be the most powerful argument against predictionism, followed by an examination of the most common defence of predictionism, and why it does not work. My defence of predictionism will be in the same spirit as the standard one, but overcomes the anti-predictionist objections. I will conclude with some suggested applications of this discussion to debates over scientific realism.

1. Clarification of the issues

First let me clarify my use of the expression 'the datum'. We will be concerned with cases in which a theory T entails a certain proposition which, either before or after the construction of T, is discovered to be true. But the mere fact that T entails a known proposition is not remarkable by itself, since all theories entail known truths. For instance, T entails the disjunction (T or P), for any known proposition P (and (T or P) can be known if P is known). So we need some restriction on which entailed truths are relevant to confirmation. I will not address

[3] Collins (1994) mentions a number of related issues, including Lakatos's account of scientific methodology, according to which one research programme can supersede another, only if it predicts new, *unforeseen* phenomena, Popper's view of science as a form of knowledge superior to other explanatory enterprises such as history or psychoanalysis, and the legitimacy of the distinction, advocated by the positivists, between the 'logic of discovery' and the 'logic of justification'.

the interesting problem of giving a general account of the conditions in which entailment of a truth counts as evidence for a theory. For our purposes we can understand entailment of data as relative to a certain *experiment*, and corresponding class of mutually exclusive possible outcomes. Relative to experiment *E*, 'the datum', refers to that proposition which specifies the unique actual outcome of *E*.

Second, we should be clear on just what the prediction/accommodation distinction is. In a typical case of successful prediction, a theory is first constructed, then tested by deriving some of its consequences, which are later discovered to be true. In a case of accommodation the datum is already known before the theory is constructed. This might suggest that the crucial distinction concerns the *temporal order* of theory construction and data discovery. But while some discussions have focused on this distinction, it seems that what really matters is not temporal order, but a causal relation. Intuitively, a theory is less well confirmed if it was *designed* to entail the datum, that is, the condition of entailing that datum acted as a constraint on the construction of the theory.[4] Of course the reason why a theory was not designed to entail a datum is usually that this datum was not known at the time. But if it was known, yet the theory was not constructed with this datum in mind, it seems that it should support the theory in the same way and to the same extent as it would have had it not been discovered until after the theory was constructed. The following definitions capture the distinction which matters here.

> A theory *T accommodated D* iff *T* entails *D*, *D* is true, and *T* was *designed* to entail *D*, (that is, the condition that the theory entail *D* acted as a constraint on the selection of *T* as the accepted theory).[5]
>
> *T correctly predicted D*, iff *T* entails *D*, *D* is true, and *T* was *not* designed to entail *D*.

We can now state the question which concerns us: In what circumstances, if any, should the information that *T* correctly predicted, rather than accommodated *D*, give us greater confidence in *T*? It is useful to distinguish a weak and a strong version of predictionism:

[4] This account of accommodation is close to what Zahar (1973) and Worrall (1985) call lack of *heuristic novelty*.

[5] Of course, a theory rarely entails any specific experimental data on its own, but only in conjunction with a set of auxiliary assumptions and background theory. So entailment here should be understood as entailment relative to a set of background assumptions. In comparisons between cases of prediction and accommodation, these background assumptions should be kept fixed.

> **Weak Predictionism**: That *T* correctly predicted rather than accommodated D, typically provides further evidence for *T*, if we are ignorant of either the content of *T* or the independent evidence that supports it.
>
> **Strong Predictionism**: That *T* correctly predicted rather than accommodated D, typically provides further evidence for *T*, even if we are familiar with the content of *T* and the independent evidence that supports it.

The weak thesis is not controversial. It is agreed on all sides that if we were to survey all the actual theories that have been proposed, we should expect to find that on average, those theories from which successful predictions had been made would be better supported by the total evidence, than those which have merely accommodated existing data. There are at least a couple of reasons for this. First, as Keynes (1921) pointed out, as a matter of practice, rarely is a theory tested by deducing its consequences unless it already has evidential support, whereas a theory will often be proposed to accommodate existing data, even if it has little or no independent support. Second, as Lipton (1991) argues, the accommodation of data often results in a clumsy, *ad hoc*, and hence less *simple* theory, one which gives a less unified account of the total evidence, especially in the case where an existing theory is modified to account for new data. So theories which successfully predict data tend to be more plausible, all things considered, than those that merely accommodate data, by virtue of their greater simplicity.

As a consequence, information as to whether a datum was predicted rather than accommodated by the theory can rationally increase our confidence in the theory, at least in so far as we are ignorant of its degree of simplicity and the additional evidence that supports it. For in this case, learning that the datum was predicted should increase our confidence in the theory by virtue of increasing our estimate of the theory's simplicity and independent evidential support. So Weak Predictionism seems clearly correct. But in a situation where we are thoroughly familiar with what the theory says (and hence its degree of simplicity), and all the evidence supporting it, we can assess its simplicity and evidential support *first hand*, and it seems that our judgement should not be affected upon learning that the datum was predicted or accommodated. So Strong Predictionism is not supported by considerations of simplicity and independent evidence.[6]

[6] Or is it? Lipton (1991) makes an interesting case for Strong Predictionism which hinges on simplicity.

Most of the debate over prediction versus accommodation has focused on the strong thesis. This thesis is open to serious objections, and its opponents often suspect that its popularity stems from a confusion with cases which only support Weak Predictionism. My account will both support Strong Predictionism and provide a new understanding of why both the strong and weak theses hold, one which goes beyond the standard explanations in the literature. First I will consider what the defender of Strong Predictionism is up against.

2. The anti-predictionist challenge

There are a number of reasons why the strong predictionist thesis seems highly dubious, some involving analyses of individual cases, and others involving general arguments. First, many of the cases which might be taken to illustrate the epistemic advantage of prediction are either historically dubious,[7] or can be diagnosed as involving some other factor, such as simplicity, which makes the epistemic difference. When we are careful to construct a case which eliminates these other differences, we often find that the epistemic advantage of prediction seems to disappear. Suppose I watch a coin being tossed fifty times, landing heads every time. After five heads, I tentatively form the hypothesis that the coin is double-headed, and correctly predict the remaining outcomes. You, on the other hand, learn of the outcomes after the sequence is completed, and similarly conclude that the coin is double-headed. Surely I have no more reason to believe this hypothesis than you, just because I made an early prediction. This case eliminates various independent features which can make an epistemic difference (we both fully grasp the theory and its evidence). It is tempting to generalize to the view that whenever we fully grasp the theory and all the evidence, whether the datum was predicted or accommodated makes no epistemic difference. Second, a rather compelling case can be made against Strong Predictionism, as we will see presently. And third, the standard and initially compelling predictionist argument turns out to be flawed on closer inspection; I will examine this argument in sections 3–4.

The following argument, based on Collins (1994), brings out just how implausible the strong predictionist thesis can seem. Suppose we know that D is true and that T entails D. We are also thoroughly famil-

[7] For example, Worrall (1989) challenges Geire's (1983) historical account of Fresnel's light diffraction predictions and Brush (1994) challenges Maher's (1988) and Lipton's (1991) accounts of Mendeleev's prediction of the elements, among many other cases.

iar with the content of *T* and all the independent evidence supporting it. Our question is whether upon learning that *T* predicted rather than accommodated *D*, we should revise our confidence in *T*. Note that the difference between accommodation and prediction consists simply in the occurrence of a certain psychological process in those who developed the theory *T*, namely the process of designing the theory to entail *D*. The answer to our question now hinges on whether the information that this psychological process occurred, should have any affect on our confidence in *T*.

The problem is that there seems to be no plausible, non-mysterious way that the fact that this psychological process took place in the theorist's head could be epistemically relevant to the truth of her theory. One way in which one fact may provide evidence that another fact obtains, is when we have reason to suspect that there is some *causal* connection between the two facts. Could there be a causal connection between the truth of the theory *T* and the theorist's not having designed *T* to entail a certain datum? Let *T* be the theory of General Relativity, which, as it happens, Einstein did not design to entail the correct degree to which light bends around the sun (though it does in fact entail it). The theory of General Relativity is true just in case certain physical states of affairs obtain, such as that space-time is curved to the degree given by the field equations, and so on. But it can hardly be that the goings on in Einstein's head are causally responsible for the structure of space-time. Nor does there seem to be a causal connection in the other direction. It is certainly not the curvature of space-time which prevented Einstein from designing his theory to entail the correct degree of light bending.

Now of course a causal connection is not the only possible basis for an evidential connection between states of affairs. But in the present case it is hard to see what other kind of evidential connection there might be. The predictionist therefore faces the following challenge: explain how the fact that the psychological process of designing the theory *T* to entail the datum occurred, can, in some plausible and non-mysterious way, rationally affect our confidence in *T*.

3. The No-Coincidence Argument for Predictionism

The third reason that I suggested as to why the strong predictionist thesis seems dubious was that the strongest argument in its favour is flawed. Let's now turn to consider this argument, sometimes called the No-Coincidence Argument. The most common line of argument for

Strong Predictionism is some version of the following.[8] If our theory *T* *correctly predicted D*, a good explanation of this fact is that *T* is true, for the truth of *T* guarantees the success of its predictions such as *D*. But if *T* is false, then it is highly unlikely to correctly predict data that we later discover; we should have to say that its predictive success was a mere *coincidence*. The fact that the truth of *T* can explain its predictive success, which would otherwise be a striking coincidence, is significant evidence for T. However, if *T* merely *accommodates D*, we do not need to invoke the truth of *T* to explain this fact. For if *T* was *designed* to entail *D*, it is no surprise that it does so, regardless of whether *T* is true or false. So when we know that *T* merely accommodated *D*, it does not gain this extra support.

The point is sometimes put in terms of two competing explanations for the fact that *T* entails the datum: (i) the *truth* hypothesis—that *T* is true, and (ii) the *design* hypothesis—that *T* was designed to entail the datum. If *T* predicted *D*, then the truth hypothesis is the only explanatory option and hence is confirmed. But if *T* merely accommodated *D*, the design hypothesis is sufficient to explain the fact that *T* entails the datum, and hence it renders the truth hypothesis otiose. Hence *T* is better supported over all, given that it predicted rather than accommodated the data.[9]

4. Problems with the No-Coincidence Argument

I will argue in this section that no version of the No-Coincidence Argument is successful. The argument involves a kind of inference to the best explanation, so it will pay us to examine just what the explanans and explanandum are. Two competing explanantia appear in the argument: theory *T*'s being true, and *T*'s having been designed to entail the datum. The explanandum has to do with the entailment relation between the theory and the datum. Unfortunately, precisely what the explanandum is taken to be varies among different versions of the argu-

[8] Versions can be found it Peirce (1931–51), Whewell (1860), Geire (1983) and Worrall (1989). Opponents of Strong Predictionism such as Keynes (1921), Horwich (1982), and Collins (1994) identify this as the major motivation for predictionism. My diagnosis of the argument differs from theirs.

[9] Another variation on the argument, found in Geire (1983) and Worrall (1989) is that prediction has an epistemic advantage because only in the case of prediction does the experiment whose outcome is specified by *D*, constitute a *good test* of *T*, that is, one which has a good chance of *falsifying T*. The underlying reasoning here is essentially the same. In a case of prediction, *T* is far more likely to pass the test if it is true, than if it is false, whereas in a case of accommodation *T* is guaranteed to pass the test. I will not discuss this version of the argument directly in what follows, but I believe that my criticisms apply equally.

ment, or is just left unclear, so we will have to survey a number of alternatives.

Taking our datum to be *D*, our first candidate for the explanandum is

(**E**) *T* entails *D*.

But this is a non-starter, since entailments are necessary; *T* would have entailed *D* regardless of the truth of *T*, or how it was 'designed', or anything else for that matter. Perhaps a more promising suggestion is

(**P**) *T* correctly predicted *D*.

This is at least a contingent fact, and hence open to explanation. Taking *P* as our explanandum, the predictionist argument proceeds as follows. In a case of prediction, *P* may be explained by the truth of *T*, and hence *P* confirms *T*'s truth. But in a case of accommodation, we have no such fact as *P* to explain; instead we have

(**A**) *T* accommodated *D*.

And we do not need the truth of *T* to explain A. Indeed A needs no explanation, since it is all too easy to get a theory to accommodate a datum, regardless of its truth. Hence in a case of prediction we have stronger confirmation for *T*.

But now note that the fact that *T* correctly predicted *D* is a conjunction of three facts: that *T* entails *D*, that *T* was not designed to entail *D*, and that *D* is true. We have just noted that *T*'s entailing *D* is not open to explanation at all. As for the fact that *T* was *not designed* to entail *D*, it seems rather implausible that this could be explained by the truth of T. For example, the fact that General Relativity is true, that is, that spacetime is curved and so on, does not explain the fact that Einstein did not design his theory to entail the datum that light bends around the sun. So it seems that the truth of *T* can explain the fact that *T* correctly predicted *D*, only by explaining *D*'s being true. This it may well do, for since *T* entails D, the truth of *T* guarantees the truth of D.

But precisely the same holds in the case where *T* merely *accommodated D*. T's accommodating *D* consists in the fact that *T* entails *D*, *T* was designed to entail *D*, and *D* is true. As with the case of prediction, the truth of *T* is irrelevant to the first two conjuncts, but entails the third. If the truth of *T* explains *T*'s correctly predicting *D*, by virtue of entailing that *D* is true, then it seems it should also explain *T*'s accommodating *D*, for the same reason. So this approach fails to bring out a difference between the weight of predicted and accommodated data.[10]

[10] This objection is loosely based on Collins (1994).

Our first suggestion failed because entailments hold necessarily between propositions, and '*T*' and '*D*' refer *rigidly* to certain propositions. Whatever proposition *D* is, *T* cannot help but entail *that very proposition*. But *T* need not have entailed *the datum*, where 'the datum' is taken to refer non-rigidly to whichever proposition describes the actual outcome of our experiment, or in other words, T *might not have been datum-entailing*. So we might take

(**DE**) *T* is datum-entailing

as a good candidate for the explanandum. The predictionist argument would then go as follows. In a case of prediction, the truth of *T* may explain DE, and hence be confirmed by DE. But in a case of accommodation, DE is adequately explained by *T*'s having been designed to entail the datum; there is no need for *T*'s truth in the explanation. And hence the inference to *T*'s truth is undermined.

Now *T*'s being true might well explain its being datum-entailing, since necessarily, the entailments of a true theory are true. The question is whether *T*'s being *designed* to entail the datum offers a rival explanation. Here the word 'design' can be misleading. We cannot design a theory to entail the datum, in the sense that we design a house to face the ocean, where that very house would not have faced the ocean had we not designed it to. A theory is a *proposition* which cannot be moulded into shape to fit the data; it has its truth-conditions and hence entailments essentially. A better metaphor for the process of theorizing is that of selecting a theory off the platonic library shelf. Theories already exist, and necessarily entail what they do, independently of our selection of one. To *modify* our current theory to fit a datum is really to *discard* it and select a slightly different one. To say that *T* was *designed* to entail the datum just means that *T* was selected under a certain restriction, namely that the chosen theory entail whatever the datum happens to be.

But now the fact that *T* was selected under this restriction does not help explain the fact that it meets the restriction, any more than Jane's *choosing* to buy a house that faces the ocean helps explain why *it* faces the ocean. That very house would have faced the ocean regardless of Jane's criteria in choosing a house. Similarly, *T* is datum-entailing just in case the possible experimental outcome which *T* necessarily entails, does in fact obtain. But of course which outcome obtains in no way depends on the theorist's method of theory selection (the degree of light bending could hardly be explained by the way that Einstein came up with General Relativity). So in the case of accommodation, *T*'s hav-

ing been designed to entail the datum does not serve as a rival to *T*'s truth, as an explanation of DE, for it does not serve as an explanation of that fact at all.

5. Introducing the role of the theorist

We have considered three candidates for the role of explanandum in the No-Coincidence Argument, all of which fail to make the argument work. The following suggestion goes beyond the standard No-Coincidence Argument, by focusing on the *theorist*. I will argue that it will not save the No-Coincidence Argument, yet it provides the basis for the successful argument presented in the next section.

Perhaps the temptation to suppose that *T*'s being designed to entail the datum explains the fact that it does entail the datum, is due to a confusion between this and another fact, namely, *that the theorist now holds a theory which entails the datum*. This might be explained by her theory selection having been restricted to datum-entailing theories, just as the fact that Jane now inhabits a house which faces the ocean can be explained by her having deliberately chosen one that does. So we should consider

(**ES**) The theorist selected a datum-entailing theory

as our explanandum. Let us call this fact *the theorist's entailment-success*. The trouble here is that our preferred explanans, namely *T*'s being true, does not explain ES. A concrete example makes this clear. That General Relativity is true, that is, that space-time is curved and so on, does not explain why Einstein came up with a theory which makes true predictions. Einstein's success had more to do with his epistemic relation to the facts, than with what those facts happened to be. It is tempting to suppose that *T*'s being true might help explain the theorist's entailment-success by helping explain why the theorist holds T, since *T*'s truth guarantees that it is datum-entailing. But *T*'s truth can help explain the theorist's holding of *T*, only if the theorist has some kind of propensity to hold true theories. But if so, then the fact that it is T which is true, is irrelevant to the theorist's entailment-success. What matters is just that she is reliably connected with the truth, that is, she will tend to accept the truth, regardless of whether the truth happens to be *T*. If, by contrast, the theorist's holding of *T* has no reliable connection with the facts, say, if it is just a wild guess, then the lucky fact that she holds a datum-entailing theory has nothing at all to do with the

truth of T. Either way, the truth of *T* is irrelevant when it comes to explaining the theorist's entailment-success.

I will canvas one last attempt to take *truth* to explain the theorist's entailment-success. Instead of taking the *truth of* T as our explanans, we might try the truth of the *theorist's theory*, where 'the theorist's theory' refers non-rigidly, or better, the fact that *the theorist holds a true theory*. This has the advantage that it does entail ES, that *the theorist holds a datum-entailing theory*. But it does not seem to explain it. That Jane owns a house facing the North Atlantic entails that she owns an ocean-facing house, but it does not *explain* it. The explanation must have to do with the *way* in which her house was chosen, for instance that she wanted a house facing the ocean and so tried hard to get one.

In any case, this does not help the predictionist's case unless we can explain why the inference to truth is undermined by the design hypothesis. If we do not know the location of Jane's house, our learning that it is ocean-facing supports the hypothesis that it faces the North Atlantic (not because either fact *explains* the other, but just because we have narrowed down the possibilities, and all houses facing the North Atlantic face the ocean). But it is not clear why the information that Jane chose her house on the condition that it face the ocean, diminishes this support. Similarly, the fact that the theorist holds a datum-entailing theory supports the hypothesis that she holds a true theory (not because one fact *explains* the other, but just because we have narrowed down the possibilities, and all true theories are datum-entailing). But now why should we suppose that the information that she chose her theory on the condition that it entails the datum, diminishes this support at all? It is not yet clear why it should.

6. A new argument for Strong Predictionism

Our attempts to save the standard No-Coincidence Argument have failed. I wish to present a new argument which is persuasive. We should still take

(**ES**) The theorist selected a datum-entailing theory

as our explanandum. As we have noted, this might be explained by the design hypothesis

(**DS**) The theorist designed her theory to entail the datum, that is, knowing the experimental outcome, she selected her theory on the condition that it entail this datum.

What other hypothesis might explain ES? We might try to explain it by supposing that she selected her theory on the condition that it was *true*, for this would guarantee that she selected a datum-entailing theory. But unfortunately, theories do not come with clear labels attached declaring their truth-value, so they cannot be straightforwardly selected by this criterion. Theory selection may, however, be more or less well *aimed at the truth*. This notion requires further analysis, but it might roughly be characterized as the degree to which the causal chain of mechanisms which led to her selection of the theory were reliably connected to the facts. Obviously this is a matter of degree, but for the sake of simplicity we can focus on the truth or falsity of the hypothesis

(**RA**) The theorist's selection of her theory was reliably aimed at the truth

by which I mean roughly that the mechanisms which led to her selection of a theory gave her a good chance of arriving at the truth. This hypothesis at least raises the theorist's chances of holding a datum-entailing theory, by raising her chances of holding a true theory.[11]

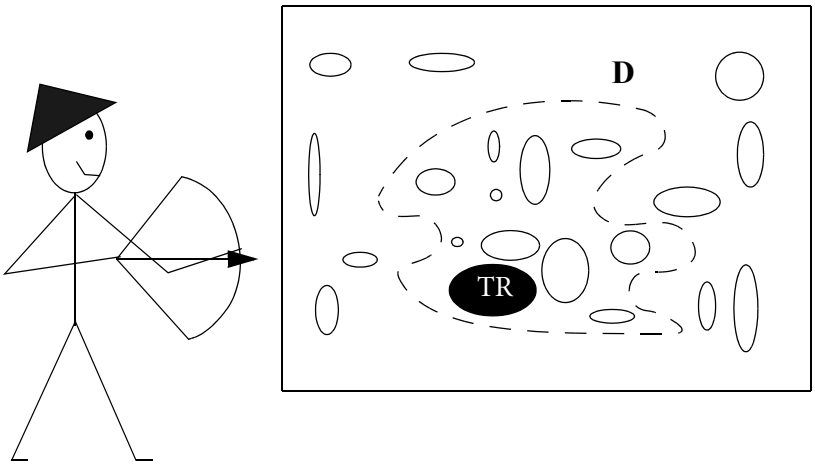
7. The archer analogy

The relations among ES, DS and RA can be illustrated by a simple analogy. We may represent our theories and data on a map of logical space, where regions on the map represent sets of possible worlds in which a proposition is true, the area of a region being proportional to the probability of the proposition. The dotted region D represents our datum, the outcome of a certain experiment. The small circular regions represent various theories. Only those which are sub-regions of D, entail the actual experimental outcome, and region TR, which contains the actual world, is the only true theory.

Now suppose that this map is drawn on the side of a barn and an archer shoots an arrow at it. We do not know if the archer is aiming at TR or even how good his aim is. Without seeing where the arrow landed, we learn that it landed within a circle in region D. The question which concerns us is whether the arrow landed in TR. The information that the arrow landed in D, should increase our confidence that the arrow landed in TR, since TR is contained within D, and D is a smaller region than the wall. (This is analogous to the way that learning that a theory entails the datum, can provide evidence that the theory is true,

[11] Maher (1988) seems to be onto a similar idea, but develops it along different lines. For criticisms of Maher's argument see Howson and Franklin (1991).

quite apart from whether the datum was predicted or accommodated).



But now consider how our confidence that the arrow landed in TR should differ depending on whether we make the following assumption.

(**DS***) The archer is reliably aiming at region D (he may or may not be aiming more specifically at TR).

Whether or not we know this to be the case, will affect whether

(**ES***) The arrow landed within D

supports

(**RA***) The archer was reliably aiming at TR.[12]

[12] DS* and RA* should be understood to be logically independent. DS* says simply that the archer, knowing the location of D, restricts his aim so that he is guaranteed to hit somewhere within D. He may or may not attempt to hit some more specific region such as TR. The denial of DS* is consistent with his reliably aiming at TR. Of course in one sense, if the archer is aiming at TR, he must also be aiming at D, since TR lies within D. But there is another sense—the one relevant to our discussion—according to which the archer may aim at TR without aiming at D, that is, without *intending to hit* D, if he does not even know where region D is, or at any rate, if his knowledge of the location of D has no influence on how he shoots. Similarly with DS and RA, the theorist may design her theory to entail the datum with or without also reliably aiming for a true theory. And she may reliably aim at the truth, without designing her theory to entail the datum, if she does not know the datum, or her knowledge of the datum plays no role in her selection of a theory.

Let's begin with the assumption that DS* is not true, indeed, let's suppose that the archer couldn't have aimed at D, since it isn't even marked on the map (he still may or may not have been aiming at TR). On this assumption, the fact ES*, that the arrow landed in D, lends some support to the hypothesis RA*, that the archer was reliably aiming at TR. For if he was reliably aiming at TR, he is more likely to hit it, and hence hit D, since TR is a sub-region of D. He is far less likely to hit TR if he was not reliably aiming at it, and hence less likely to hit D.

But on the assumption of DS*, he is guaranteed to hit D, regardless of whether he is aiming more specifically at TR. So given DS*, RA* does not render ES* more probable, and hence ES* provides no support for RA*. Hence the reliable aiming theory is better supported by ES*, on the assumption that the archer did *not* aim at region D. And this in turn renders it more probable that the arrow landed on TR, given that the archer did not aim at D.

We can summarize the reasoning here as follows. Upon learning that the arrow landed within D, we should increase our confidence that it landed on TR, since TR lies within D and we have narrowed down the region in which it might have landed. Upon further learning that the archer was not restricting his aim to regions within D, we have a *further* reason to suppose that it landed on TR. The fact that the arrow landed within D should increase our confidence that the archer was reliably aiming at TR (since his aiming at TR would make him more likely to hit within D), and hence increase our confidence that he hit TR.[13] If, on the other hand, we learn that the archer restricted his aim to regions within D, we have no grounds to further increase our confidence in his aim at TR, or his hitting TR. For in this case his hitting within D is no further indicator of his aim at TR (since he was bound to hit D, regardless of whether he was aiming more specifically at TR).

The analogy should be clear. We can think of the process of theory selection as like shooting an arrow at logical space, where we are uncertain as to how well the theorist is aiming at the truth, that is, the reliability of the process by which she selected her theory. That the theorist holds a datum-entailing theory ES, supports the reliable aim hypothesis RA, but only if we can rule out the design hypothesis DS. The analogy between DS* and DS is as follows. The archer, if he knows where region D is, can restrict his aim to circles within this region, with or without

[13] It is important here that TR is not just any sub-region of D, but a *salient* target, one which stands out by being painted black. We do not know if the archer is aiming at any small region or how good his aim is, but if he is aiming, he is somewhat more likely to aim at TR, since it stands out from the surrounding regions.

aiming more specifically at TR. Similarly, the theorist, if she knows the datum, can restrict her theory selection to theories which are datum-entailing, with or without aiming more specifically for the truth.

On the assumption of not-DS, the fact ES, supports the hypothesis RA. For the theorist is more likely to select a true theory, given RA, and a true theory is more likely to entail the datum than a false one. But ES does not support RA on the assumption DS. For on this assumption it is to be expected that she will select a datum-entailing theory, regardless of how well she was aiming at the truth. Hence the reliable aiming hypothesis RA, is better supported by ES, on the assumption that the theory was *not* designed to entail the datum. And this in turn renders it more probable that her theory is *true*, given that it was not designed to entail the datum.

As with the archery analogy, we can summarize the reasoning here as follows. Regardless of whether the theorist designed her theory to entail the datum, upon learning that her theory does entail the datum, we should increase our confidence in its truth, since necessarily, true theories are datum-entailing, and we have narrowed down the region of logical space in which the theory is contained. But now upon further learning that the theorist did not restrict her theory selection to datum-entailing theories, we have a *further* reason to suppose that she selected a true theory. The fact that a datum-entailing theory was selected, should increase our confidence that the theorist was reliably aiming at the truth (since her aiming at the truth would make her more likely to select a datum-entailing theory), and hence increase our confidence that she selected a true theory.[14] If, on the other hand, we learn that the theorist restricted her theory selection to datum-entailing theories, we have no grounds to increase further our confidence in her aim at the truth, or her selection of a true theory. For in this case her selection of a datum-entailing theory is no further indicator of her aim at the truth (since she was bound to come up with a datum-entailing theory, regardless of how well she was aiming at the truth).

7.1 Two objections[15]

Objection 1: Your argument might seem to involve a kind of illicit double counting of evidence, by using the evidence to provide direct induc-

[14] Here it is important that the truth is a *salient* target. We do not know if the process of theory selection was directed toward any specific kind of theory, but insofar as it was, it is most likely to have been directed toward the truth (it would be odd for the theorist to try to construct a *specific* kind of false theory).

[15] These were suggested by an anonymous referee for *Mind*.

tive support for the theory, and for the theorist's aim at the truth, which in turn supports the theory.
Reply: Here it is important to note that there are two independent sources of evidence for theory *T*. The datum *D* may provide direct evidence for *T*. But it is not *D*, but ES, the theorist's holding a datum-entailing theory which provides evidence that she was reliably aimed at the truth, and hence for the truth of her theory *T*. It is only the significance of this latter item of evidence which is affected by whether *D* was predicted or accommodated.
Objection 2: One factor which determines a theorist's aim at the truth is how much evidence she has to go on. A theorist will have more evidence to go on if she accommodates a datum than if she chooses her theory without that datum, but predicts it. Hence, it might seem, in the a case of accommodation, the theorist has a better aim at the truth and hence we have more reason to believe her theory.
Reply: Notice that whether the theorist predicts or accommodates a datum gives us no indication at all as to her ability at assessing evidence, the reliability of her equipment and methods, or anything along those lines. That the theorist accommodated datum *D* entails that she is aimed at the truth just in this respect: she will come up with a theory which entails *D*, and since *D* is true, she is somewhat more likely to hit upon a true theory than if her theory did not entail *D*. But in this respect, she is no better off than if she had predicted *D*, for in either case her theory entails *D*.

8. The lottery prediction example

The advantage of prediction can best be illustrated with a case in which the datum *D* provides little or no evidence for the theory *T*, when *T* merely accommodates *D*. Compare the following two cases:

> **Accommodation**: We read in the paper that Jane won the national lottery. Fred proposes the following theory to explain this fact: the lottery was rigged in Jane's favour.
>
> **Prediction**: Before the lottery is even drawn, Fred proposes the theory that it is rigged in Jane's favour. We later discover that Jane won.

In the second case we are far more inclined to believe Fred's theory than in the first. In the second case we suspect that he must have been onto something, that he must have had some kind of reliable access to the facts concerning the lottery setup, to have been able to predict the lottery's outcome.

First let's look briefly at why the datum

(*D*) Jane won

does not render the theory

(*T*) The lottery was rigged in Jane's favour

very probable if *T* merely accommodates *D*. While *T* does entail *D*, it does so only at the expense of being highly improbable. Theory *T*, we might say, inherits the arbitrariness of *D*, for even if the lottery was rigged, we have no more reason to suppose that it would be rigged in *Jane's* favour, than we have to suppose that Jane would win just by chance. Indeed the fact that Jane won hardly calls for an explanation in the first place; someone had to win, and it could just as easily be Jane, as anyone. In Bayesian terms, we can note that there is a weaker theory *T**, which states simply that *the lottery was rigged*, which is not confirmed at all by Jane's winning, since Jane is no more likely to win given that the lottery was rigged. But now since *T* entails *T**, *T* can be no more probable than *T**. That is, Jane's winning renders the hypothesis that the lottery was rigged in Jane's favour no more probable than that the lottery was rigged at all.

But there is something else which we might want to explain, apart from Jane's winning, namely Fred's holding of a theory which entails her winning. Or rather, *Fred's holding a theory which entails the actual outcome of the lottery*, (his holding a theory that entails *Jane's* winning is significant only if Jane was the actual winner). The question that strikes us is, out of all the possible theories concerning the mechanics of the lottery, how did Fred manage to get one into his head which happens to entail the actual lottery result? Now of course in the accommodation case, the answer is straightforward. Since Fred knew that Jane won, he could select his theory under the constraint that it must entail this outcome. Apart from this constraint, his theory construction need not have been aimed at the truth, it may have been just a wild speculation.

In the prediction case, Fred did not select his theory under the constraint that it entail the data, so we need a different explanation. The natural hypothesis that comes to mind is that Fred was somehow reliably hooked up to the facts. On this assumption, he is far more likely to come up with a theory which entails the actual outcome. It would be an extraordinary fluke, if he just guessed a theory which entailed the actual outcome. So in the case where Fred's theory predicts the data, we have reason to suppose he was reliably hooked up to the facts, which in turn gives us reason to suppose that he is right.

9. Meeting the anti-predictionist challenge

We are now in a position to see how the new account avoids the problems of the standard No-Coincidence Argument and meets the anti-predictionist challenge. We have two potential explanatory hypotheses for the fact ES that the theorist chose a datum-entailing theory: the design hypothesis DS, and the reliable aim hypothesis RA. The design hypothesis renders the reliable aiming hypothesis unnecessary, with respect to explaining the theorist's entailment-success. For what we have here are two causal hypotheses concerning the process by which the theory was selected, each of which potentially explains the result of the selection. This is a case of *causal pre-emption*. Perhaps the theorist's process of theory selection had a good chance of producing a true, and hence datum-entailing theory. But in a case of accommodation, this causal explanation is pre-empted by the fact that non-datum-entailing theories were not even *open* to selection. The fact that, knowing the datum, the theorist restricted her theory selection to datum-entailing theories guarantees that she would select a datum-entailing theory, and no further hypothesis regarding her aim at truth is necessary to explain her doing so.

We can now see the plausible non-mysterious way that information concerning a certain psychological process in the theorist's head, namely designing her theory to entail a certain datum, is epistemically relevant to the truth of her theory. This information, DS, is relevant in that it *screens off* the confirmation of the hypothesis that the theory was reliably selected, by the fact that the theory entails the datum. In doing so, it diminishes the support that the theorist's entailment-success provides for her theory.

A very simple Bayesian analysis brings this out, by comparing the relation between ES and RA, first on the assumption of ¬DS, and then assuming DS:

$$P(ES|RA \wedge \neg DS) > P(ES|\neg DS)$$

$$\text{and so, } P(RA|ES \wedge \neg DS) > P(RA|\neg DS) \qquad (1)$$

that is, relative to ¬DS, ES confirms RA. However:

$$P(ES|RA \wedge DS) = P(ES|DS)$$

$$\text{and so, } P(RA|ES \wedge DS) = P(RA|DS) \qquad (2)$$

that is, relative to DS, ES and RA are independent. So DS screens off the support that ES provides to RA. Furthermore, adding the assumption that:

$$P(RA|DS) \leq P(RA|\neg DS) \qquad (3)$$

that is, without knowing whether the theorist holds a datum-entailing theory, her designing her theory to entail the data, makes it *no more* likely that her theorizing was reliably aimed at the truth, it follows from (1)–(3) that:

$$P(RA|ES \wedge \neg DS) > P(RA|ES \wedge DS)$$

that is, given the theorist's entailment-success, her having designed her theory to entail the datum, renders it less probable that her theorizing was reliably aimed at the truth, and hence less probable that her theory is true, which is the thesis of Strong Predictionism.

10. The degree and circumstances of the epistemic advantage of prediction

Granted that the successful prediction of a datum can, in principle, have an epistemic advantage over the mere accommodation of that datum, it remains to be seen in what range of circumstances this holds and to what degree. In particular, we should address a certain worry, namely that I have shown only that the weak predictionist thesis is true (which was never in dispute anyway) but not the strong thesis. Recall that according to Strong Predictionism, the fact that *T* correctly predicted rather than accommodated *D*, typically provides further evidence for *T*, even if we are familiar with the content of *T* and all the background evidence supporting it. Now according to the new account, information that a datum was predicted by a theory, can rationally affect our confidence in the theory, by indicating something about how well the process of theory selection was aimed at the truth. But this theory selection process just consists in the evaluation of *evidence*. So it might seem that in a case where we know what the theorist's evidence is, we can see for ourselves how well her theorizing was aimed at the truth, and hence any other indications of her aim, such as whether she designed her theory to entail the data, will be irrelevant.

In response, it must be granted that our knowledge of the theorist's total evidence *diminishes* the relevance of whether her theory predicted or accommodated a particular datum. For knowing what evidence she had to go on gives us at least a good indication of how well her theorizing was aimed at the truth. The crucial question is whether knowledge of the theorist's evidence, *entirely screens off* the relevance of further

information concerning her theory selection process, such as whether a certain datum was predicted or accommodated.

It seems clear that knowledge of the theorist's total evidence does not entirely screen off the relevance of this further information. For while information concerning the evidence that the theorist had to go on is very relevant to how reliably her theorizing was aimed at the truth, it does not settle the matter. It is useful here to consider two important factors linking evidence and theory. First, there are certain a priori epistemic constraints on how evidence should be *assessed* in forming theories. We can think of this in terms of a range of degrees of confidence that an ideal epistemic agent might have in a theory, given a body of evidence (how wide this range is, that is, how tight the a priori epistemic constraints are, is an open question). Second, there are various *causal* relations which are not knowable a priori: these include the reliability of our perceptual faculties, the trustworthiness of various sources, the accuracy of our measuring instruments, and so on. The crucial point here is that the degree of reliable aim of theorizing depends on both factors, neither of which is entirely *transparent* to us.[16]

Concerning the first, since we are not ideal epistemic agents, we are fallible in our assessment of evidence. For instance, construction of a theory might involve complex mathematical derivations where there is plenty of opportunity for errors, even if we double check our work. In some cases, the inference from evidence to theory involves intuitive judgements, the principles of which are not easy to spell out. A particularly striking case of this is our ability to 'read' a person's facial expressions, even though we cannot easily say how we interpret the visual cues on which our judgements are based. We have no trouble forming such judgements, but the degree to which the visual evidence supports our judgement may be in doubt, and no amount of double checking our inference can help us.

Suppose now the theorist comes up with theory T via complex derivations and intuitive inferences from a multifarious collection of background evidence E. T entails D, a possible outcome of a crucial experiment. Upon later discovering that D is true, we have reason to increase our confidence in her assessment of the evidence E. For if her assessment of the evidence was well attuned to the actual degree of epistemic support between the evidence and the various candidate theories, she had a better chance of hitting upon a true, and hence datum-entailing theory. Suppose on the other hand that knowing the data in advance, she narrowed down the pool of candidate theories, by elimi-

[16] Lipton (1991) makes a similar claim that the theory's *simplicity* need not be transparent.

nating those which do not entail the datum, and then applied the mathematical calculations and intuitive judgements to select among this narrower pool, in the light of background evidence *E*. In this case it is no surprise that the resulting theory entails the datum, and hence we have no *further* grounds for confidence in her assessment of the evidence.

Similar points apply to the various causal relations which the process of theory selection involves. For instance in choosing theories we often rely on the use of measuring instruments, opinion of experts, and our own perceptual faculties, the reliability of which is open to question. That the theorist came up with a datum-entailing theory may indicate that such causal connections were indeed reliable, but only if her theory selection process did not involve narrowing down the candidate theories in the light of the known datum. For as before, if the measuring devices, opinions of colleagues and so on, did not lead to theory *T* on their own, but only when various non-datum-entailing theories were already eliminated, it is no surprise that the selected theory entails the datum, and hence the theorist's entailment-success would be no indication of the reliability of these devices.

So Strong Predictionism is vindicated. Even if we know all the evidence on which the theorist based her theory, the fact that a certain datum was predicted rather than accommodated, may provide further evidence for the theory. When it comes to our actual theoretical practices however, the strong thesis is not particularly relevant, since we typically do *not* know all the evidence on which a theory was based. The multifarious considerations which lead to the acceptance of a theory are often too subtle and complex to be easily communicated. Indeed, even in our *own* case, we do not typically keep careful track of all the reasons we ever had for adopting a certain theory. We might come to question just how good all our reasons were, in which case the predictive success of our theory can suggest that our reasons were not bad at all.

The degree to which prediction has an epistemic advantage over accommodation can be seen now to depend on how well we understand the process by which the theory in question was selected, and the bearing of the new predicted or accommodated data. In some cases a datum might provide overwhelming support for the theory by itself, even if we know nothing of the independent evidence, and hence whether the datum was predicted or accommodated can make little difference. The coin's landing heads fifty times for instance, was all the evidence we ever needed to conclude that it is double-headed; whether

these landings were predicted or accommodated can make little further difference. In other cases the new datum may be less than conclusive, but we have a very firm grip on the other factors which led to the theory's acceptance, and hence learning that the datum was predicted will affect our estimation of the reliability of the process of theorizing only to a negligible degree. It is in cases where the new datum provides less than conclusive support for the theory, *and* we have either limited knowledge of the background evidence, or limited abilities in assessing it, that the information that the data was predicted, rather than accommodated is most significant.

11. Super-strong Predictionism

What I have called 'Strong Predictionism' is strong enough to be disputed by many, while weaker than that which some predictionists may accept, and some anti-predictionists see as their real target. The structure of my argument for Strong Predictionism was similar to the arguments for Weak Predictionism. In each case the epistemic significance of prediction to the truth of a theory was in a certain sense *indirect*. The fact that the theory predicted rather than accommodated the datum increases our estimate of some intermediate factor, which in turn confirms the theory. In the case of Weak Predictionism, the intermediate factors were the theory's simplicity and background evidential support. For the strong thesis it was the theorist's aim at the truth.

We can imagine an even stronger thesis, call it Super-strong Predictionism, according to which the fact that a theory predicted rather than accommodated a datum is evidence all by itself for the theory, quite apart from our estimates of these intermediaries. It is not clear that anyone has explicitly endorsed this thesis, but it may be what some have in mind, and may be what many who consider themselves opponents of predictionism are really opposed to. So it is worth noting that my discussion in no way supports Super-strong Predictionism. Indeed my account of when and how prediction matters might undermine the temptation to accept the super-strong thesis. For this temptation may be due to a simplistic generalization from cases where predicted data seem to carry more weight, the antidote to which is a more subtle account when and how prediction matters. In this way my defence of predictionism may be welcomed by those in the broadly anti-predictionist camp.

12. The No-Miracles Argument for scientific realism

I will conclude with some very general suggestions as to how my account of the epistemic advantage of prediction might be applied to a defence of scientific realism.[17] Those who believe that our current scientific theories are true, or at least approximately true, must face the fact that no matter what data we have supporting a theory, there are numerous alternative theories which entail the same data (or indeed coincide in all their observable entailments). The challenge for the realist is to explain why some theories are more likely to be true than the many other theories which entail the same observable data. One response to the challenge is to appeal to further principles of confirmation; for instance, it might be argued that some theories provide a *better explanation* of the data than others, where the criterion for a good explanation goes well beyond mere *entailment* of data. This is not the strategy I wish to focus on. There is what I believe is supposed to be an *independent* argument, which claims, in Putnam's (1975) words, that scientific realism 'is the only philosophy which doesn't make the success of science a miracle.' (p. 73) It would be a *miracle*, it is sometimes suggested, that an aeroplane should fly me safely home to Sydney if the aerodynamical theories on which its design is based were not *true*. Aeroplane flight is a tricky business. Of course it is *possible* for a wildly false theory to entail the correct results—but why would any sane person get in a plane if he did not think the theories on which its design was based were close enough to correct?

The No-Miracles Argument involves an inference to the best explanation of the *success* of science, where this success just consists in the fact that our theories entail certain data concerning say, aeroplane behaviour. The claim is that the *truth* of these theories can explain their success, which seems fair enough, since necessarily, truths entail truths. But let's look more closely at what this success consists in. Let D specify the aeroplane behaviour that our theory T entails. Since T necessarily entails D, T is successful if and only if D is true. So to explain the fact that T is successful, is just to explain D. But now there is something odd about the idea that *truth* could play some explanatory role, over and above the facts which obtain, if our theories are true. If anything explains the fact that aeroplanes stay up, it is (roughly) that the pressure on the underside of a moving airfoil is greater than the pressure on its

[17] The version of scientific realism that I am concerned with here states that our best current scientific theories are (approximately) true. The kind of antirealist I have in mind is one who, like van Fraassen (1980) insists that our theories should 'save the phenomena', that is, fit all possible observational data, but remains agnostic as to which of the many theories which do so, are true.

overside. There is no *further* explanation to be found by appealing to its being *true* that the pressure on the underside of a moving airfoil is greater than the pressure on its overside.[18] The appeal to truth seems superfluous, and hence this line of argument carries *no more* force than just the appeal to the explanatory strength of theory *T*.[19]

To see more clearly what it going on here, compare the following two explanatory arguments.

	Original Argument	**No-miracles Argument**
Explanandum:	*D*	*T* is successful
Explanans:	*T*	*T* is true

Originally, *T* is defended on the basis of its explaining the datum *D*. Rightly or wrongly, the antirealist is not thereby persuaded of *T*'s truth, but only of its empirical adequacy. It is a mistake to suppose that the No-miracles argument provides some *further* case for the truth of *T*. For the explananda and explanatia of the two arguments are equivalent. Indeed, the No-miracles argument seems to carry *less* force against the antirealist challenge. The antirealist is already convinced that *T* is *empirically adequate* (not on the basis of a fancy philosophical argument, but the regular evidence found in the physics journals). So she believes that the truth lies somewhere in the class of theories empirically equivalent to *T*. Our job is to persuade her that it is T which is true in particular. No matter how strong our argument is in support of *T*, if an argument of equal strength can be proposed in support of each of *T*'s empirically equivalent rivals, then our argument is irrelevant to meeting the antirealist challenge. So any argument which has any hope of answering the antirealist must appeal to some feature of *T* which distinguishes it from its empirically equivalent rivals. But of course the property of being predictively successful is shared by *all* of *T*'s empirically equivalent rivals. Any empirically equivalent theory T' will also entail the datum *D*, and hence be successful. So we could just as well argue that the success of T' is best explained by its being true, and so on, for all of *T*'s empirically equivalent rivals.

[18] The example is adapted from Levin (1984), who uses it to make a somewhat different point.

[19] This does not depend on a controversial deflationary account of truth, but merely that propositions of the form *it is true that P* and *P*, are known to be necessarily equivalent, and hence one cannot explain any more than the other does.

Clearly we are making no progress in meeting the antirealist challenge this way. At least in the original argument, where we were not merely appealing to the truth of *T*, but spelling out the details of this specific theory as an explanation of *D*, there was some hope of showing that *T* had an explanatory advantage over *T′*. For in seeking to explain *D*, we may appeal to features of *T* in which it differs from *T′* and other rivals. The No-Miracles Argument fails to discriminate among these rivals.

13. Reconstructing the No-Miracles Argument

So the No-Miracles Argument is not very satisfactory as it stands. Nevertheless, it carries considerable intuitive force. Perhaps we can reconstruct it along more promising lines, following my explanation of the epistemic advantage of prediction. Note that the No-Miracles Argument assumes, at least implicitly, that successful *predictions* have a special epistemic status.[20] For if our theories simply *accommodated* masses of data collected about aeroplane flight (suppose we never designed planes, but Martians gave them to us long before we could grasp aerodynamic theory) the argument would have much less bite. For what then would be the miracle? Our theories would entail these data because we made sure they do. And, the antirealist will be quick to add, there are plenty of other theories we are free to choose from which do likewise.

The key to reconstructing the No-Miracles Argument is to see that what needs to be explained is not just the fact that our aerodynamical theories entail that aeroplanes stay up, or that aeroplanes do stay up, but that scientists have managed to hit upon theories which entail that they do, and in general, that they have hit upon theories which are remarkably successful in a variety of practical applications. This would be a *miracle*, if there were *no* explanation for it, and an obvious candidate explanation is that the methods of science tend to get at the truth. This line of argument has at least the potential to meet the antirealist challenge, because it does discriminate between empirically equivalent rival theories. *T* and *T′* may entail the same observational data, but it is by choosing *T* that the scientist managed to successfully predict the data, and if her doing so suggests that her theorizing was reliably aimed at the truth, this will support *T* over *T′*. Our argument does not directly support *T*, but supports the reliability of the scientist's theory selection.

[20] Leplin (1997) is one recent proponent of the argument who is explicit about it.

T is thereby supported over T' and other rivals, since T is the theory that the scientists actually accept.[21]

14. Limitations of the reconstructed No-Miracles Argument

The reconstructed No-Miracles Argument is far from conclusive. It appeals to scientist's propensity to get at the truth, to explain their empirical success. But perhaps we can explain this by supposing that they have a propensity to get at empirically adequate theories, without the more specific tendency to get at the truth. Of course, scientists cannot straightforwardly select their theories to entail all the data, since a fair amount of the data is not known when the theory is chosen (it is their predictions, not accommodations which need explaining). Still, there might be some indirect means by which the process of scientific theorizing tends toward successful theories, without tending more specifically toward the truth. A thorough defence of realism will have to show why truth-propensity provides a better explanation of scientific success than empirical-adequacy-propensity.[22]

A thorough assessment of these competing explanations is beyond the scope of this paper. But it is interesting to consider the prospects of a more aggressive antirealist response, which tries to show in more detail *how* the success of science need have nothing to do with an aim at the truth. Here it is worth considering the neo-Darwinian response which van Fraassen (1980) offers to the No-Miracles Argument:

> The success of current scientific theories is no miracle. It is not even surprising to the scientific (Darwinist) mind. For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive. (p. 40)

The force of this objection depends on how we construe the argument against which it is directed. Against the standard No-Miracles Argument, it carries no force at all. The fact that unsuccessful theories are selected against may explain why no such theories are currently held, that is, that all currently held theories are successful. But it does not

[21] On the surface, Richard Boyd's (1984) subtle version of the No-Miracles Argument seems to be more along the lines I am suggesting, as his proposed explanandum is not the success of scientific theories, but the success of the *methodology* in producing successful theories. However, instead of appealing to the truth-aim of the methodology, to explain its success, which in turn supports the results of the methodology, Boyd explains the methodology's success in terms of the approximate truth of the background theories on which the methodology crucially depends.

[22] Note that this response does not undermine the argument for Strong Predictionism. A theorist's aim at the truth is still more strongly confirmed given that a datum was predicted than if it was accommodated, even if this difference is diminished by the alternative empirical-adequacy aim hypothesis.

explain why *those very theories* are successful. (Compare: the fact that pens which do not work are quickly discarded, explains why no such pens are used, that is, that all currently used pens work. But it does not explain why *those very pens* work). There remains the question, for each currently accepted theory, why *it* entails the data.[23] And a potential explanation for this fact is that the theory is true (even though, as we have seen, this does not by itself meet the antirealist challenge).

Something like Van Fraassen's response, however, might seem to undermine the reconstructed version of the No-Miracles Argument. For here the explanandum is the fact that scientists happen to hold successful theories. And it looks as though we can explain this fact by noting that unsuccessful theories are quickly discarded. But here we need to be careful. What we need to explain is not just that scientists currently hold *only* successful theories, that is, that they hold no unsuccessful ones—this would be the case if they held no theories at all. We also want to explain why they hold many successful theories. The fact that Jane drives only reliable used cars might be explained by her policy of getting rid of cars that do not work. But we might still wonder how she managed to get a reliable used car at all since (let's suppose) the vast majority of used cars on the market are lemons. The answer might lie in her skill in checking the engine, transmission, and so on, and hence being able to pick a good car from a bad one.

Recall the striking fact that Fred came up with a successful theory concerning the rigging of the lottery. Of course, if Jane had not won, then Fred's theory that the lottery was rigged in her favour would have been quickly discarded. But the fact that Fred currently holds a successful theory, is hardly explained away by the fact that if it had not been successful, then Fred would not still hold it. We are still struck by the fact that he managed to come up with a successful theory in the first place. To take the Darwinian analogy seriously, it *would* be a miracle if a single mutation produced an entirely new complex organism well adapted to its environment. We could not dispel its amazingness by noting that if the organism had not turned out just the way it did, it would not have survived.

What is required to dispel the miracle, is a vast plenitude of such mutations, most of which are selected against. The *plenitude* of mutations explains why there are some 'successful' organisms, while the

[23] Something like this has been noted by proponents of the No-Miracles Argument, such as Musgrave (1988), Lipton (1991), Leplin (1997) and Psillos (1999). This might partly explain why they insist on the standard version of the argument, the version which I have argued fails to meet the antirealist challenge.

selection explains why there are *only* successful ones. Perhaps if Fred had been in the habit of proposing lottery rigging theories every day for years, we should find it less remarkable that he managed to come up with a successful one eventually (similarly, perhaps Jane's success in picking reliable used cars has nothing to do with automotive expertise, but rather her policy of buying *dozens* of cars, and ditching those that stop working). The plenitude of Fred's attempted theories would not explain why, in the case of *Jane's lottery* he managed to hit upon a successful theory. Here there may be no explanation at all, that is, he was just lucky. But they would make his success less *surprising*, and hence less in need of an explanation in terms of his reliable connection to the truth.

So we at least get a glimpse at how we might render it non-miraculous that scientists manage to come up with theories which are marvelously successful in applications such as aeronautics. Very roughly, it might be a matter of trial and error. Plenty of dud theories are suggested and rejected when they fail, rendering it unremarkable that scientists currently hold successful ones. There are two components to the explanation, as there are two facts to explain:

(**S1**) Scientists currently hold *some* successful theories

(**S2**) Scientists currently hold *only* successful theories

Van Fraassen's point about the harsh environment into which theories are born explains S2, while the multiple attempts at successful theories explain S1.

This mere glimpse however, seems insufficient to ground any conclusions. We need to look at such factors as just how remarkable the successes of science are and the ratio of successful to unsuccessful theories, among those theories that have been *proposed*, not merely those currently held. At least on the face of it, it does not seem that scientists have been churning out so many unsuccessful theories that a few startling successes are to be expected by chance. What I hope to have done in the preceding discussion is clarify how the debate should go. As matters stand, it seems that we need to explain, or at least render non-miraculous, the fact that scientists manage to come up with so many remarkably successful theories. An explanation might lie in the mechanisms of theory selection being directed toward the truth, and the acceptance of such an explanation should increase our confidence in the truth of scientific theories in general. But this explanation threatens to be undercut by the alternative explanation that the mechanisms of

theory selection are directed toward successful theories, without being more specifically directed toward the truth. Van Fraassen's attempt to show how this might work is suggestive, yet not compelling without a lot more detail. How matters will stand when the details are considered is an open question.[24]

Department of Philosophy ROGER WHITE
New York University
503 Silver Center
100 Washington Square East
New York, NY 10003
USA
rlw5@nyu.edu

References

Achinstein, Peter 1994: 'Explanation v. Prediction: Which Carries More Weight?', in Hull, Forbes, and Burian (eds) 1994, pp. 156–64.

Boyd, Richard 1984: 'The Current Status of Scientific Realism' in Leplin (ed.) 1984, pp. 41–82.

Brush, Stephen 1994: 'Dynamics of Theory Change: The Role of Predictions' in Hull, Forbes, and Burian (eds) 1994, pp. 133–45.

Collins, Robin 1994: 'Against the Epistemic Value of Prediction over Accommodation'. *Noûs,* 28, pp. 210–24.

Duhem, Pierre 1956: *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Earman, John, (ed.) 1983: *Minnesota Studies in the Philosophy of Science*, Vol. 10, Minneapolis: University of Minnesota Press.

Fine, A and Leplin, J. (eds) 1988: *Proceedings of the Philosophy of Science Association* Vol. 1, East Lansing Mich.: Philosophy of Science Association.

Giere, Ronald N. 1983: 'Testing Theoretical Hypotheses', in Earman 1983 pp. 269–98.

Gooding, D., Pinch, T. and Schaffer S. 1989: *The Uses of Experiment*. Cambridge: Cambridge University Press.

Hull, D., Forbes, M., and Burian R. (eds) 1994: *Proceedings of the Philosophy of Science Association* Vol. 2, East Lansing, Mich.: Philosophy of Science Association.

[24] Thanks to Adam Elga, Ned Hall, Patrick Hawley, Bob Stalnaker, Judy Thomson, participants in my 2001 Epistemology Seminar, and an anonymous referee for *Mind*, for helpful comments on drafts.

Keynes, John Maynard 1921: *A Treatise on Probability*. London: Macmillan.
Leplin, Jarrett (ed.) 1984: *Scientific Realism*, Berkeley: University of California Press.
——1997: *A Novel Defense of Scientific Realism*. Oxford: Oxford University Press.
Levin, Michael 1984: 'What Kind of Explanation is Truth?' in Leplin 1984, pp. 124–39.
Lipton, Peter 1991: *Inference to the Best Explanation*. London: Routledge.
Horwich, Paul 1982: *Probability and Evidence*. Cambridge: Cambridge University Press.
Howson, Colin and Franklin, Alan 1991: 'Maher, Mendeleev and Bayesianism', *Philosophy of Science*, 58, pp. 574–85.
Maher, Patrick 1988: 'Prediction, Accommodation, and the Logic of Discovery' in Fine and Leplin (eds) 1988, pp. 273–85.
Mill, John Stuart 1904: *A System of Logic*. London: Longmans, Green and Co.
Musgrave, Alan 1988: 'The Ultimate Argument for Scientific Realism', in Nola (ed.) 1988, pp. 229–52.
Nola, R. (ed) 1988: *Relativism and Realism in the Sciences*. Dordrecht: Kluwer Academic Press.
Peirce, Charles Saunders 1931–51: *The Collected Papers of Charles Saunders Peirce*. C. Hartshorne and P. Weiss (eds) Cambridge Mass.: Harvard University Press.
Pitt, Joseph C. (ed.) 1985: *Change and Progress in Modern Science*. Dordrecht: D. Reidel.
Psillos, Stathis 1999: *Scientific Realism: How Science Tracks the Truth*. London: Routledge.
Putnam, Hilary 1975: *Mathematics, Matter and Method: Philosophical Papers Vol. 1*. Cambridge: Cambridge University Press.
Schlesinger, George N. 1987: 'Accommodation and Prediction', *Australasian Journal of Philosophy*, 65, pp. 28–42.
Whewell, William 1860: *Philosophy of the Inductive Sciences*. London: John W. Parker.
Worrall, John 1985: 'Scientific Discovery and Theory Confirmation', in Pitt (ed) 1985, pp. 301–31.
——1989: 'Fresnel, Poisson and the White Spot: The Role of Successful Predictions in the Acceptance of Scientific Theories', in Gooding, Pitch, and Schaffer (eds) 1989, pp. 135–57.

Van Fraassen, Bas C. 1980: *The Scientific Image*. Oxford: Oxford University Press.
Zahar, Elie 1973: 'Why did Einstein's Programme Supercede Lorentz's? (I)' *British Journal for the Philosophy of Science,* 24, pp. 95–123.