Anthony O'Hear (ed) *Karl Popper: Philosophy and Problems* (CUP, 1995) 31-44.

Popper and Reliabilism

Peter Lipton

Karl Popper attempted to give an account of scientific research as the rational pursuit of the truth about nature without any appeal to what he took to be the fictitious notion of non-demonstrative or inductive support. Deductive inference can be seen to be inference enough for science, he claimed, once we appreciate the power of data to refute theory. Many of the standard objections to Popper's account purport to show that his deductivism actually entails a radical scepticism about the possibility of scientific knowledge. Some of these objections appear unanswerable in the context of the traditional analysis of knowledge as justified true belief; but this is neither a conception of knowledge that Popper himself accepted nor one that is currently in fashion. Reliabilism, the view that knowledge is a true belief generated by a reliable method, is now a popular replacement for the traditional analysis and one that is closer to Popper's own conception of knowledge. My aim in this essay is to consider in brief compass the prospects of a reliabilist reading of Popper's account of science. Such a reading makes it possible to turn some of the standard objections helps to show which of Popper's views should be accepted and which rejected.

The Standard Objections

Popper's philosophy of science is naturally seen as a radical response to Humean scepticism about induction. According to the sceptical argument, no form of non-demonstrative reasoning is rationally defensible, since any argument to show that such reasoning is generally truth-preserving or reliable would itself need to be a non-demonstrative argument and would hence beg the question. Our inductive practices have presumably been at least moderately reliable up to now, else we would not be here to consider the problem, but what needs showing is that they will continue to be reliable in future. The claim of future reliability is however a prediction that could only be justified inductively. What we have observed hitherto does not deductively entail that induction will work in future, but to give an inductive justification for that prediction is to argue in a circle. Our situation seems analagous to that of a party of hikers who could only get across a wide chasm with the aid of a rope neatly coiled on the far side.

Hume's own response to his sceptical argument is that inductive inferences are rationally indefensible but psychologically unavoidable. He goes on to give a description of the what he took the psychological mechanism behind them to be, a process of Pavlovian conditioning or habit formation. Most epistemologists have instead rejected Hume's scepticism about induction and have attempted to show by more or less desperate means what is wrong with his remarkably resilient argument. Popper, by contrast,

simply accepts the sceptical argument: induction is irrational. Unlike Hume, however, he does not retreat to the descriptive psychological project. Instead, Popper sets out to show that, scepticism about induction notwithstanding, scientific inquiry is rational, by showing that, appearances to the contrary, scientific inferences are purely deductive.

Popper's central idea is that although the scientific evidence never entails that a theory is true, it may entail that the theory is false. If we have a hypothesis of universal conditional form, no number of positive instances will entail that the hypothesis is true, but a single negative instance will entail that it is false. No number of black ravens entails that the hypothesis that all ravens are black is true, but a single white raven entails that the hypothesis is false. More generally, if a theory entails a prediction and the prediction is found to be false, then the theory must be false as well, since to say that the argument from theory to prediction is deductively valid is just to say that, if the conclusion is false, at least one of the premises must be false as well. Scientists can thus know that a theory is false, without recourse to induction. Moreover, faced with a choice between two competing theories, they can exercise a rational preference with respect to the goal of discovering the truth, if one of the theories has been refuted but the other not, since it is rational to prefer a theory that might be true over one known to be false. Induction never enters the picture, so Hume's argument is defused.

Popper thus seeks to accept Humean scepticism about induction without accepting scepticism about science. Many of the standard objections to Popper's position attempt to show that he cannot have it both ways: insofar as his account really does abjure induction, it makes scientific knowledge impossible. Let me remind you of four of the most familiar of these objections. Firstly, according to Popper scientists are never justified in believing that the observation statements they accept are true. Scientists adopt certain procedures for accepting data that they hope will lead to accepting mostly true statements, but a hope is not a reason. The only thing that could justify an observation statement would be the scientist's experience but, according to Popper, only statements can justify statements. Consequently, if an accepted observation statement contradicts a hypothesis, we are not justified in claiming that the hypothesis is false, only that either the hypothesis or the observation statement is false. On a traditional conception of knowledge, this makes it impossible to use the mechanism of empirical refutation to gain knowledge that any hypothesis is false. The data cannot be known to be correct and as a result theories cannot be known to be false. This is the problem of unjustifiable data.

The second standard objection is that knowledge of the falsity of a theory remains impossible even if we grant knowledge of the truth of the data. This is the problem of the auxiliaries. Theories do not entail predictions outright, but only with the aid of various often ill-defined auxiliary assumptions. From a logical point of view, what follows from the falsity of

the entailed prediction is only that at least one of the premises is false, not which ones. Consequently, since the premise set includes statements apart from the theory under test, knowledge of the falsity of conclusion does not make it possible to know that the theory is false.

The third objection is the problem of application. Even if we granted that the mechanism of refutation made it possible to know that certain theories are false, Popper's account still does not make possible the rational application of science. In particular, we would no reason to prefer for practical purposes the predictions of unrefuted theories over those of their refuted rivals. The source of the problem is the inapplicability of Popper's argument for rational preference between theories to preference between their predictions. According to Popper, we are to prefer the unrefuted theory, because it may be true whereas its refuted rival is known to be false. Every set of false statements has however indefinitely many true consequences, so we cannot say that the prediction from the false theory is itself false, and the basis for preference is lost. It is impossible to know that any of a theory's predictions are true, or that the prediction of one theory is more likely to be true than the incompatible prediction of another.

The final objection is that Popper's account provides no reason to believe that science is moving towards the truth. This is the problem of the bad lot. However assiduously scientists may eliminate and replace false theories, there is no reason to believe that the new theories are better than the old ones. Perhaps all the theories we will ever generate are false, and perhaps from among those it is the ones we fail to eliminate that are furthest from the truth. It is impossible to know that later theories are better than the ones they replace.

If these objections are sound, Popper's philosophy entails a profound scepticism about science. Nor is this something a Popperian ought to accept with equanimity: we must not confuse scepticism with fallibilism. A fallibilist account of knowledge is the ground we all wish to occupy: neither theory nor data are ever certain. Hume, however, was no mere fallibilist about induction. He did not claim that the conclusions of inductive inferences are uncertain: he claimed they were epistemically worthless. Similarly, if the objections to Popper's philosophy are sound, they show that the results of science are worthless, not merely uncertain.

Popper was well aware of the objections I have mentioned, but his replies are unsatisfying. His response to the problems of unknowlable data and of the auxiliaries is, to put it baldly, that scientists should pretend they do not exist. Scientists should pretend that the observations statements they accept are known to be true, and they should pretend that a failed prediction refutes the theory directly. Even if this were good advice to scientists, which is doubtful, it does not meet the epistemic difficulties the objections raise. Popper's reply to the problem of practical preference is to claim that, if we must rely on any theory, it is rational by definition to rely on the best-tested

one, which is one that has not been refuted. But if we are careful to free the notion of `best-tested' from any inductive associations, this semantic solution is no more satisfying than the parallel semantic solution to the Humean problem, according to which using induction is part of what we mean by acting rationally. In both cases, the natural response is that, if this is what `rational' means, what we care about is not being rational, but being right. Finally, Popper's reply to the problem of the bad lot seems to be that we may hope that his method of conjecture and refutation takes scientists towards the truth and that there is no better alternative. But scepticism is not avoided by calling it unavoidable, and it is unclear on what basis Popper could argue that his method is any more likely to generate true theories than random guessing.

Reliabilism

The standard objections to Popper's account of science have considerable force. Each of them casts serious doubt on the power of a Popperian methodology to generate scientific knowledge. In each case, the argument for the impossibility of knowledge is based on the impossibility of justification. This suggests is that at least some of the objections might be turned by using an analysis of knowledge that does not depend on the notion of justification, a thought that fits well with Popper's general hostility to such a notion. In recent epistemology, the most discussed such analysis is reliabilism. Can Popper's philosophy of science be improved by combining it with a reliabilist theory of knowledge?

Reliabilism was not originally developed as a response to Humean scepticism, though it was later so applied. Instead, it was motivated by the thought that having a justification seems neither necessary nor sufficient for a true belief to count as knowledge. Perceptual knowledge is the obvious example of a kind of knowledge that does not require justification, at least if justification is understood as explicit argument. Over-intellectualising philosophers to the contrary, when I see that my pen has fallen on the floor, my knowledge of its present location is not based on argument or inference. Even cats and dogs have perceptual knowledge, though they are incapable of giving reasons. Other plausible examples of knowledge without justification include knowledge from memory and, at least for humans, from testimony. There can be knowledge without justified true belief.

Conversely, the standard source of examples for justified true belief without knowledge are the famous Gettier cases. Gettier constructed a conceptual machine for generating examples of justified true beliefs that are not knowledge. The machine works simply by deducing true beliefs from justified but false beliefs, capitalising on the logical point, already noted above, that every false statement has innumerable true consequences. These derived true beliefs will not in general be cases of knowledge. For example, suppose I had the justified but, as it happens, false belief that my wife took our car to London for the day, from which I deduced that it would not be at

our house when I came home. The deduced belief was true as it happens, but only because our car was stolen during the day. My belief that the car would not be at home was true and justified, but not a case of knowledge.

According to reliabilism, the justification condition that a true belief is knowledge just in case one has good reasons for the belief should be replaced by the condition that the true belief was generated by a reliable method or process. Knowledge is reliably produced true belief. Different reliabilists have analysed the notion of a reliable method in different ways. For example, some have understood a method to be reliable just in case it tends to produce true beliefs, others have construed reliability in terms of various counterfactuals concerning the resultant belief, for example that one would not have held the belief, had it been false. Not all true beliefs are knowledge, since it may be a matter of luck that the belief is true, but if the truth was reliably produced it counts as knowledge, even if the reliable method did not involve argument or justification. Thus if our perceptual mechanisms or those of some other animals reliably generate true perceptual beliefs, those beliefs count as knowledge. And a justification may fail to generate knowledge by failing to satisfy the appropriate reliability condition. I did not know that our car would not be at the house, in spite of having a good reason to believe this, because it was a fluke that my belief was true, and this is reflected in the fact that I would still have had that belief if the car had not been stolen.

How best to articulate the notion of reliability suitable for the analysis of knowledge is an interesting and difficult question which, though clearly central to the development a reliabilist theory of knowledge, is one I hope to avoid in this essay. There are, however, three general features of reliabilist accounts of knowledge that are worth emphasising. The first is that reliabilism does not drop the truth condition on knowledge. To say that a method is reliable is to speak of its propensity to generate true belierfs, not merely beliefs that are useful or otherwise attractive, whatever the other connotations of the term `reliable'. Secondly, what counts for knowledge is reliability in fact, not having reasons to believe that one's methods are reliable. The cat knows birds when it sees them, but of course can give no reason to believe that its visual system is a reliable bird-detector. Finally, although reason and inference are not required for knowledge, there is such a thing as inferential knowlege. What makes this knowledge, according to the reliabilist, is just that the methods of inference employed are in fact reliable. Inference is one set of methods for generating beliefs among others.

Although not constructed for this purpose, reliabilism offers a quick non-Popperian solution to scepticism about inductive knowledge. Suppose that Hume is right in claiming that there is no possible justification for the claim that our inductive practices are reliable. Even so, from a reliabilist point of view, the impossibility of knowledge does not follow. For inductive knowledge to be possible, it must be possible that our inductive practices are

reliable in fact, a possibility Hume does not deny. Certainly he could not argue that our inductive practices will be unreliable, since this would be a prediction that could itself only be justified inductively. So inductive knowledge is possible and if, as we all believe, induction is at least a moderately reliable method of acquiring beliefs, then it is actual as well.

Knowing what is False

My topic in this essay, however, is not how reliabilism may vindicate induction, but rather how it may vindicate Popper's emphasis on negative methods, by helping to turn the four standard objections to his position. Let us begin with the first two, the problems of unjustifiable data and of the auxiliaries, both of which threaten the claim that scientists can know that the theories they reject are false. Popper would have liked falsification to have the certainty of proof, as his use of the word `refutation' suggests. Such certainty is possible in the relation of logical incompatibility and Popper would have scientists act as though the rejection of theories has a similar status when motivated by that relation. But on Popper's view scientists do not know that the data they accept are true or that the theory is to blame for failed predictions, so they cannot know that the theory they reject is false. From a reliabilist point of view, however, falsification becomes reliable rejection. If using the deductive relation enables scientists reliably to reject falsehoods and not to reject truths, then they can know what that the theories they reject are false. Neither the unjustifiability of the data nor the presence of additional premises rule this out.

The unjustifiability of data by experience poses no special difficulty for the reliabilist, since knowledge does not require justification. At the same time, experience can be more than a motivation for accepting observation statements, if it is part of the method which causes those statements to be accepted. On this view, the senses are detection devices which generate knowledge insofar as they are reliable. This is not to show that they are reliable, but such a demonstration is not required in order for the data to be known. Reliabilism does not show that scientists do know that their data are correct, but it improves the basic Popperian scheme by showing how such knowledge is possible and by explaining the epistemic relevance of experience, without abandoning the Popperian stricture against the notion of the justification of statement by experience.

Reliabilism also helps with another challenge to the possibility of knowing the data that Popper was among the first to identity, the theory-ladeness of observation. The essential role of theoretical beliefs in the generation of data is no bar knowledge of the data. Of course if the lading theories are false and are so related to the data that this entails that the data must be false as well, then the data cannot be known. But the presence of theories in the generating mechanism is compatible with the reliability of that

mechanism. For example, when scientists use theories of their instrumentation to get data from those instruments, what counts for knowledge is whether the composite mechanism, consisting of both physical and intellectual technology, tends to generate only correct data. The reliability of this mechanism does not depend on a theory-neutral description of the evidence; indeed it may not require that the theoretical components be justified, known or, in certain circumstances, even true.

The reliabilist takes a similar line on the problem of auxiliaries. Where premises apart from the theory under test are needed to deduce a prediction, the falsity of the prediction does not entail the falsity of the theory, but this is no bar to using the Popperian method of refutation to gain knowledge of the falsity of theory. What counts is that the scientists' attribution of blame be reliable: that they tend to blame the theory only when it is to blame. Now Popper has often been criticised for saying that scientists ought generally to blame the theory, but this is a criticism the reliabilist can easily absorb. Scientists have complex and poorly understood ways of apportioning blame, of deciding whether to apply to theory or auxiliaries and to which part of either, but insofar as these practices are reliable, they can yield knowledge of what is false.

The problems of unjustifiable data and of the auxiliaries show that falsification, unlike incompatibility, cannot be understood as a purely deductive relation. A reliabilist account of knowledge shows how one may nonetheless claim that scientists can know that the theories they reject are false. We should now ask to what extent this reliabilist picture of the rejection of theories remains recognisably Popperian. It does appear to be able to respect Popper's proscription on justification. Knowledge of the falsity of a theory does not depend on the justification either of observation statements or premises not rejected. But can the reliabilist respect Popper's asymmetry between confirmation and refutation? Reliabilism is in itself not hostile to induction and, as the problems of the data and the auxiliaries show, neither the reliabilist nor Popper can make sense of refutation as a purely deductive operation. Nevertheless, several important asymmetries between positive and negative methods remain from a reliabilist point of view. To put the matter crudely, the reliabilist can explain why falsification is easier that verification, why it is easier for scientists to acquire reliable methods for determining that a particular theory is false than for determining that a theory is true.

Popper's logical asymmetry is based on a contrast between part and whole. To determine that a universal generalisation is true, we need to know about all its instances, whereas to determine that it is false we need only know about one. Similarly, to determine that a theory is true we need to know the truth value of all its consequences, whereas we need only know that one consequence is false to determine that the theory is false. This asymmetry survives the complications of falsification that we have considered and the

reliabilist response to them. Whatever the method for determining the truth value of a statement, it will be easier to determine the truth value of one than of many, at least where the many includes the one. Similarly, it is easier to construct a reliable method for detecting the truth value of some of the consequences of a theory than of all of them.

There is an additional source of asymmetry between negative and positive methods from the data side that the reliabilist can account for. Inductive support depends on what is known as the `total evidence condition': it must be assessed with respect to all the available evidence. This is so because, unlike deductive reasoning, inductive reasoning is `non-monotonic'. If a deductive argument is valid, it will remain so whatever additional premises are added. By contrast, an argument that we judge to be inductively strong may no longer be so judged when additional data is added. To take a Popperian example, having seen many black ravens I may, if I indulge in induction, infer that all ravens are black, but I will retract the inference if I later see a non-black raven, without rejecting any of my ealier data. This is why inductive assessment is normally only taken to be proper relative to all the available evidence. The situation is different with respect to falsification. Having convinced myself that I have found a genuine counterexample to my hypothesis, I do not need to take into account all the other available evidence: more white raven's won't undo the damage caused by the black ones. From a reliabilist perspective, this provides another reason why we should expect it to be easier for scientists to construct reliable methods of rejection than for acceptance, since it is easier to construct a method that can use limited inputs than one that must accept and assess all the available data in one go. Negative methods can thus be seen to have a double advantage over positive ones, requiring less output and less input to be effective.

Falsification is Necessary for Positive Knowledge

So far I have argued that while Popper's account of falsification makes it impossible for scientists to know that a theory is false on a justificationist theory of knowledge, a reliabilist account makes such knowledge possible and also helps to explain why it should be easier to falsify than to confirm. What I want now to show is that falsification is not just possible but also necessary if there is to be any positive scientific knowledge. As we will see, any reliable method of scientific discovery depends on the reliable elimination of false hypotheses.

Given that reliabilism in no way excludes inductive inferences, one might wonder why there couldn't be a method yielding positive scientific knowledge that did not depend essentially on the elimination of falsehood. The history of science provides overwhelming evidence that science does in fact depend on elimination, but there are also reasons of principle why this should be so. One is a central Popperian theme, the impossibility of inductivism. Most scientific theories appeal to entities and processes not

mentioned in the evidence: those theories are not simple extrapolations and interpolations. Hence there is no algorithmic route from data to theory. Moreover, again as Popper has emphasised, a theory usually needs to be generated before collecting the data that best tests it, since the scientist can often only tell which data are relevant in light of that theory.

There is a world of difference between rejecting inductivism -- the view that there is a mechanical procedure for moving from available data to theory best supported by that data -- and rejecting the possibility of induction or non-demonstrative support. Popperians have not always been careful to distinguish these two things. Nevertheless, the impossibility of inductivism and the temporal priority of theory to data do explain the necessity of negative methods.

If there is no mechanical route from data to theory and theories must be generated before scientists can find the data that would test them, there is no non-miraculous way for scientists to generate only true theories. Hence scientists must rely on methods of elimination, however much they may also depend on inductive support. Moreover, since the method of generation is bound to throw up some false theories, it must be designed so that the false theories that are generated can be reliably eliminated, if positive knowledge is to be possible. This is the reliabilist's version of Popper's falsifiability requirement on scientific theories.

Negative methods are thus an essential feature of any way of doing science that would yield positive knowledge. Indeed there may even be more of a role for negative methods in science than Popper's own account suggests, because of the constant use of negative filters in hypothesis generation. This is supressed in Popper's discussion, because of his artificial separation of the contexts of generation and evaluation and his consequent neglect of the process of theory-generation. Although theories usually need to be tested after they are proposed, for the reasons just given, there are strong constraints on generation, including just the same mechanisms that later lead to the elimination of theories that are generated. Many theories never make it to the testing-stage, because the scientist who thinks of them can eliminate them on already available information. Theories that nobody generates can in one obvious sense not be eliminated, but many of the constraints on generation are equivalent to those of elimination.

Falsification is Sufficient for Positive Knowledge

Reliable falsification is thus necessary for positive scientific knowledge; I now want to suggest why, perhaps more surprisingly, it is also sufficient. Let us suppose that scientists' eliminative methods are in fact reliable, yielding knowledge that various theories are false. What follows from this? Recall the problem of the bad lot: if all the theories we generate are false, then eliminating some of them won't leave us with the truth. Could not scientists be perfectly reliable falsifiers, never rejecting what is true, yet

be quite hopeless at generating truths? They would then have plenty of negative but still no positive knowledge, which would still be a strong form of scepticism about science. At first glance, this possibility appears to remain open, but this is an illusion: if scientists are reliable falsifiers, they must also be getting at the truth.

The connection between negative and positive knowledge is hidden on a naive deductivist picture. If scientists had some theory-independent way of determining the truth of the data, and some of those data contradicted the theory outright, then they could reliably eliminate some false theories even if they had no way of determining which theories are true or indeed even if none of the theories scientists generate are true. Popper never thought this was our situation, but his tendency to have us pretend as if it were has suppressed the connection between negative and positive knowledge.

For falsification to be reliable, scientists must know that the data are true: the method of data acceptance must itself be reliable. But given the extent to which the method of accepting data depends upon background theories, and given the extent to which the data are, as Popper emphasised, theory-laden, reliable data acceptance is only possible if our background is largely correct.

The same conclusion follows from the role of auxiliary statements in prediction. Unless those premises were largely true, a practice of laying blame on the theory under test would not be reliable. In practice, the situation is more complex since, as we have already noted, scientists often do not blame the theory for a failed prediction. This however strengthens the present point. Scientists' selective attribution of blame could not be reliable unless they were reliable judges of when the background is correct and when it isn't. To put the point in general terms, elimination requires not just that scientists accept data but also that they accept theory, and the elimination can only be reliable if the acceptance is.

This argument does not show that science does take us to the truth: it does not solve Hume's problem. What it does purport to show, however, is that knowledge of which theories are true and knowledge of which theories are false are intimately connected. Popper may have been over-optimistic about scientists' ability to detect falsehood, but he was also over-pessimistic about what follows from this about the ability to detect the truth. Scientists cannot be good a detecting one without being good a detecting the other.

What then of the asymmetry between negative and positive knowledge? If what I have just claimed is right, and scientists could not know what is false unless they also knew what is true, then how could what I have said earlier be right, that it is easier to determine that a theory is false than that it is true? The answer lies in a difference of scope or scale. It is in general easier for scientists to show that a particular hypothesis is false than that a particular hypothesis is true, because showing falsehood requires the determination of the truth value of only a part of the hypothesis and because it

requires only the use of part of the available evidence. Nevertheless, a method of falsification can be generally reliable only if there are also in use reliable methods of generating true beliefs concerning both data and theory.

Whence Induction?

I have suggested three main ways that a reliabilist account of knowledge may help to vindicate the Popperian emphasis on a negative methodology of science, in the face of various standard objections to Popper's position. Reliabilism shows how falsification is possible, why it is necessary for positive scientific knowledge, and why it is sufficient. Can it, however, be used to defend Popper's wholesale rejection of induction? I think not.

Induction is just non-demonstrative inference. It cuts across the distinction between acceptance and rejection, between positive and negative methods. Some acceptance, such as acceptance of of observation statements, may be non-inductive because non-inferential. Some rejection may be non-inductive because, though inferential, the falsity of the claim is strictly entailed by the truth of the data. But at least most and perhaps every case of the acceptance or rejection of a scientific theory involves non-demonstrate inference. The scientist has reasons for the judgement, and those reasons are not conclusive.

There is thus no reliable route to falsification that does not use induction. Nor can scientists do without inductive methods that yield positive results else, as we have seen, their negative methods would not be reliable either. Moreover, there is in my view no adequate response to the problem of application that does not concede positive inductive argument. To deny that scientists ever know that any of the unobserved consequences of their theories is scepticism. To accept that such knowledge is possible is to accept inductive inference to a prediction, however mediated by the technique of conjecture and refutation.

Induction is unavoidable, so Popper's solution to the problem of induction fails. But Popper's case against inductivism stands, as does his emphasis on the importance of a negative methodology. Reliabilism shares Popper's focus on the search for truth, takes the unavoidability of induction in its stride, and explains why Popper was right to put such weight on the role of theory elimination in science. It also brings out the way
the negative and positive results come together in science. If science generates knowledge at all, it can only do so by determining what is false, but if it can determine what is false, it can also determine where the truth lies.

Peter Lipton is a Lecturer in the Department of History and Philosophy of Science, Cambridge University, and a Fellow of King's College. He is the

author of *Inference to the Best Explanation* (Routledge, 1991).

-The method of difference is a good example of how accept and reject combine.

[From beginning of `False is possible'] In particular, I will argue that a reliabilist perspective enables us to see how failed predictions enable scientists to know that a theory is false, why the method of falsification is indispensible to the growth of positive scientific knowledge, and how scepticism about science can be avoided by showing that, if we can come to know which theories are false, we must also be able to know which theories are true. [Polish this later.] [Add that I will discuss baby/bathwater worry.]