Proceedings of the Aristotelian Society, Supplementary Volume LXXIV (2000) 179-206.

Tracking Track Records

PETER LIPTON

ABSTRACT From a reliabilist point of view, our inferential practices make us into instruments for determining the truth value of hypotheses where, like all instruments, reliability is a central virtue. I apply this perspective to second-order inductions, the inductive assessments of inductive practices. Such assessments are extremely common, for example whenever we test the reliability of our instruments or our informants. Nevertheless, the inductive assessment of induction has had a bad name ever since David Hume maintained that any attempt to justify induction by means of an inductive argument must beg the question. I will consider how the inductive justification of induction fares from the reliabilist point of view. I will also consider two other well-known arguments that can be construed as inductive assessments of induction. One is the miracle argument, according to which the truth of scientific theories should be inferred as the best explanation of their predictive success; the other is the disaster argument, according to which we should infer that all present and future theories are false on the grounds that all past theories have been found to be false.

> `And here it is constantly supposed, that there is a connexion between the present fact and that which is inferred from it. Were there nothing to bind them together, the inference would be entirely precarious.' (David Hume, *Enquiry* IV)

I

The Inductive Assessment of Inductive Methods

We often engage in the inductive or non-demonstrative assessment of inductive methods, investigating the track record of a given method to gauge its future prospects. Testimony is a clear example. Relying on someone's word is an inductive method in my broad sense, since the truth of what people say does not follow deductively from the fact that they said it. And we often assess the method of testimony inductively. One way that we decide whether to accept people's testimony is by investigating their track record

-- how trustworthy they have been in the past. This is an inductive assessment of an inductive method, a second-order induction. The calibration of instruments is another example. That an instrument gives a particular reading does not of course entail that the reading is correct, since instruments sometimes malfunction. The general use of an instrument is thus an inductive method. The calibration and testing of instruments are also inductive methods, so when we calibrate and test our instruments we are engaging in the inductive assessment of an inductive method.

I enjoyed a particularly stark example of the inductive assessment of an inductive method some years ago, during a hiking trip in the Teton mountains of the Western United States. Sitting down on a large rock to have my lunch, I was joined by another hiker, who turned out to be a meteorologist. Our conversation turned to predicting the weather, and of the comparative reliability of the various methods in use. She informed me that the method of `persistence forecasting' still compared favourably with other, more sophisticated techniques. Replying to the obvious question, she explained that `persistence forecasting is where you say the weather will be the same tomorrow as it was today'. Perhaps she was kidding, but the thought that we might use the track record of persistence forecasting to assess its future prospects is perfectly coherent. Another suggestive example is the use of three identical computers on Apollo spacecraft. The system was designed so that the computers performed the same calculations at the same time. If one of the three computers gave an answer different from that given by the other two, the system would follow the output of the two that agreed. In effect, the computers in this system were assessing each other's reliability.

In these examples, the inductive assessment of inductive methods seems relatively unproblematic, by which I mean no more problematic that inductive arguments generally. And this is what one might expect, since the claim that an inductive method is reliable is tantamount to a physical hypothesis that there is a correlation between the output of the method and the state of the world. If there can be legitimate inductive arguments concerning the expansion of metals when heated, there can be legitimate inductive arguments about the reliability of thermometers.

Another reason for supposing that inductive assessments of inductions are no worse than first-order inductions is the very strong interpenetration of method and substantive claim in our cognitive economy. Our methods take us to our claims, but our claims also play an indispensable role in the constitution of our methods. Instrumentation again provides a clear example. A measuring device depends not just on

the hardware and its application, but also on the theories that fix what the output of the machinery should be taken to be measuring. More generally, what we take to be evidence is typically determined by our substantive background theories. It seems implausible that methods should have a fundamentally different epistemic status than do substantive claims. If we can use induction to evaluate the one, we ought to be able to use induction to evaluate the other.

II

Hume's Circle

Nevertheless, the idea of an inductive justification of induction has had a very bad reputation at least since David Hume's sceptical argument against induction. The Humean argument concludes that no inductive arguments are legitimate, but an essential step is the claim that the inductive justification of induction would be even worse than the rest, since it, unlike mundane first-order inductions, is viciously and worthlessly circular, not merely unjustifiable. Even if one has independent evidence that induction has been reliable in the past, this is supposed to provide no reason for saying that induction is reliable generally, since such an argument would itself be inductive. Hume of course does not argue that induction is generally unreliable, since that negative claim would itself require an inductive argument, leading Hume into the sort of circle he deplores in others; but his claim that any inductive argument for induction `must evidently be going in a circle, and taking that for granted which is the very point in question' (Hume 1777, Sec. 4) is the key step in the argument that our inductive practices are not just unjustified but unjustifiable. The thought is that, even if first-order inductions were somehow justified, inductive justifications of induction would still be illegitimate.

I have no solution to the problem of induction that fully satisfies me, but fortunately my aims in this paper are more modest. One of them is to use a general discussion of the broad class of inductive assessments of inductive methods to show that the inductive justification of induction is worth more than Hume's discussion suggests, though of course it will not move anyone who refuses to engage in inductive inference altogether. Second-order inductions are no worse that first-order induction. Some inductive assessments of inductive methods are themselves strong arguments and others weak, and the inductive justification of induction that Hume dismissed need not be one of the weak ones. I will go on to consider two philosophical arguments that are

tantamount to inductive assessments of inductive methods, both taken from the debate over scientific realism. One is the miracle argument, according to which the truth of scientific theories should be inferred as the best explanation of their predictive success; the other is the disaster argument (better known as the `pessimistic meta-induction'), according to which we should infer that all theories scientists ever produce are false, on the inductive grounds that all past theories have been found to be false. It will turn out that, unlike the simple inductive justification of induction, both of these inductive assessments are seriously flawed.

III

Inductive Strength and Tracking

To assess second-order inductions, we need to say something about what the difference between strong and weak inductive arguments generally. The project of providing a general account of inductive strength has proven to be remarkably difficult, but we are not entirely in the dark. We know, for example, quite a bit about the pathologies of induction: there are many familiar sources of weaknesses in the inductive support that available evidence provides for a given scientific hypothesis. As Nelson Goodman taught us (1983, ch. III), the hypothesis under investigation may be unprojectable (e.g. `All emeralds are grue'), so that its positive instances do not support it. Coming closer to everyday scientific concerns, evidence that does offer genuine support may nevertheless be too sparse, too imprecise, too little varied, or without suitable controls.

A familiar feature of inductive reasoning is its defeasibility. In contrast to deduction where, starting with a valid argument and adding premises always leaves a valid argument, a strong inductive argument may become weak one when premises are added. Observations of many and varied ravens, all found to be black, may provide strong inductive support for the hypothesis that all ravens are black, but everything change when the next raven is observed to be white. This is why inductive reasoning is subject to what is sometimes called the `total evidence condition': assessments of inductive support should be made relative to all the available and relevant data. The point should be generalised. Inductive strength must be assessed relative to all relevant background belief, not just to the data. This condition brings out other familiar inductive flaws. For example, one may believe that the process by which the data were collected was biased. If my assistants, knowing of my psychological investment in my pet ornithological

hypothesis, adopt a policy of only showing me black ravens, committed to the suppression of any non-black ravens they might discover, the fact that my data include only black ravens is not good evidence that all ravens are black.

There are two other types of inductive pathology that I flag here, because of the roles they will play later in the essay. The first of these is another and common way background belief may defeat an otherwise strong inference. In this sort of case, we believe that the phenomena under investigation are dynamic so that having a certain feature at one time is not a good reason to suppose that they will have the same feature at another time. Knowing what I do about human development, I know that the fact that every minute of my life has been followed by another is not a good reason for saying I am immortal (Quine & Ullian 1978, 86). Similarly, knowing what I do about my children's educational activities, I forebear inferring that what they do not know about today they will not know about tomorrow. The flaw of ignoring relevant dynamical information will become important when I consider the disaster argument against scientific realism.

The final inductive flaw I will mention is peculiar to certain inductive assessments of induction -- I will call it `self-certification'. Consider the argument that that a particular instrument is generally reliable, on the grounds that it has given accurate readings in the past. The past reliability of the instrument might be good grounds for claiming general reliability, but only if there is some independent reason for claiming that the instrument has indeed been reliable in the past. If by contrast one only claims that the instrument has been reliable because one trusts it, the argument is worthless. There are many analogous cases. You might argue that I have a mind on the grounds that everyone else does, but this carries no weight if you have no independent reason for supposing that other people do indeed have minds. Believing someone's testimony because he also says `Trust me!' may also be an analogous case.

Self-certification is an egregious flaw, to be sharply distinguished from legitimate inductive assessments of induction, such as an argument for the general reliability of an instrument on the basis of independent evidence that it has functioned well in the past. The flaw of self-certification plays an important role in my general strategy for legitimating the inductive justification of induction. Having distinguished clearly legitimate inductive assessments of inductive methods from worthless self-certification, my aim is to legitimise the inductive justification of induction by associating it with

benign inductive assessments of induction and disassociating it from vicious self-certification.

Many of the aspects of inductive assessment implicated by my list of inductive pathologies can naturally be captured from a counterfactual reliabilist perspective. The details are bound to be messy, but the general idea is that a strong inductive argument is one that is both an instance of a method that is generally reliable and is also an argument that is counterfactually reliable in this instance. The premises track the conclusion, in Nozick's sense of the word (1981, ch. 3, esp. 222-23). Slightly more generally, in a strong inductive inference, if the conclusion had not been true, we would not have made the inference. In a strong inductive argument, the premise, or what it describes, is a reliable symptom of the conclusion. Thus a reliable thermometer can be used to provide a strong inductive inference about the present temperature here because it not only gives what is in fact the correct reading, but also because it would not have given that reading had the temperature been different. The inference from Koplic spots to measles is a good induction, because the patient would not have had Koplic spots if she did not have measles.

From this point of view, we may view our inductive methods as the ways we attempt to make ourselves into reliable measuring instruments, where what we are measuring is the truth value of hypotheses. A good thermometer is sensitive to variations in temperature; a good inductive method makes us sensitive to variations in truth value. In a good inductive inference, if the conclusion had not been true, you would have noticed. If your evidence would have been just the same whether or not your hypothesis were true, the inductive connection is weak. Thus my prediction that it will rain is only strong insofar as my evidence would not have been just the same if it were not going to rain. The inductive weaknesses I mentioned above are from this point of view weaknesses because arguments that exhibits them are unlikely to satisfy the reliability requirements.

I am going to work with the simple counterfactual idea that a strong inductive inference is one where, had the conclusion been false, you would not have made the inference; but while I hope this will be good enough for my suggestive purposes in this essay, I know it is too simple for a fully adequate account of what makes for strong induction. For one thing, we want to allow for strong inductive arguments with conclusions that are alas nevertheless false, and my simple counterfactual does not appear to apply to these cases. For another, those who are adverse to backtracking

counterfactuals may have to analyse reliability with different conditionals, because of the prevalence of inductive inferences that move from past to future. But I do want to emphasise that it is my view no good dropping a counterfactual or subjunctive requirement altogether and trying to make do with statistical reliability of method. One salient reason is that such an account would fail to exclude vicious self-certification. To see this, suppose that an argument of enumerative form from the reliability of an instrument in the past to its future reliability has true premises and is itself an instance of a generally reliable form of inference. This is as may be, but the inference is still worthless if there is no independent check on the past reliability of the instrument. Intuitively, the counterfactual requirement is needed to capture the idea that a good inductive argument brings new or additional evidence to bear on the conclusion, something self-certifying arguments signally fail to do.

Because of the intimate connection between causes and counterfactuals, the tracking requirement covers the many good inferences that depend on the causal relation, such as the inference from Koplic spots to the measles. Reliable symptoms support strong inductive inferences, because in such cases the symptom would not be present if the underlying condition were absent. At the same time, the tracking requirement accounts for the precariousness of certain causal inferences in cases of overdetermination. I put the dinner in the oven, but the inference from hot dinner to my preparation is precarious because, had I not put the dinner in the oven, my wife would have. The tracking requirement may also usefully be more general than the causal relation. For example, we often infer from the proportion of a trait in a sample to the same proportion in a larger population. This is not obviously a causal inference, but it is sound when the tracking requirement is met, so that the sample would not have had that proportion of the trait if the population as a whole did not. Good sampling technique establishes a tracking relationship; biased or otherwise poor sampling does not.

The tracking requirement catches many of the inductive flaws mentioned above. Evidence that is sparse, insufficiently various, imprecise, uncontrolled or inadequately shielded from disturbing influences is bad evidence because it fails to establish an appropriate tracking relation to the hypothesis in question, and aspects of experimental practice designed to improve inductive warrant can naturally seen as attempts to establish or improve tracking. When such measures succeed, what they succeed in doing is making it more likely that, had the conclusion been false, the evidence would been different. Turning to the new riddle, the fact that all observed emeralds have been grue

is not good evidence that all emeralds are grue, since if future emeralds are not grue (that is, not blue), present emeralds would be still be grue (that is, green). Faced with the new riddle of induction, it is natural to feel that the problem is a failure of similarity: the inferred cases (blue emeralds) are dissimilar from the evidential cases (green emeralds). But we cannot require that the conclusions of good inferences be similar to their premises, since this would rule out all `vertical' inference, notably all scientific inference from observable evidence to unobservable properties and processes. The tracking requirement clearly does permit such inferences. Perhaps electrons are unobservable, but the inference from cloud chamber trail to electronic behaviour passes the tracking requirement so long as the trail would not have been there or that way without the electron.

The tracking requirement has other attractive features. In bringing out the way good inductive arguments turn us into instruments that reliably measure the truth value of hypotheses, it clarifies the sense in which what makes for a good inference is a physical rather than a logical matter, and so gives a natural explanation for the many familiar failures of logical accounts such as the hypothetico-deductive model. A second related attraction is that is gives a natural explanation for the importance of background beliefs in inductive inference, since those background beliefs will help to determine judgements of what would track what. Finally, it seems to me that the tracking requirement fits quite well with the idea of inference to the best explanation, an account to which I am particularly attracted (cf. Lipton 1991). For one way to judge whether the premises (the data) would track the conclusion (the hypothesis) is to ask whether the hypothesis would explain the data.

IV

The inductive justification of induction

Strong second-order inductions satisfy the tracking requirement. Thus independent checks on an instrument's reliability support a strong inductive inference about its general reliability insofar as, had the instrument not had that general degree of reliability, it would not have had the reliability we observed in the examined cases. Similarly, independent checks on an informant's veracity in some cases may be a good guide to his trustworthiness in other cases, because if he were going to lie in the one case, he would have lied in the others. This suggests that something like the traditional inductive justification of induction may also meet the requirements on a strong second-

order induction. Indeed it is difficult to see how first-order inductions to various empirical hypotheses could meet the tracking requirement and yet in principle the inductive justification of induction could not since, as I mentioned at the start of this essay, the claim that an inductive method is reliable is tantamount to an empirical hypothesis of lawful correlation between the output of the method and various states of the world. Moreover, to deny the relevance of independent evidence about the track record of induction seems rather like denying the inductive relevance of additional evidence. This will not of course move a sceptic who refuses to use any inductive arguments, but it may show that the inductive justification of induction is no worse than other inductive arguments. An argument may be worthless against a sceptic, yet not epistemically impotent for those willing to follow it.

The intuitive pull of the claim that induction cannot be used to justify induction is strong; yet most of us continue to feel that induction's track record is relevant to its future prospects. The fact that induction has been rather reliable in the past (witness our survival) does seem to give some reason to believe that it will continue to be reliable, and yet the circularity point seems to show that it cannot. I think that the tracking requirement helps to explain this tension in our thinking, and helps to give an accurate assessment of the real status of the inductive justification of induction. To see this, we need to look at the nature of epistemic circularity.

The simplest sort of circular argument is one of the form `P therefore P' -- deductively valid and cognitively worthless. More interesting are cases where the dependence of the premise on the conclusion is more subtle. A nice example is Descartes' argument (1641, Meditation Three) that every event must have a cause, on the grounds that a cause must have at least as much `reality' as its effect, a principle that would be violated if an event had no cause, since something always has more reality than nothing. The reality principle may itself be dubious, but the argument does not appear circular until Hume (1739, Sec. I.III.III) points out that the reality principle simply would not apply to uncaused events, and to treat `nothing' as if it would be a feeble cause rests on the tacit assumption that every event has a cause, which is what was to be established.

The inductive justification of induction, however, seems quite different from these cases. The premise of the argument (induction has been reliable) is different from the conclusion (induction will be reliable); nor do we believe the premise because we believe the conclusion. The alleged circularity of the inductive justification of induction rests not in the relation between the conclusion and premise, but between the conclusion

and enumerative `rule' of inference: the rule seems to presuppose the conclusion. This distinction between premise and rule circularity is sometimes cited to exonerate the inductive justification of induction; but without further argument it is unclear why rule-circularity should any less vicious than premise circularity.

What we should like at this stage is a good general account of circularity, but such a thing is surprisingly hard to produce. It will not do to say that an argument is circular when, if you do not believe the conclusion then you will not believe all the premises, since this runs the risk of tarring all valid arguments, and I take it that many deductive arguments are not circular. Even an argument where the conclusion is logically equivalent to the conjunction of the premises need not be circular: the free-will dilemma may be a case in point, since the first premise is a tautology and the latter two are material conditionals whose consequents are the same as the conclusion. Nor will it do to say that the conclusion of a circular argument provides an essential reason for a premise (or a rule, for that matter). For this seems itself to be circular, since the notion of reason depends on the notion of circularity.

Fortunately, we can make some progress on our particular problem without a full analysis of circularity. Intuitively, the judgement that an argument is viciously circular is often based on the judgement that the argument would be pathologically ineffective against anyone who did not already accept its conclusion. This feature is then taken to show that the argument is cognitively impotent generally, whatever one's prior attitude to the conclusion. We may, however, be able to prise apart these two notions, so that an argument that is ineffective against the sceptic may yet have some cognitive value. The inductive justification of induction is clearly not going to move any sceptic who refuses to indulge in non-demonstrative inference, but it may not follow that it has no cognitive value for those of us who already do so indulge. To see this, we may focus on cognitive impotence directly. The idea is that if we can show that the inductive justification of induction is not cognitively impotent, then we have shown that it is not viciously circular either.

One way this approach has been used in an attempt to exonerate the inductive justification of induction is by arguing that the inductive justification of induction is not circular because it could be used to gain knowledge of its conclusion, where knowledge is understood as reliably generated true belief (cf. Braithwaite 1953, ch. 8; Van Cleve 1984). The idea is that, if induction is indeed generally reliable, then someone who used induction could use the inductive justification of induction to move from mere belief to

knowledge of its general reliability. Arguments give us knowledge by having conclusions that are true and where we gain a reliable belief from premises themselves reliably believed and along inferential routes that are reliability preserving. In this sense, good arguments work by providing a reliable method of generating or sustaining a belief in their conclusions.

This reliability gain criterion is attractive. Cogent deductions and ordinary inductions rightly pass it; arguments that are clearly premise circular fail it. Thus an argument of the form `P therefore P' fails, because one would need already to have reliable belief in P in order to gain reliable belief in P. More simply, one couldn't use that argument to gain reliable belief that P. Cogent deductions pass, because we are not deductively omniscient: we can know the premises of a deductive argument without already knowing the conclusion, so going through the deduction may yield a gain in reliable belief. And ordinary inductions pass: if our inductive practices are in fact reliable, they can be used to gain reliable beliefs in conclusions from premises that may be independently known. The reliability gain condition thus appears to make the correct discrimination between circular and non-circular argument, and so the fact that the inductive justification of induction meets it seems to show that the inductive justification of induction is not circular.

As we have already seen, however, it matters how we construe reliability. Suppose that you believe that induction will be reliable in the future because you believe that it has been reliable in the past, but your only reason for saying that it was reliable is that this is what your inductive practices told you. Thus if the past inferences were to predictions, you never bothered to check them; if they were to laws you never went on to test them further. These are examples of the inductive flaw of self-certification, which I introduced above. Self-certifying arguments are viciously circular, yet also appear to permit the gaining of reliable belief. If induction is in fact generally reliable, then you did come to know the conclusions of those past inductions without checking. You then went on to infer future reliability, again on a reliable route, so it appears that you have gained knowledge of the future reliability of induction. The reliability gain criterion thus appears too permissive. It wrongly certifies worthlessly self-certifying arguments, so the fact that it also certifies the inductive justification of induction is cold comfort.

The inductive justification of induction, where one has independent knowledge of induction's track record, seems quite different and better than self-certification, but how are we to capture this difference? The difference is in the tracking. The self-

certifying argument is worthless because it is completely insensitive to the reliability of our rules. Perhaps if our rules were not generally reliable they would not have been reliable in the past, but in the self-certifying case we would not have noticed, having never checked. In short, although the self-certifying argument might meet a statistical reliability requirement, it does not meet the tracking requirement, while the inductive justification of induction does (or could). Just as a good ordinary induction establishes a tracking relationship between premise and conclusion, so does the inductive justification of induction. That is why the inductive justification of induction is not viciously circular. The past performance of induction may be a reliable symptom of its future prospects.

As promised, I am adopting a two prong strategy for legitimising the inductive justification of induction. One is sharply to disassociate the inductive justification from self-certifying arguments; the other is strongly to associate it with clearly legitimate second-order inductive assessments of specific inductive methods. I have been focussing just now on the first prong, but I want briefly to return to the second. One may well feel that there is a fundamental disanalogy between the inductive justification of induction and the other second-order inductions I have mentioned, involving testimony, instrumentation and weather prediction. The difference is that where these other inductions are indeed inductive assessments of inductive methods, only in the inductive justification of induction is the method assessed the same as the method doing the assessing. My response is twofold. Firstly, I am not sure that all of the other examples really do exhibit a clean separation between method used and method assessed. When we project the reliability of predictive forecasting, we are making an enumerative induction of a method that is itself very close to enumerative induction. And in the case of Apollo computers, we again have something very close to a method checking themselves. Analogously perhaps, one might imagine a thermometer capable of checking its own temperature. (Perhaps this is what every ordinary thermometer does.)

My second response is to re-emphasise the way in which the sort of self-assessment exhibited by the inductive justification of induction is not vicious. This is the point I have been attempting to make through the contrast with self-certification. It is not vicious because it may meet the tracking requirement; unlike a self-certifying argument, it tests its conclusion in a way that it might fail. Because we have independent evidence for the track record of induction, we might in principle have found that induction has not been terribly reliable in the past.

One might nevertheless still feel that, although the inductive justification of induction is indeed importantly different from self-certifying arguments, it is also importantly different from mundane second-order inductions. For if induction is in fact unreliable, then would not its past failure be irrelevant to its future prospects? The track record of induction is only relevant if induction is reliable, so here we seem to have an argument that is only strong if its conclusion is true, and this seems to mean that examining the track record of induction is not a real test after all, since if we did find that induction were unreliable in the past, we could infer nothing from this. Unlike self-certifying arguments, we do have an independent assessment of the truth of our premises; but we do not have the appropriate independence of inferential method and conclusion.

It is difficult to see the situation here clearly, but I want to suggest that although this link between conclusion and inference may generate a certain asymmetry between good and bad track records, this is not pernicious. In any inductive argument, the inference must track if it is to be strong. So if induction does not in fact track in the case of the inference from past to future performance of induction, then we are out of luck. I doubt however that this entails that any weakness in the track record of induction would automatically preclude inductive projection in the future, partly because reliability is a matter of degree, and partly because the counterfactual tracking requirement is `local', depending on the particular inference. But suppose that we accepted that a poor track record would tell us nothing about induction's future prospects. Even so, a good track record could tell us something. For as I never tire of saying, a good track record may satisfy the tracking requirement, such that if induction were not generally reliable we would not have inferred that it is. And this does not require that, had the track record been poor, we would have inferred future unreliability. It is enough that, if induction were not generally reliable, we would have witnessed a bad track record, and we would have properly turned agnostic about future prospects.

The general way I have been looking at the inductive justification of induction helps to explain our ambivalence towards it, our feeling both that appeal to the track record of our methods must be epistemically relevant and that it cannot be. We feel it cannot be relevant because such an appeal would not move the sceptic, and perhaps also because we fail to distinguish the inductive justification of induction from its degenerate, self-certifying cousin. At the same time, we feel that the track record is relevant, rightly in my view, because we recognise that this second-order argument may satisfy the same

tracking requirement that makes for good first-order arguments. The inductive justification of induction still will not move the sceptic, but it does not follow from this that the argument has no cognitive value for us. Of course if Hume is right in suggesting that nothing does bind one event to another, so that no premise of a non-demonstrative argument ever does track its conclusion, then the inductive justification of induction, like all inductive arguments, is worthless. But the inductive justification of induction is standardly supposed to suffer a special liability of circularity that ordinary inductions do not. What I have suggested is that this is wrong: the inductive justification of induction need be no worse than any other induction. The sceptic will not employ any of them, but those of us who are happy to use ordinary first-order inductions may also enjoy cognitive benefit from the inductive justification of induction.

V

The Miracle Argument

I turn now to the first of the two particular philosophical second-order inductions I wish to assess. This is the miracle argument for scientific realism, according to which we should infer that predictively successful scientific theories are true, since truth is the best explanation of this success (cf. Putnam 1978, 18-22). The miracle argument is clearly inductive, since the success of a scientific theory does not entail its truth. It is also virtually equivalent to an inductive justification of the methods by which the theory in question was supported: it is in that guise that I consider it a second-order induction. Moreover, the miracle argument may well meet the tracking requirement. Nevertheless, I will argue that the miracle argument is weak, because even if it is the case that the data would have been different if the theory in question had not been correct the miracle argument fails to provide any independent or additional evidence for the truth of the theory or the reliability of the methods that lead to its acceptance. Unlike the inductive justification of induction, which involved new and independent data about the track record of our methods, the miracle argument simply reiterates the scientific case for the theory.

According to the miracle argument, it would be miraculous, and hence very unlikely, for a scientific theory to generate numerous, varied and precise predictions, all of which turn out to be correct, unless that theory were at least broadly correct. We are supposed to be entitled to infer the truth of a scientific theory from such predictive success. The argument is often couched as an inference to the best explanation, a mode

of induction that is supposed to characterise many ordinary scientific inferences. According to inference to the best explanation, given the available data, we are entitled to infer a hypothesis which would, if correct, provide the best explanation of those data. Being the best explanation does not guarantee being correct, but this is just to say that the reasoning is inductive: being the best explanation is on this view supposed to confer high probability. In application to the miracle argument, the idea is that the truth of a theory would provide an excellent explanation of the success of its predictions, those predictions being deductive consequences of the theory, and there is no better explanation. Hence we are entitled to infer the truth of a successful theory.

The similarities between the miracle argument and the inductive justification of induction are striking. They are both inductive arguments, and both supposed to be distinct from the first-order inductive arguments made by scientists and ordinary folk. They are also both intended to provide justifications for those first-order practices, providing a reason for saying that they are generally reliable. Moreover, like the inductive justification of induction, the miracle argument has been accused of vicious circularity (cf Laudan 1984, 242-43; Fine 1984, 85-6). Assuming that inference to the best explanation does characterise scientific inference, those who are sceptical about the truth of scientific theories are sceptical about the reliability of inference to the best explanation as a route to truth. The miracle argument is however itself supposed to be an inference to the best explanation. Hence the miracle argument is using an inference to the best explanation to justify inference to the best explanation, apparently just the same sort of rule circularity as the inductive justification of induction's use of an enumerative induction to justify enumerative induction.

So it seems that my vindication of the inductive justification of induction should apply directly to the miracle argument. Although neither argument has any force against the sceptic, both may provide believers with reasons for their belief, because both argument have premises that may track their conclusions. Indeed the standard presentation of the miracle argument lends itself very naturally to a tracking gloss. To say that we may infer the truth of a successful theory on the grounds that it would be a miracle for the theory to enjoy this success if it were false appears tantamount to saying that, in such a case, the theory would not have enjoyed this success if it had been false, which is just the tracking claim applied to this case.

In my view this analogy holds only up to a point. It may show how the miracle argument avoids vicious circularity, but the miracle argument has also been accused of a

quite different sort of deficiency. The accusation is that, circularity aside, the miracle argument does not meet our own inductive standards, in this case the standards of inference to the best explanation, because truth is not in fact the best explanation of predictive success. The most straightforward way to prosecute this case is to generate competing explanations. Two such explanations seem to me particularly interesting. The first is the selection explanation, according to which the reason our current theories are successful is that scientific practice consists in the practice of testing theories to destruction, eliminating them as they fail. Small wonder then that the latest theories have been successful: had they not been, we would no longer have them (Van Fraassen 1980, 39-40). The second alternative (or class of alternatives) are the competitors explanation. Since a successful theory's predictions do not entail the theory, we know that, for any successful theory, there exists in principle many other theories, incompatible with the first yet sharing all the successful predictions. The truth of any of these competing theories would explain the truth of the predictions just as well as the truth of our original theory, since they all entail those predictions. So the truth of our theory is not the *best* explanation of its success and so by our own standards ought not to be inferred.

How threatening are the selection and competitors explanations to the inference from a theory's predictive success to its truth? In spite of what it sounds like, inference to the best explanation does not require that one infer only one explanation, so long as the explanations inferred are compatible. And the selection explanation is compatible with the truth explanation: a theory may both have been selected for its predictive success and be true. Nevertheless, one may feel that the existence of the selection mechanism undermines the inference from success to truth, much as noticing that my computer is unplugged undermines my reason for inferring from its lack of behaviour that it is broken. From the point of view of the tracking requirement, however, it appears that the selection explanation does not undermine the inference to the truth explanation. The truth of the theory's predictions either track the truth of the theory or they do not, and the fact that we ended up with this theory as a result of a selection mechanism does not affect the tracking question. The situation is quite different in the computer case. There the fact that the computer is unplugged means that my evidence (e.g. nothing appearing on the screen) does not track the hypothesis that the computer is broken. Being unplugged, the computer would display the same lack of behaviour whether or not it was broken.

Unfortunately, the competitors explanations are more of a threat. One might say that the truth of the predictions either track the truth of our theory or not, irrespective of the existence of competing theories, but this is cold comfort, since their existence appears to undermine any basis for saying that it is *our* theory that the evidence tracks. The point as I see it is not that all theories that share certain predictions are equally well supported by those predictions. Our inductive practices and especially our judgements of tracking are far more nuanced than this, taking account of many other sorts of information and background belief. The miracle argument, however, effaces all this additional information, by describing an inference simply from a theory's success to its truth, without regard for the theory's content. At this extremely low level of resolution, there is nothing to choose between the truth of our theory and the truth of the competing theories.

To avoid this consequence, we may re-interpret the miracle argument, taking it as a summary of the first-order scientific evidence for our theory, rather than as a distinct but extremely general inference. In this case, the inference may be an excellent one, insofar as the scientific evidence tracks the correctness of the theory. If we do have tracking evidence for our theory, however, this will not be simply because that evidence is a consequence of the theory: that is what the competing explanations show. More importantly for present purposes, whatever the value of the evidence, the miracle argument adds nothing to it and so gives no independent or additional reason for realism. Nor does it give any reason to say that the first-order evidence tracks or that the inference based upon it should be construed realistically.

Here we find an important disanalogy with the inductive justification of induction. Before considering the track record of our inductive practices, we have no reason to say that induction will work in unexamined cases. When we take on the past success of our practices as evidence, however, we may infer future success on that basis, and this inference may be strong, since its premise may track its conclusion. By contrast, before the miracle argument we already have the first-order evidence for our theory, and the introduction of the miracle argument adds nothing to that evidence.

To remove this disabilitating disanalogy, we might historicise the miracle argument into an argument that infers the truth-status of our current theories from premises about the success of *previous* theories. There is a well-known second-order induction of this form in the philosophical literature, but it is used against the realist, not

for her. This is what is sometimes called the pessimistic meta-induction but which I have chosen to call the disaster argument, and it is to this argument that I now turn.

VI

The Disaster Argument

According to the disaster argument, past scientific theories have been found to be false, so present and future theories are probably false too. Like the miracle argument, the disaster argument is tantamount to an inductive assessment of the inductive methods of science, but here the assessment is negative. But this too is a weak induction. It may be that, like the miracle argument, the disaster argument fails to bring substantially new data to bear on the issue, since the falsity of past theories in a given area of science would follow simply from the fact that those theories form a series of logical contraries, on the assumption that the last member of the series was true. The argument will also fail to satisfy the tracking requirement, since even if present theories were true, the past theories would still have been false. Another way of bringing out the weakness of the disaster argument is to see that the history of science is just the sort of dynamic system over which enumerative inductions are unreliable.

One of the main lessons of the history of science, according to proponents of the disaster argument, is that theories have a sell-by date. Indeed in light of present science, virtually all theories more than say 100 years old are seen to be, strictly speaking, false. In many cases, the conflict between present and past science is pronounced. Past theories appealed to entities, processes and relations whose existence we now deny. No more crystalline spheres, humours, phlogiston, caloric, ether; no more earth at the centre of the universe or mass independent of velocity. Some of these past theories were, moreover, predictively successful in their day, which places another nail in the coffin of the miracle argument. But the disaster argument itself is more direct than this. Stripped down and unqualified, it is a simple enumerative induction: all past theories have been false, so all present and future theories will be false as well. On this pessimistic view, scientists are like guileless cousins of the boy who cried wolf: they keep claiming the correctness of their latest theories, but those theories keep turning out to be wrong.

The miracle argument and the disaster argument are both inductive, but where the miracle argument projects from truth of part of a theory (the successful predictions) to the whole, the disaster argument projects from the falsity of some theories to the falsity of others. In this respect the miracle argument is closer in structure to the

inductive justification of induction: both are enumerative arguments inferring from the outcome of past inferences to predict the outcome of future inferences. Of course where the inductive justification of induction projects success, the disaster argument projects failure.

Most of the standard responses to the disaster argument are strikingly concessive. A common realist reaction is to retreat from truth to verisimilitude. By claiming only that our best theories are approximately true, we may admit both that past theories are false and that from this we should infer that future theories will probably be false as well -- that is, we may admit both the premise and the conclusion of the disaster argument -- while still retaining the realist's attachment to the truth. Although realism may well require a notion of verisimilitude or approximate truth, however awkward its analysis, this response seems to me at once both to underestimate how badly wrong, by present lights, many past theories have been, and to overestimate the force of the disaster argument. Another common response to the disaster argument is to retreat to a form of semi-realism, according to which we should commit ourselves only to the truth of those aspects a theories that have shown marked stability over the history of science. This response is also problematic, partly because it is quite unclear what these stable elements are. Entities and abstract structures are the most common candidates, but neither seems in general to have been suitably stable. More importantly, from my point of view, this response again concedes too much to the disaster argument. Its retreat to the least common denominator is a form of conservatism that is epistemically indefensible. Suppose we could divide our current theory into an old (i.e. stable) part and a new part. What reason to we have to place more trust in the old part? The old part has not therefore been better tested, and the new part is supposed to mark the epistemic advance on what has come before.

The respect with which the disaster argument is commonly treated by realists is somewhat surprising, given its tricky structure. (It also suffers from a false premise -- not all past theories are false by present lights -- but I will not depend on this in what follows.) Although it has the superficial form of a simple enumerative induction, the disaster argument has a number of peculiar features. For example, it seem to be an exercise in what might be called judo epistemology, in honour of my childhood judo teacher, who claimed that judo would enable me to use my opponents' strength against them. (This is a phenomenon I never managed to elicit.) From a realist perspective, the disaster argument seems an attempt to use the great progress of science against itself.

We believe that past theories are false because we take present theories to be true, but then the disaster argument uses the admitted falsity of past theories to undermine our claims about the truth of present theories. A similar point comes out at the evidential level. This can be seen most simply if we imagine an idealised `crucial' experiment, where the evidence simultaneously refutes the past theory and supports the current one. The disaster argument has the effect of transmitting the negative force of the refutation forward onto the current theory, so that in this case we have a strange situation where the very same evidence both supports (directly) and undermines (indirectly) the very same theory. The only `evidence' that the disaster argument relies upon is the observation that successive theories have been contraries, whereas the issue is what this shows about what science may achieve.

From the point of view of the tracking requirement, the concessive responses to the disaster argument seem wildly inappropriate, because the disaster argument is a weak induction. The argument would have us infer that present theories are false because past theories are false. Tracking requires of a strong argument that, had the conclusion been false, the premises would have been false as well. In the present case, this comes to the requirement that, if present theories had been true, past theories would have been true as well. This conditional is obviously false, however, since present and past theories are contraries. If present theories were true, past theories would have to be false. So the falsity of past theories gives no inductive grounds for believing that present theories are false, and the disaster argument fails.

The tracking requirement thus appears satisfyingly to destroy the disaster argument in one blow. Alas, things are not quite this simple, though we are on the right track. For consider the case of testimony. The fact that you have told me falsehoods in the past may be a good reason for me not to trust you this time. And this may obviously (perhaps even especially) be the case if your present testimony contradicts your past testimony. Yet in such a case, if what you just said were true, your former testimony would have still had been false. Here it is easy to see what has gone wrong. The tracking requirement has been misapplied, by holding the content of the testimony constant. The reason that your past unreliability is good evidence that you will be unreliable in future is that, if you weren’t going to lie in future, you wouldn’t have lied in the past: you would have made different and true statements.

Thus the defender of the disaster argument may in desperation maintain that the tracking requirement is in fact met, since if our current theories were true, we would have

come up with different, true theories in the past too. Desperate, because this seems false. To take a simple case entirely in the past, consider the theory that the planets move in ellipses, in the time when Kepler propounded it. If that theory had been true, would scientists have propounded different theories before Kepler? It is difficult to see why. Presumably, if that theory had been correct, Kepler would have had the same data he actually had, since those data supported the elliptical hypothesis. And I see no obvious reason why earlier scientists would have had different data, or anyway different data that would have led them to different hypotheses. So it seems that if the latest in a series of false theories had been true, the earlier members of the series would have been the same, and the tracking requirement remains unmet.

If a defender of the disaster argument wants to argue that the tracking requirement is met, she should I think argue that past theories would have been different not because the available data would have been different, but because the scientists would have been different: they would have been better at doing science. This comes out naturally if we recast the disaster argument into the form of a pessimistic inference to the best explanation. The comparison with testimony is again instructive. The inference to the unreliability of an informant is naturally seen as having two steps, the first of which is an inference to the best explanation. From past unreliability we infer that the speaker is dishonest or incompetent, from which we infer that future testimony will be unreliable too. Similarly, though honesty is not here at issue, the defender of the disaster argument will quite naturally describe the inference as moving from past falsity to incompetence to future falsity. Our scientific track record gives us good reason to believe that we are just not very good at doing science. After all, if we were any good at it, we would have gotten something right by now. And our incompetence gives us good reason to believe that we will not get things right in future either. It is because the inference from past to future falsity is mediated by incompetence that the tracking requirement is met. We would only get science right in future if we were competent, and if we were competent some of the previous theories we would have generated would have been correct as well.

This is a stronger version of the disaster argument than the simpler enumerative induction from past to future falsity. Nevertheless, it fails to meet the tracking requirement for two reasons, which correspond to the two steps in this version of the disaster argument, the inference from past falsity to past incompetence, and the inference from past incompetence to future incompetence. Past failure is not a sign of

incompetence, but of science working by its proper quasi-evolutionary methods in a way that may take us to the truth. Past incompetence would not in any case be good evidence for present or future falsity, because of all the inductively relevant differences between past in present science. Both the continuities and the discontinuities in the history of science militate against the disaster argument.

What undermines the inference from past falsity to general incompetence in science is what we know about how science develops. We know that this process is in some ways akin to natural selection, with theories that are generated and selected. At least to this extent, Popper was right: theories are generated and then replaced by contraries. Consequently, we should expect to find a stretch of a contrary series of false theories whether or not current or future theories are true, whether or not the practices we have for investigating the world will take us to the truth. This means that the inference from past falsity to incompetence is unwarranted. It does not track, because we would have past falsity even if we were competent. The selection mechanism that drives science thus blocks the inference from falsity to incompetence. The disaster argument is a weak induction because it suffers from yet another of the general inductive flaws I mentioned near the start of this paper, the flaw of extrapolating for systems we know to be dynamic and developing in relevant ways. This differentiates the scientific case from the case of the unreliable informant. We have no reason to believe that people who keep telling us falsehoods are doing so because they are committed to a method of conjecture and refutation, so there is nothing to block the inference from consistently false testimony to incompetence or dishonesty.

I want to consider two objections to this `Darwinian' reply to the disaster argument. The first is that, even with a Darwinian mechanism, a series of theories that will eventually yield some true theories should already have done so. The fact that all past theories have been false thus remains a good reason for saying that future theories will be false too. A history of unrelieved falsity would not however be surprising early on in a series that eventually yielded truths, and if we consider the amount of science produced in this century (however this might be measured) as compared with everything that has come before, and if in addition we project this growth into the future, there is some basis for the claim that what has come so far is only the beginning. In a way, this is the core question: if we were going to reach the truth, would we already have reached it? It is far from clear that the answer is yes, and in any event this is what the disaster argument effectively assumes, whereas it is what that argument needs to show.

The second objection is that to rebut the disaster argument with the Darwinian objection is to move briskly from the frying pan to the fire. A Popperian mechanism of conjecture and refutation may effectively block the inference from past falsity to future falsity, but the cost is a straightforward scepticism about science, since there can be in Popper's system no reason ever to believe that any theory we produce is correct, since such an assessment would depend on a kind of inductive inference, and induction is what Popper abjures. This is I think a fair comment on Popper's own position, and one that he would have accepted. (One might add that, from an *echt* Popperian perspective, the disaster argument never gets off the ground, since it is of course an inductive argument and all inductive arguments are declared worthless.) But to agree with Popper as I do that science relies in part on a process of conjecture and refutation, suitably articulated, is not to agree with his wholesale rejection of induction, something I obviously do not accept. Eliminative induction is induction nonetheless, and does not require one to say that knowledge is impossible. The issues here are too complex for me to develop in this paper, but we may have to reach a compromise. What I have in mind is that the negative methodology may indeed block any straightforward inference from predictive success to the truth of high-level scientific theories, but may yet allow the view that the mechanism is truth-tropic over the long term. For present purposes, it is enough that the Darwinian mechanism blocks the disaster argument: we must leave the exploration of its further epistemic consequences to another occasion.

The Darwinian objection to the disaster argument exploits a feature of scientific practice that remains constant across its history: its negative methodology. A final objection I wish to make against the disaster argument focuses instead on how science changes. Inductive arguments are unsafe when the predicted cases are relevantly different from the evidential cases. It is for example unsafe to infer from the fact that all observed British ravens have been black to conclusions about the coloration of the elusive Arctic raven. The tracking condition reflects this feature of good inductive practice, since where there are relevant differences in general the data would have remained the same even if the conclusion had been false. Thus if Arctic ravens had been white, British ravens would still have been black, since the reason that Arctic ravens would have been white would have concerned their local environment.

There obviously are relevant differences between past and present science, where relevance here is relevance to truth-value. Thus we have much more data supporting current theories than we had for past theories. We have better instrumental

technologies. We have the knowledge of the weaknesses of past theories, which we have exploited in order to construct current theories, and we have often saved the best bits. These sorts of consideration hardly show that present theories are true, but they do show that the inference from past to future falsity is unsafe. Even if we had been incompetent at doing science in the past, it would be hasty to infer that we will not learn from our mistakes. So my assessments of the miracle argument for realism and the disaster argument against realism are both negative. These second-order inductions are weak. At the same time, I have argued that other second-order inductions are strong, and that the traditional inductive justification of induction may be among them.

Acknowledgements

I have discussed earlier versions of various parts of this essay with diverse audiences at the Universities of Cambridge, Edinburgh, Leeds, Nottingham, Sheffield and Tulane, The London School of Economics, Birkbeck College, The Belgian Society for the Philosophy of Science and The British Society for the Philosophy of Science. I am very grateful for the many and helpful questions, comments and suggestions I received.

REFERENCES

Braithwaite, R., 1953, *Scientific Explanation* (Cambridge: Cambridge University Press), reprinted in Swinburne, R. (ed.), 1974, *The Justification of Induction* (Oxford: Oxford University Press).

Descartes, R., 1641, *Meditations on First Philosophy*, Donald Cress (trans., 1979) (Indianapolis: Hackett).

Fine, A., 1984, `The Natural Ontological Attitude', in J. Leplin (ed.), *Scientific Realism* (Berkeley: University of California Press).

Goodman, N., 1983, *Fact, Fiction and Forecast*, 4th edn. (Indianapolis: Bobbs-Merrill).

Hume, D., 1739, *A Treatise of Human Nature*, L.A. Selby-Bigg & P.H. Nidditch (eds., 1978) (Oxford: Oxford University Press).

Hume, D., 1777, *An Enquiry Concerning Human Understanding*, L.A. Selby-Bigg & P.H. Nidditch (eds., 1975) (Oxford: Oxford University Press).

Laudan, L., 1984, `A Confutation of Convergent Realism', in J. Leplin (ed.), *Scientific Realism* (Berkeley: University of California Press).

Lipton, P., 1991, *Inference to the Best Explanation* (London: Routledge).

Nozick, R., 1981, *Philosophical Explanations* (Cambridge MA: Harvard University Press).

Putnam, H., 1978, *Meaning and the Moral Sciences* (London: Routledge).

Quine, W.V. & Ullian, J.S., 1978, *The Web of Belief*, 2nd edn. (New York: Random House).

Van Cleve, J., 1984, `Reliability, Justification and the Problem of Induction' in French, P., Uehling, T. & Wettstein, H. (eds.) *Midwest Studies in Philosophy IX* (Minneapolis: University of Minnesota Press) 555-67.

Van Fraassen, B., 1980, *The Scientific Image* (Oxford: Oxford University Press).