JOHN WORRALL

SCIENTIFIC DISCOVERY AND THEORY-CONFIRMATION*

1. INTRODUCTION

Although I find most recent challenges to older 'positivistic' views in philosophy of science either unchallenging or unconvincing, there is one respect in which the new 'post-positivists' are, I believe, definitely right and the older 'positivists' definitely wrong. Reichenbach, Carnap, Popper and others all agreed that philosophy of science is exclusively concerned with the logical analysis of the merits of theories already 'on the table'. Of course, these thinkers were ready to allow that the question of how a theory arrived on the table could be a fascinating one, but they held that it was a question of no interest to a philosopher as such. In particular, to hold that the origins of a theory have any relevance for the appraisal of its scientific merits was, according to these philosophers, to commit one form of the 'genetic fallacy'.

While I agree with the 'post-postivists' that this thesis is wrong, I also agree with Laudan, who, in a recent and hard-hitting article (Laudan 1980), claimed essentially that, while the 'post-positivists' have produced a good deal of bluster, they have produced no really solid argument for just why considerations of theory-discovery should have any relevance for any specifically philosophical issue. I shall try here to supply some solid argument for this claim.

In fact I shall argue the following specific *thesis*: in order to decide whether a particular empirical result supports or confirms or corroborates a particular theory the way in which that theory was developed or constructed needs to be known – more especially, it has to be checked whether or not that empirical result was itself involved in the construction of the theory.[1] If this thesis is correct then it follows, of course, that heuristic questions are indeed of central philosophical significance - for the question of the rationale for scientific change is surely a philosophical one, and any adequate rationale for scientific change is surely bound to give a substantial role to consid-

J. C. Pitt (ed.), Change and Progress in Modern Science, 301–331.

erations of empirical support. So, if indeed empirical support, probably *the central* notion from the 'context of justification', is discovery-dependent, then justification and discovery are not separate but rather interdependent contexts.

2. THE 'HEURISTIC VIEW' OF EMPIRICAL SUPPORT

When does a piece of empirical evidence support a scientific theory? This clearly has a lot to do with the logical relation between the theory and the evidence. Ideally, the evidence statement will assert that certain initial conditions held and that the experiment or observation had a certain outcome *and* the theory together with the statement of initial conditions will logically entail the outcome statement. Equally clearly, support or confirmation cannot be *simply* a question of whether or not the theory entails the evidence (I shall, from now on, talk about the theory entailing the evidence as shorthand for the above more complicated statement concerning initial conditions and outcome). For one thing, a theory which entails a given piece of evidence e could be created quite trivially by taking any theory T of which e is independent and forming a new theory T' by simply conjoining T and e. T' obviously entails e but surely no one would want to say that it is confirmed by it – at least not if the judgment is meant to carry with it any consequences, no matter how tentative or hedged, for the rationality of believing T' or using it in technological applications. One reason why such a 'cobbled up' T' is not regarded as thus confirmed by e despite entailing it is, of course, that exactly the same trick could be pulled starting from a theory S with radically different observational consequences than T: we could just as well 'tack' e onto this different theory S, and the resultant S' would guide applications quite differently than does T'. This is, of course, the venerable 'tacking paradox', so often raised, only to be swept under the carpet.[2]

Scientists and historians tend to ignore the 'tacking paradox' as a philosophers' plaything: no scientist would ever cheat so blatantly as by tacking. There are, however, real historical cases of procedures not so very different from tacking. Take the case of the early nineteenth century revolution in optics. The story usually given is that this revolution occurred because the wave theory of light turned out to make many predictions (for example about diffraction and polarisation effects) which proved correct observationally and which the rival, longer established corpuscular theory either contradicted, or, more often, simply could not match. Such accounts make it easy to see why scientists would rationally prefer the wave theory; unfortunately a little historical research shows them to be factually quite inaccurate. Far from diffraction and interference effects being new phenomena which surprised and defeated the early nineteenth century corpuscularists, the phenomena had been known since at least the mid-17th century and corpuscularists had been in the business of explaining them since Newton. They attributed the effects to close-range forces of 'inflexion' which ordinary, gross matter exercises on the minute corpuscles of light, thus bending them away from their naturally rectilinear paths. Since the 'inflecting' forces are alternately attractive and repulsive at different distances from the body, light and dark *fringes* are produced. As for polarisation effects, again many of these (those connected with the two rays produced by birefringent crystals) had been known since before Newton, who suggested that they could be explained by endowing the light-corpuscles with 'sides' or 'poles'. Newton's suggestion was turned into a highly elaborate theory by the French scientist J. B. Biot. Biot's theory was of polarised light-particles revolving around 'axes of polarisation', axes which themselves could perform complicated movements.[3]

I do not claim that a corpuscular theory was ever actually produced which was fully observationally equivalent to the wave theory as, say, developed by Fresnel in the 1820s. I *do* claim that many of the so-called crucial effects could in fact be explained within the corpuscular approach and that many more could have been brought within the corpuscularists' net by straightforwardly developing the ideas of Newton and Biot. The reason this exercise was not completed was not its impossibility but the fact that, as corpuscularists eventually realised, it was not scientifically worthwhile. Corpuscularists had managed, by various contrivances, to account within their theory for certain effects which are entailed quite naturally by Fresnel's wave theory. Yet neither the scientific community of the time nor subsequent generations of scientists took this achievement of the corpuscularists at all seriously. The great majority of scientists took it that these effects which the corpuscularists had forced into their framework still provided no genuine support for their theories. In other words, the wave theory was regarded as well ahead of its rival in terms of empirical support even before the corpuscularists finally surrendered; the wave theory's superiority was taken as established even by some phenomena, a correct description of which fully-fledged versions of the corpuscular theory entailed.

I labor this point because it does seem to me that we must capture this sort of intuitive judgment above all others if we are to have any hope of explaining scientific change as rational. Clearly if our formal account of empirical support *is* to capture this intuition then it must make derivability of the evidence at most a necessary and *not* a sufficient condition for empirical confirmation.

Someone who realised this rather clearly long ago is William Whewell. Writing about this same optical revolution he said:

> When we look at the history of the emission theory of light, we see exactly what we may consider as the natural course of things in the career of a false theory. Such a theory may, to a certain extent, explain the phenomena which it was contrived to meet; but every new class of facts requires a new supposition – an addition to the machinery; and as observation goes on, these incoherent appendages accumulate, till they overwhelm and upset the original framework. Such was the history of the hypothesis of solid epicycles; such has been the history of the hypothesis of the material emission of light. In its simple form, it explained reflection and refraction; but the colours of thin plates added to it the hypothesis of fits of easy transmission and reflection; the phenomena of diffraction further invested the particles with complex hypothetical laws of attraction and repulsion; polarisation gave them sides; double refraction subjected them to peculiar forces emanating from the axes of crystals; finally, dipolarisation loaded them with the complex and unconnected contrivance of moveable polarisation; and even when all this had been assumed, additional mechanism was still wanting. There

> is here no unexpected success, no happy coincidence, no convergence of principles from remote quarters; the philosopher builds the machine, but its parts do not fit; they hold together only while he presses them; this is not the character of truth. (Whewell 1837, II, 340)

The reason why Biot's theory of fixed and moveable polarisation is not genuinely confirmed by the polarisation effects seems obvious: the theory was actually constructed using those facts, it was rigged precisely to yield the known phenomena, or, in Whewell's phrase, it was 'contrived to meet' the known results. It had to look for real support to facts other than these, but no such support was forthcoming. To take a more recent example, the reason why classical physics as modified by the incorporation of the Lorentz-Fitzgerald contraction hypothesis is not generally taken to be supported by the result of the Michelson-Morley experiment is surely this. It is generally assumed that the amount of contraction postulated by Lorentz and by Fitzgerald was "read off" the null experimental result.[4] In other words, *the experimental outcome was used in the construction of the theory and hence cannot then be used again in its support.*

This seems the obvious lesson to draw from such cases and several philosophers have at first drawn it. Here, for example, is a passage from Moritz Schlick, which is also quoted with approval by Karl Popper in an early work:[5]

> the confirmation of a prediction means nothing else but the corroboration of a formula for those data which were not used in setting up the formula. Whether these data had already been observed or whether they were subsequently ascertained makes no difference at all.

There is, however, no trace of any heuristic element in the formal accounts of confirmation developed by members of Schlick's Vienna Circle and its later followers. And, as for Popper, he and later Popperians quickly switched to the 'historical' or 'temporal' account of confirmation.[6] According to this temporal account, the actual time-order of theory and evidence is a crucial factor in support. If a theory entails some empirical result but the result was already known to hold when the theory was first formulated, then the theory receives much less confirmation than if the empirical effect was found to occur only as a result of the theory's predicting it. (Indeed on Lakatos's extreme version of the 'temporal view' a theory is confirmed *only* by correct predictions and *not at all* by any evidence which it yields but which was already known.) (Lakatos 1978, I, p. 36)

The temptation to embrace this temporal view is strong: it captures some of the intuitive confirmation judgments which I have highlighted and makes them depend on a factor which is reasonably clear-cut and objective. We know that the Michelson-Morley result was obtained in 1887, while the Lorentz-Fitzgerald contraction hypothesis was formulated only in 1892. On the other hand, the existence of the planet Neptune was definitely *pre*dicted by Newtonian theory, the existence of the star-shift by general relativity theory, and so on. The important factor in determining the extent of confirmation is, at any rate usually, unambiguous. But if we have to look at the way that a theory was constructed in order to decide on confirmation, then things seem to become very messy. Wouldn't we need to have Einstein's psyche available for inspection in order to know how he arrived at his theory of relativity? Surely the philosopher should shun such subjective matters. And if we did know how theories were arrived at, mightn't it sometimes be true that two scientists arrived at the same theory in two quite different ways, and wouldn't this mean that a theory as arrived at by scientist *A* might be confirmed by evidence *e* on this heuristic view, while the *selfsame* theory as arrived at by scientist *B* is not confirmed by the self-same evidence? Surely this is a *reductio ad absurdum* if ever there was one?[7]

But the 'new predictions count more' view, whatever its appeal, is wrong. First, it seems altogether mysterious just why the time-order of theory and evidence *should* matter. My intuitions here are all on the side of John Stuart Mill who, in an often quoted passage wrote:[8]

> it seems to be thought that an hypothesis ... is entitled to a more favourable reception, if, besides accounting for all the facts previously known, it has led to the anticipation and prediction of others which experiment afterwards verified ... Such predictions and their fulfillment are, indeed, well calculated to impress the

> ignorant vulgar, whose faith in science rests solely on similar coincidences between its prophecies and what comes to pass. But it is strange that any considerable stress should be laid upon such a coincidence by persons of scientific attainments.

Mill's sentiments were equally famously echoed in this century by John Maynard Keynes, who wrote:

> [the] peculiar value of prediction ... is altogether imaginary ... The question of whether a particular hypothesis happens to be propounded before or after examination of [its experimental consequences] is quite irrelevant. (Keynes 1921, p. 305)

Mill and Keynes were surely right that the time-order suggestion lacks both inherent plausibility and intuitive justification. There is however another strong argument against it: it fails to capture all of scientists' intuitions about particular cases. It is certainly true that *some* results which a theory can account for but which were already known before the theory was developed are not regarded as lending the theory support. But this is far from *always* being the case. The facts about Mercury's perihelion advance, for instance, had long been known and been at the center of astronomical concern before the development of general relativity. Yet that theory is regarded as receiving enormous support from the facts about Mercury's orbit – quite as much support as it receives from the newly predicted facts about star-shift. Similarly, Newton's theory's precise account of the already known facts of the moon's orbit and of the earth's oblateness seem to have been taken as supporting his theory quite as strongly and dramatically as any novel prediction.[9]

It seems that while all novel, correct empirical predictions count significantly in favor of the theory which made them, *not* all empirical results which count significantly in favor of a theory are novel. The time-order of theory and evidence cannot *in itself* be the important factor in distinguishing genuine and spurious confirmations.

Where else might the missing factor be found? Let's next try the mainstream tradition in confirmation theory, namely Bayesianism. One usual Bayesian account is that the important quantity so far as confirmation is concerned is the ratio $p(h/e)/p(h)$: confirmation is the greater the higher is this ratio. One problem is that this account is principally designed for the case of statistical hypotheses and even some of its best friends would admit that it leaves something to be desired in the case of deterministic theories (the only ones I am concerned with in this paper). Nevertheless, let's see what Bayesianism can tell us about this case. Since $p(h/e)=p(e/h).p(h)$, the above ratio can also be expressed as $p(e/h)/p(e)$. I am interested only in the case in which empirical results are entailed by the relevant theories, that is in which $p(e/h)=1$; and so the Bayesian confirmation ratio is always going to be just $1/p(e)$ – one over the "prior probability" of the evidence e. The question, as so often with Bayesianism, is 'prior to what'?

If $p(e)$ is taken as an 'absolute probability', as some measure simply of the logical strength of e then, while such intuitive judgments might be delivered as that several pieces of independent evidence confirm h more strongly than a single piece, there seems no hope of making the sort of discrimination which I have argued is necessary among the confirming effects of 'single', 'atomic' pieces of empirical evidence.

If we interpret $p(e)$ as the probability of e, given everything that was accepted at the time the relevant h was proposed, then the Bayesian approach seems simply to incorporate a formalised version of the temporal account of confirmation – an account which, I have argued, is unacceptable.

The only hope that I see of successfully utilising the Bayesian system is to identify $p(e)$ with the probability of e, relative not to the whole of 'background knowledge' at the time h was proposed, but relative only to that part of 'background knowledge' which was used in the construction of h. We could then say that h is not confirmed by e unless $p(h/e)/p(h)>1$. Assuming that h entails e, this ratio would be greater than one if $p(e)$ is less than one, and would be equal to one only if $p(e)$ equals one; and $p(e)$ would equal one on this construal precisely if e were already entailed by some results used in the construction of h. While this is the right result, it clearly amounts to nothing more than a reformulation of the heuristic view, and giving it this Bayesian formulation has in itself taken us no further forward. We are still under an obligation to explain precisely what it means for empirical results to be used in the construction of a theory and to face certain difficulties with the view. Once these problems have been solved, that is, once the hard work has been done, then we certainly can, if we choose, express the result in Bayesian terms.

The Bayesian might resort to other functions of probabilities to represent confirmation values. The only other factor at his disposal, however, in building these functions seems to be $p(h)$ – the 'prior probability' of the hypothesis. The confirmation function might be so constructed that, if the prior probability of h is sufficiently low, even a large body of positive evidence leaves h without appreciable confirmation. If assignments of prior probabilities to hypotheses are not to be simply reflections of the quirks of individual scientists, then the suggestion amounts, in less formal terms, to the proposal that account should be taken not only of whether a theory entails the right results, but also of whether the theory is sufficiently plausible or simple or unified.

The suggestion that plausibility or simplicity or some kindred notion is the key to this problem is an attractive one, and one which several philosophers, both Bayesian and non-Bayesian, have adopted. The general idea would be that once a theory becomes sufficiently complex, it is not confirmed even by empiricial results which it correctly entails. In the particular case of early 19th Century optics, the suggestion would be that the fully-fledged corpuscular theory of light, complete with alternately attractive and repulsive forces of inflexion and particles with axes of polarisation which perform jerky and irregular movements, is just too implausible and complicated for the fact that it entails correct results about diffraction and polarisation effects to count in its favor.

Biot's theory is undeniably enormously implausible and complex, intuitively speaking. Moreover, there is, as I shall indicate later, a close connection between the heuristic view of confirmation and the intuitive idea of simplicity. Still simplicity is *not* the obvious (and old) solution to the problem that I have been focussing on.

First, it is by no means obvious that we would do ourselves any favors by invoking plausibility or simplicity or the like. Allowing that heuristics play a role does indeed threaten to make confirmation a dangerously unclear and subjectivist notion. But does invoking plausibility or simplicity instead really improve matters? After all, centuries of effort aimed at producing clear and objective notions of plausibility and especially simplicity have failed notoriously to bear fruit.

Secondly, if something like simplicity/complexity *were* the key to this problem, then there would have to be no cases of the following sort. A theory is proposed which is (*i*) intuitively simple, (*ii*) logically entails an accepted evidence statement *e*, and yet (*iii*) is intuitively not supported by *e*. But there are such cases. The following is an artificially simple, yet instructive, abstract example. There are two theories T and T' which appear equally simple; and in the area of the two theories only one relevant piece of evidence is known. This evidence, *e*, is that the two variables x and y take on the values 2 and 10 respectively in a certain circumstance. One of the theories, T say, yields *e* directly and without artifice. The relevant equation yielded by T', however, is $y = ax$, where a is, at this stage, a free parameter. We can now use *e* to fix the value of a at 5 – thus creating a new theory T'' which (in the sense is which I am using the word in this connection) entails *e*. There are now, therefore, two theories, T and T'' each of which entails the known relevant evidence. I claim we should nonetheless say that – provided other things are equal, as we are assuming they are – *e* favors T above T'': T was genuinely tested by *e* (or so we are assuming), but T'' on the contrary was *not* genuinely tested by *e* – T'' *had* to get *e* right because of the way that it was constructed. If the invocation of simplicity did indeed solve this problem by yielding this judgment, it would have to be clear that T'' is more complex than its predecessor T', that is, simply filling in the value of an initially free parameter would have to increase automatically the complexity of a theory. This is surely not right.

I have argued, then, that the attempts to avoid drawing what I claim is the obvious lesson of cases like Biot's theory all fail. This 'obvious' lesson, remember, was that even if a theory entails an accepted empirical result, the theory is not supported by that result, if the result was *used in the construction of* the theory. I claim that we must return to this obvious lesson (all else having failed) and face up squarely to any difficulties it brings in its wake.

Perhaps the most obvious difficulty with this heuristic view of empirical support is its vagueness: what exactly does it mean for an empirical result to be used in the construction of a theory? I shall try to show that this is at any rate not a hopeless problem – by showing that there are some especially clear-cut and clearly describable cases. I shall then outline and try to rebut various more specific criticisms that have been made of this heuristic criterion.

3. SOME CLEAR-CUT CASES OF THE USE OF EMPIRICAL RESULTS IN CONSTRUCTING THEORIES

(a) "*Exception-Incorporation*"

Thomas Young did not produce one version of his celebrated principle of interference, but several quite different versions. The principle began life in 1802 in the quite general form that *any two* near-parallel beams of light which affect the same area will there produce alternating light and dark interference fringes.[10] Several critics, notably Young's arch enemy Henry Brougham, pointed out that this general proposition is quite readily refuted by a host of everyday facts – for example, if the proposition were correct, two closely contiguous candles casting their light on some nearby wall ought to produce fringes. Young had accounted for external straight-edge diffraction fringes by assuming that they are caused by the interference of two 'portions' of light – a 'portion' of direct light and another 'portion' 'inflected' by the diffracting object.[11] The obvious difference between this case and the two candles case is that the two candles are two *separate* sources of light, whereas the two 'portions' in the diffraction case originate in the same source. Young therefore responded to Brougham's objection by switching to the claim that interference occurs when, *and only when*, the two interfering 'portions' of light originate in the same source. (It is important to remark here that Young had no clear ideas about what came to be called the coherence of different light beams.) The modified principle entails, of course, the experimental result that the two contiguous candles will *not* produce interference fringes. While this is precisely what is observed, no one (least of all Young himself) would claim that it constitutes empirical support for his modified theory.

The logical structure of this episode is clear. It involves a rather stronger version of what Lakatos in his studies of mathematical discovery called 'exception barring'

(Lakatos 1976), – an appropriate name might be 'exception-incorporation'. An exception having been found to some general principle, some characteristic is sought which the exception does not possess, but which known positive instances do, and the original principle is then claimed to hold *when and only when* this characteristic is satisfied.

To avoid misunderstanding, let me say immediately that there is of course no question but that Young was right to take the empirical facts into account. His modified principle is consistent with experiment and hence is clearly preferrable to the original principle which is inconsistent with experiment. I claim only, first, that this is a clearly articulable way in which an experimental result may be used in the construction of a theory, and, second, that the intuition is clear that, although the modified principle entails the candles result which refuted its predecessor, the modified principle is not supported by the result. Moreover, should a different theory be produced which, directly and without artifice, entails the same results, then this new theory should, I claim, be judged superior to Young's modified theory *on this very ground* – that is, without necessarily requiring that the new theory be confirmed independently of these results.[12]

(b) *Parameter-adjustment*

A second type of case has already been touched upon and is, perhaps, clearest of all. In cases of parameter-adjustment, a theory is first proposed which has some free parameters – these may be parameters in the usual specific sense (constants in some mathematical equation) or in a more general vaguer sense. For example, the basic idea behind the mainstream corpuscular-theoretic account of diffraction – that the phenomenon was due to alternately attractive and repulsive forces of 'inflexion' – gave corpuscularists not a few degrees of freedom: the distances at which the force turned from attractive to repulsive and back again, the intensity of the force at given distances, the mass of the corpuscles themselves could all be juggled with in an attempt to accommodate the known experimental results. It is clear that there will be whole sets of results which the theory can be made to yield by judiciously reading off suitable values of the parameters from those very results.

It is easy to see the methodological defects of the original corpuscular theory with all its free parameters. It has a low, perhaps zero, degree of testability. But what of the theory with all the free parameters filled in? It certainly is testable in the logical sense – it has consequences which are directly comparable with, and indeed which compare favorably with, experiment. But this is no wonder since these consequences were all 'written into' the theory *via* the parameter-fixing. Many of Popper's most perspicacious remarks are, I believe, based on an *intuitive* notion of testability which readily applies here. The theory with the fixed parameters is certainly not intuitively testable against the observations used to fix the parameters; since, *because of the way it was constructed*, it was never at risk of refutation from these observations. *This fact cannot be discovered by simply inspecting the logical form of the theory*. Hence Popper's formal account of testability, which is in logical terms (set of 'potential falsifiers'), fails to do full justice to the notion; and Popper has never, I think, fully and clearly realised that his intuitive remarks about testability rest on heuristic considerations. It is easy to make a theory testable in the logical sense of having it entail empirical results which are already known. It will then, of course, be true that *were* these results not correct, the theory would be refuted. But this is no genuine *test*, since the results were already known to hold before being incorporated into the theory. Perhaps the clearest way to state my thesis is, then, this: Popper is right that a theory is genuinely supported only by passing real tests, but in order to decide whether some empirical result constituted a real test of some theory, we have to look at how the theory was constructed.[13]

Once again I should emphasize that I am *not* claiming that there is anything unscientific about parameter-adjustment – many of the best scientific theories contain parameters whose values had to be 'read off' the facts. (An example is the classical wave theory of light and the precise wavelengths of light of different colors.) These theories then went on to make *further* empirical predictions which did indeed lend them support. I *do* claim however that the results which were used to fix the parameter values provide no such support; and, especially, that when one theory has accounted for a set of facts by parameter-adjustment, while a rival accounts for the same facts directly and without contrivance, then the rival does, but the first does not, derive support from those facts.[14]

Many cases in which an empirical result was used in constructing a theory are cases of parameter-adjustment in either the more specific or broader sense. The reason why the Lorentz-Fitzgerald hypothesis, for example, is not regarded as supported by the Michelson-Morley result is that Lorentz and Fitzgerald seemed to provide themselves with a length contraction parameter, only to adjust this parameter nicely to yield the known null result. (But see footnote 4 above.)

(c) *The 'Correspondence Principle' as a Heuristic Device*

Perhaps the most important way in which empirical results may be used in the construction of a theory is rather more indirect than the two considered so far. Some theories in science have been developed *via* the use of the 'correspondence principle' as a heuristic device. It has proved difficult to give a precise general formulation of this principle. But it does seem clear that the idea that a new theory should, in some way or other, explain the *empirical* success of its predecessor has operated in the history of science – and it has operated, not just as an *adequacy requirement*, but sometimes as a heuristic device in the actual development of that new theory. Here empirical results play an indirect role – the correspondence principle guarantees that the new theory will 'go over' to the older one in those domains where empirical results have shown that the older theory is correct. I shall try to make this rather vague talk more precise by considering a famous historical example.

No logical claim in the history of physics has excited more comment than Newton's claim to have deduced his theory of gravitation 'from the phenomena' supplied by Kepler. It was for a long time widely believed that Newton had done precisely what he claimed to have done and that his deductive method should serve as a paradigm for other scientific innovations. On the other hand, it is clear that Newton's theory and Kepler's laws are logically inconsistent with one another. Although Newton was clearly aware of this plain fact, it was only pointed out plainly by Duhem in 1904 and was subsequently re-emphasized by Popper and others. (Duhem 1904 and Popper 1973) It is, of course, impossible for a valid deduction to lead from consistent premises to conclusions which contradict those premises.

Scientists, though, have either ignored Duhem's point or regarded it as an over-fastidious piece of logic-chopping. Their chief spokesman was Max Born who insisted that, whatever Duhem and formal logic might say, Newton did indeed infer his theory from Kepler's laws.

This confused situation can be clarified in a way which should satisfy both logician and scientist and which, at the same time, throws a great deal of light on the process of scientific discovery.

First, there is no doubt, of course, that Duhem was correct that Newton's theory logically contradicts Kepler's laws – the path which a planet ought to follow according to Newton's theory (and given uncontested initial conditions) is *not* an ellipse, for example. Newton could not have validly derived his theory from Kepler's laws. However, the deduction of various results from Kepler's laws did indeed play an important role in Newton's discovery of his theory of universal gravitation – as Newton himself explained clearly in the *Principia*. He for instance proved that if the sun is considered either at rest or in uniform rectilinear motion, and if Kepler's laws are strictly true of a planet orbiting the sun, then the planet is subject to a net force directed towards the sun, a force which is inversely proportional to the square of the distance between the planet and the sun, and directly proportional to the mass of the planet. The deduction involves Newton's own second law of motion but no other synthetic assumption. This result does not of course imply the principle of universal gravitation – indeed the assumption on which the result is based, that the sun is at rest or in uniform rectilinear motion, is logically inconsistent with the gravitational principle. Nonetheless, instead of some enormous and inexplicable leap being needed for Newton to arrive at his theory, the gap between this result (arrived at deductively remember) and the gravitational theory is intuitively quite small.[15] The gap is bridged by the third law of motion. Again Newton did not simply assert this law, but instead argued for it. He had in fact two main arguments: one largely observational and the other theoretical. The latter is especially simple. It starts from the Cartesian principle of the conservation of total momentum, applied to a two body system. Newton gave this principle the vectorial form:

$$m_1 \vec{v_1} + m_2 \vec{v_2} \quad \text{constant.}$$

Differentiating with respect to time gives

$$m_1 d\vec{v_1} /dt = -m_2 d\vec{v_2} /dt;$$

that is, given the second law of motion,

$$f_{12} = -f_{21},$$

which is the third law for the two body case.

It follows, of course, that if a planet is acted on by a force directed towards the sun, the sun must be acted on by a force directed towards the planet and hence will not be unaccelerated. Kepler's laws were predicated on the assumption that the sun is motionless. Hence again we have inconsistency between Newton and Kepler. This inconsistency is, however, surely unimportant when compared to the following facts. First, Newton knew that since planetary masses are very small compared to that of the sun, the latter will be *almost* unaccelerated according to his theory. Second, Newton showed that it follows from the assumption that a planet describes a strict ellipse relative to the sun and with the sun at one focus that (*a*) the center of gravity *G* of the system consisting of the sun and planet is either fixed or moves uniformly, (*b*) the planet and sun each describe ellipses relatively to the other and these ellipses have a common focus at *G*, (*c*) the force acting on the planet is constantly directed towards *G* and therefore towards the sun, and (*d*) this force is inversely proportional to the square of the distance between the planet and *G* and therefore to the square of the distance between planet and sun.

Newton still needed to assume that this inverse square force operates between *any two* bodies; and this assumption leads to a contradiction with the assumption of perfectly elliptical planetary orbits. Nonetheless, this extension from the two body case is the obvious one to make: Newton's achievement does not of course lie in his having made this step, but instead in the earlier and purely mathematical, deductive steps to the single planet results.[16]

This example points to many methodological lessons, not all of which have yet been drawn clearly – despite intensive study of the case. Some lessons are, however, clear. The case shows for example that the requirement that a new theory match its predecessor's success may *automatically* be satisfied in certain cases ('by construction'). It hence explains what Lakatos may have had in mind – if only in a vague and confused way – in making the seemingly extreme claim that in science the *only* empirical consequences of a new theory which count are those which go beyond its predecessor. (Lakatos 1978, I, 36) *Of course* it matters that the new theory get right what its predecessor already did, but this may sometimes be *guaranteed* by the method of construction of the new theory. In that case the sole question of interest may well indeed be whether the new theory is also empirically correct where it transcends the old.

The central points which this example illustrates are, however, these. First that it is indeed a fact, *and a plain one*, that at any rate weak versions of Kepler's laws were used by Newton in the development of his theory. Second that this fact needs to be taken into account in assessing correctly that theory's empirical support. The first of these facts underscores the point that the heuristic path leading to a theory may well be open to public inspection and not be hidden in the recesses of some scientist's psyche.[17] Other equally clear examples can be cited – for example, it is a matter of public record that Planck used the correspondence principle in his development of the relativistic version of the second law of motion. The requirement that the new theory go over to the classical law as v/c tends to zero figured as an explicit premise in Planck's development of the new theory (Zahar 1973).

The second point concerning empirical support is not quite as clear in this case as it is in some others. First of all, there is a measure of *overdetermination* here – Newton's theory can be derived without making use of the full power of all three of Kepler's laws. So the fact that Newton's theory yields all three laws as approximately correct should clearly be regarded as lending *some* support to the theory. Moreover, at least after the demise of Cartesian vortex theory, Newton had no real competitors and so not much hangs in this case on whether or not we judge those parts of Kepler's laws which Newton did use as supportive.[18] Given that Newton's theory made extra predictions which proved correct – notably about the corrections of Kepler's laws, so-called perturbations, but also about the oblateness of the earth, movements of the moon's axis and so on – it won't matter much if we say that this independent support licenses counting even those facts used in the construction of the theory as supportive. The principles of support-accounting advocated by the heuristic view really come into their own only when two rivals are vying – one really predicting results, the other accommodating itself to those results. Nonetheless, it is surely better even in this case to make a distinction between the results *used* in the construction of the theory and those predicted by the theory once constructed: the extra predictions (perturbations, earth's oblateness, *etc.*) surely count as stronger support for Newton's theory than do the results already taken to support Kepler's laws. I think this reflects the attitude of most scientists. Here, for example, is a very revealing passage from a recent textbook on Newtonian Mechanics:

> If universal gravitation had done no more than to relate planetary periods and distances, it would still have been a splendid theory. But like every other good theory in physics, it had predictive value; that is, it could be applied to situations besides the ones from which it was deduced. Investigating the predictions of a theory may involve looking for hitherto unsuspected phenomena, or it may involve recognizing that an already existing phenomenon must fit into the new framework. In either case the theory is subjected to searching tests, by which it must stand or fall. With Newton's theory of gravitation, the initial tests resided almost entirely in the analysis of known effects – but what a list! (French 1971, pp. 5-6)

And the author goes on to give the list which includes the bulging of the earth, the variation of the gravitational acceleration with latitude and the change in direction of the earth's axis of rotation. The passage conforms precisely to the heuristic account: the theory 'stands or falls' by the results of 'searching tests' on it; a test can concern *either* 'hitherto unsuspected phenomena' *or* 'already familiar phenomena' (both stand entirely on a par), the only empirical consequences which do *not* supply tests are those from which the theory was 'deduced'.

3. SOME DIFFICULTIES FACING THE HEURISTIC ACCOUNT OF EMPIRICAL SUPPORT

(a) *Is the Account Subjectivist and/or Relativist?*

The heuristic account of confirmation is far from new. Hints of it, usually mixed in with elements of other views, are to be found in several earlier philosophers. The account's present lease of life perhaps began in 1973 – fathered by some remarks in a paper by Elie Zahar (1973). The account has already met a fair amount of criticism (Musgrave 1974 and 1978, Gardner 1981).

Underlying several of these criticisms is the fear that the heuristic path which led to a theory may not be fully articulable – who knows what really went on in Einstein's or Newton's or Fresnel's minds before they produced their great theories? I hope that what I have said already about the clear-cut cases will have allayed this fear and indicated that we don't in fact need as much information about scientists' psyches as might appear. Scientists usually *argue* for their theories, trying to show that they follow 'naturally' from a combination of certain general metaphysical principles and certain empirical results. They *design* their theories using certain materials and we can always ask whether or not certain particular experimental results were among those materials. Moreover, this question is not usually so difficult to answer. It is clear that versions of Kepler's laws played a role in the genesis of Newton's theory, while the earth's equatorial bulge, say, did not. It is clear that general relativity theory arose from Einstein's extension of his relativity programme to the case of accelerated frames and that Einstein needed no results about the orbit of Mercury, for example, to fix any aspect of his theory. Hence the fact that general relativity yields the (more or less) correct result about Mercury's orbit is surely a great feather in its cap – independently of the fact that the facts about Mercury's orbit were already well-known *and* independently of the historical fact that Einstein was hoping that his theory would indeed deal with the Mercury anomaly. (For it is no part of the heuristic view that it should matter what Einstein was worrying about at the time he produced his theory, what matters is only whether he needed to use some result about Mercury in order to tie down some part of his theory.)

Even when we have no account from the author himself of how the theory was constructed it may not be so difficult to see whether or not some particular empirical result was involved in the process. The case of Thomas Young and his modified principle of interference is again instructive. The modified principle says, remember, that interference occurs beween two 'portions' of light *when, and only when,* certain special conditions apply – in particular the condition that the two 'portions' of light were derived from the same source. Young was not saying that interference fringes are *observable* only under these conditions, but rather that interference *occurs* only under these conditions. Why are these restrictions necessary? This seems a 'natural' question to ask in view of the fact that the basic idea behind the wave theory would seem to make any such restriction superfluous. If light consists of waves in the ether then *any* two light beams affecting the same area ought to interfere in the sense that the resultant disturbance should be the vector sum of the disturbances which would be produced there by either beam separately. Young gives no answer to this 'natural' question and the only conclusion that can be drawn is that the restriction was 'read off' the known refutations of his original, general interference principle.

Of course, an answer to the question was eventually given within the framework of the wave theory. Fresnel argued that the disturbance produced at any point of the ether would indeed be the vector sum of the disturbances from all the sources producing light. If, in particular, these disturbances happened to interfere destructively at that point then darkness would be produced there. We would expect this darkness to be fleeting, however, because the movements of matter creating the light seem likely to be subject to unpredictable fluctuations. This means that there would be no correlation between the fluctuations in different light sources, and so, in this case, the places at which destructive interference occurs would be constantly changing. Hence, since our visual apparatus is not fine-tuned enough to register these quick changes, the illusion of constant illumination is produced. It will only be when a change in the oscillations making up one 'portion' of light is matched automatically by changes in the other 'portions' that the interference pattern produced will be constant and so observable. And matching changes in oscillations will occur only when the various 'portions' left the same source at more or less the same time – matching changes will certainly *not* occur in the light from two different sources, like the two candles. Given Fresnel's theory, then, we would expect *in advance* observable interference fringes in the case of light from the same source and no observable fringes in the case of light from two different sources. It is surely reasonable, then, to regard Fresnel's theory as confirmed by the fact that both these expectations prove correct empirically – quite independently of the historical accident that both results were known before Fresnel formulated his theory. But in Young's case, and independently of any facts about his psyche, he failed to give the scientific community of his time any reason for the restriction of his principle to the case of 'portions' of light originating from the same source, beyond the fact that this was the way things had already been observed to be. So Young's principle is less well supported by these facts than Fresnel's even though both entail the facts. Most importantly, the factors necessary to decide whether or not there is support in the heuristic sense can be discovered in Young and Fresnel's works – no doubt a certain amount of interpretation is necessary to discern them, but no more, I would claim, than is required in discerning precisely what theory is being propounded by a scientist.

If the heuristic path which led to a theory can indeed be specified objectively, then the alleged subjectivism or relativism of my account of support disappears. As for the purported *reductio* (that the same theory may be arrived at in two different ways and hence receive different degrees of support on this account), this can be seen to be in fact a straightforward consequence of the three-term nature of the proposed support relation. If support is a relation not just between an empirical result and a theory, but between an empirical result, a theory, and the way in which the theory was arrived at, then any apparent contradition here disappears: e may support T relative to one heuristic path, while the same e does not support the same T relative to a different heuristic path. Indeed, far from this being a *reductio* of the approach, it is precisely what is needed to capture our intuitions: while we may not have historical cases of exactly the same theory being arrived at by different routes, we certainly do have such cases of theories with exactly the same set of observational consequences and yet which intuitively are differently supported by the facts.[19]

One final internal objection to the heuristic account of empirical support runs as follows. 'Admittedly in the case of Thomas Young, say, the historical order seems clear: first he produced a very general principle, then noticed that it was experimentally refuted, and then produced a modified principle by "incorporating the exceptions". But what if Young had not published until this final stage? Worse still, what if Young had simply *dreamed-up* the principle in its already modified form without having to take any notice of empirical results?' It would seem that my account penalizes a careful scientific thinker like the real-life Young who first formulates a general theory and then takes observed exceptions into consideration, and that it rewards a lucky imaginary Young who just happened to dream up the modified, and empirically correct theory.[20]

My answer to this powerful-sounding objection can, I think, only be this: science just isn't like that. Scientific theorizing is *never* a question of throwing out conjectures, more or less at random, most of which are then refuted. On the contrary, as I have already argued, scientific theories are *designed* to meet certain requirements – one of the foremost of which is that they capture in some way or another their predecessors' empirical success. It is massively improbable that an empirically successful theory from one of the more developed sciences will simply have been dreamed-up fully fledged; and scientists in fact seem to take this improbability as an impossibility.

Consider, for example, some successful theory containing parameters fixed at certain values – say the wave theory of light with its precise values for the wavelengths of monochromatic light from particular parts of the solar spectrum. It is, of course, logically possible that some early wave-theorist one night in his sleep received a message from beyond that the wavelength of some monochromatic light from the extreme violet end of the spectrum is some precise number (around 4×10^{5} cm) – this is possible but massively improbable, and apparently never condoned as a practical possibility in science. The scientist will want a *reason* for this precise value. Sometimes the answer will be that the value was dictated by basic and general theoretical considerations; more often, as in this particular case, the answer will be that it could be 'read off' one of several possible experimental results. If so, then this needs to be taken into account in determining empirical support.

(b) *The Question of Justification. Why Should it Matter How a Theory was Arrived At?*

Finally, I want to see what can be done to meet a different type of critic than the ones considered so far. This new critic is ready to allow that the ideas behind the heuristic criterion can be made sufficiently precise, and even perhaps might allow that it is this criterion which scientists intuitively apply, but he questions the *justification* for the criterion. Perhaps scientists do, as a matter of fact, act as if it mattered how a theory was constructed, but why on earth *should* they act in this way?

John Stuart Mill raised something like this point against William Whewell. (Mill 1843, p. 328) As I already suggested, the basic idea behind the heuristic criterion of confirmation is undoubtedly to be found in Whewel's writings. Unfortunately in presenting his ideas on what he called the 'consilience of inductions',[21] Whewell sometimes conflated the heuristic view with the 'purely temporal' account. This gives some plausibility to Mill's complaint that, although successful predictions are well calculated to impress the 'ignorant vulgar' and while they might even, *as a matter of fact*, be given special weight by those of 'scientific attainments', this fact is nonetheless 'strange' – that is, unjustified. As I already said, I agree with Mill that there is no reason on earth why the time-order of theory and evidence should in itself have any relevance for the question of how well a theory objectively stands up to the evidence.

However, once the heuristic account is clearly differentiated from the temporal, and once it is made clear that it is the heuristic account only which is being defended, then the situation surely changes. The justification for the *negative* aspect of the heuristic account seems almost self-evident. If the theory was adjusted so as to yield a certain result, then its yielding that result tells us something only about the ingenuity of man; it tells us nothing about the likelihood that the theory reflects some part of the blueprint of the universe, or even about its 'rational acceptability'. This is especially clear in the most blatant form of adjustment to known results: namely, 'tacking'.

Take some theory T (perhaps a wild one with all sorts of implausible, but untested consequences – though the point is independent of this supposition). Now 'tack' onto T any independent empirical result e which is already known to hold. If someone were to point to the fact that the resulting theory entails *e* as a reason for 'accepting' the overall theory and residing some degree of 'rational belief' in the theory's other consequences, then he would rightly be laughed out of court.

The real (and of course ancient) problem is to develop some contrast with the negative case by giving some reason why we *should* regard correct, genuine predictions as lending positive support to the theory which makes them. (Again assuming that such a judgment of positive support has some consequences – no matter how heavily hedged – for the rationality of working on, or applying, the theory concerned.)

In a recent paper Laudan (1981) accuses Whewell of doing nothing more than 'handwaving' in response to this problem. But Whewell *did* attempt an argument. He claimed that, underlying the attribution of special weight to predictions, whether temporally novel or not, is the assumption that the theory cannot have made the correct prediction *by chance*. He wrote:

> Men cannot help believing that the laws laid down by discoverers must be in a great measure identical with the real laws of nature, when the discoverers thus determine effects beforehand in the same manner in which nature herself determines them when the occasion occurs. Those who can do this must, to a considerable extent, have detected nature's secret ... Such a coincidence of untried facts with speculative assertions cannot be the work of chance, but implies some large portion of truth in the principles on which the reasoning is founded. (Whewell 1837, II, p. 64)

And again:

> No accident could have given rise to such an extraordinary coincidence. No false supposition could, after being adjusted to one class of phenomena, exactly represent a different class, where the agreement was unforeseen and uncontemplated. (Whewell 1837, II, p. 68)

What are we to make of this argument? Does it indeed amount to anything more than 'handwaving'?

I am sure that Whewell is right that when confronted with a successful prediction (in the more general sense which includes predictions of already known results) we intuitively dismiss the possibility that this success is due to chance. The problem – as Mill already essentially pointed out – is that the dismissal of this possibility is certainly not based on pure logic: all false theories have true consequences, a theory which has been strongly supported on the heuristic criterion *may* nonetheless be radically false and we just happened to have tested 'unrepresentative' consequences. As Mill pointed out, the fact that the wave theory of light, for example, has made successful empirical predictions does not prove it to be correct. Indeed:

> Though twenty such coincidences should occur, they would not prove the reality of the undulatory ether; it would not follow that the phenomena of light were the results of the laws of elastic fluids, but at most that they are governed by laws partially identical with these... (Mill 1843, Ch. XIV, pp. 326-331)

And Mill's point has, of course, been driven home by the subsequent development of science. Science now sees the classical wave theory of light as, strictly speaking, false. It follows that Whewell's claim that predictive success was never enjoyed by a theory which subsequently proved to be false is itself false.

In defense of Whewell, it should be remarked that a rather liberal notion of truth is implicit in his view. He continues to speak, for example, of Kepler's laws being true while explicitly acknowledging that Newton's theory "corrected" them; and in the first passage quoted (p. 28) he speaks of successful predictions *not* as showing that the theory is true, but rather that it 'must be *in a great measure* identical with the real laws of nature', or that it contains 'some large portion of truth'. So his claim that successful predictive evidence has 'never yet been produced in favour of a falsehood' is not as easily dismissed as might at first appear. Whewell's claim should, perhaps, be interpreted as asserting that no theory was ever positively supported on the heuristic criterion but subsequently found to be "totally false" or found *not* to contain "important truth".

This saves Whewell's claim from falsity only by rendering it vague. While it is clear what it means for a theory to be true or false in the strict sense, a great deal of effort (especially of late) has so far failed to provide any sort of precise acceptable sense for the notions of 'approximate truth' or 'non-radical falsehood'. Indeed I remain sceptical that any precise notion of 'approximate truth' will ever be produced which will allow that once accepted but now superseded theories may be 'approximately true'. There is undoubtedly an important element of continuity in science *at the empirical level*, but this continuity does *not* extend to the theoretical level (at any rate to the level of what is sometimes called 'interpreted theory'). The ether, and therefore waves in it, are totally rejected by modern science, and so it is hard to see how the wave theory of light is, on present standards, anything other than outright false. Of course, many of its empirical consequences are correct – so we can certainly continue to say that light behaves in certain respects *like* a wave in a medium. But latest theory tells us that light *is not* such a wave. Unless we espouse some positivistic reduction of theories, we are, I think, stuck with sharp discontinuities at the fully-fledged theoretical level in science.[22]

The question of whether Whewell's claim can be modified so as to rationalize the weight given to successful prediction in science is one that I can only raise here and must leave open. It involves difficult problems about induction and about scientific realism. But, while leaving the question of its justification open, I still claim that the heuristic criterion is the one which is, as a matter of fact, applied intuitively by scientists. Whether or not convincing general arguments can be produced for the practice, we do, both in science and in everyday life, intuitively dismiss the chance explanation of predictive success.

The London School of Economics

NOTES

*At various points in this paper (and especially in section 2(c)) I am indebted to Elie Zahar. I completed the paper during a most enjoyable and stimulating stay as visiting fellow at the Center for Philosophy of Science, University of Pittsburgh. I should like to thank Larry Laudan, Nicholas Rescher and especially Adolf Grunbaum for making that stay possible. I received interesting comments on an earlier version from my fellow visiting fellows, especially Ron Giere and Ron Laymon.

[1] I should say immediately that this thesis is far from new. As will be seen, it is to be found, though in a rather impure form, in Whewell's writings. It was revived by Elie Zahar (1973) and I tried to develop Zahar's arguments and respond to some early criticisms (Worrall 1978).

[2] One recent philosopher who has insisted that the tacking paradox must be given due attention is Clark Glymour (1980). While very much agreeing with Glymour about the importance of the problems he raises, I believe he exaggerates the novelty (and adequacy) of the solutions he advocates (Worrall 1982).

[3] For more details of the corpuscular optics program see (Worrall 1976).

[4] The *real* situation was rather different – see especially (Zahar 1973).

[5] Popper quotes this passage from Schlick in his (1979) - written in the 1930s but only recently published. I was directed to the passage by Elie Zahar.

[6] See especially (Musgrave 1974).

[7] This objection was raised by Alan Musgrave (1974).

[8] J. S. Mill (1843), p. 328. In later editions 'ignorant vulgar' was replaced by some more tepid phrase.

[9] See, for example, French (1971, pp. 5-6).

[10] The restriction to near-parallel beams was based on Young's theory that the optical disturbance is *longitudinal*. For details see (Worrall 1976).

[11] This was Fresnel's first theory too – it is of course incorrect. It was Fresnel and Fresnel alone who eventually discovered the "true" wave-theoretic account of diffraction in which the diffracting object plays no role except that of absorbing the light which falls on it.

[12] Fresnel's theory of interference was indeed such a theory – although it was also independently supported.

[13] I learned to phrase the thesis in this way during discussions with Ron Giere. Giere also holds a version of the heuristic view, but one which differs from mine in some important respects. See especially (Giere 1983).

[14] I have, throughout this paper, simplified the presentation by assuming that the important question is always whether or not a given, *single* result was used in the construction of a theory. But, quite often, the situation is that there is a whole class of results, any of several subsets of which could tie down the free parameters, hence producing a theory which predicts the other results in the class. To take a simple example, say that *T* implies the equation $y = ax + b$ and that the known relevant empirical results are *e*: $y = 5$ when $x = 1$, *e*': $y = 8$ when $x = 2$, and *e*": $y = 11$ when $x = 3$. Any two of the three results could be used to fix parameters and the resulting relation $y = 3x + 2$ used to *predict* the remaining result. Or, to take a real example, the wave theory of light makes several predictions of fringe-spacings as one-to-one functions of the wavelengths of the light involved. Any of the corresponding experiments can be used to fix the wavelength and the results of the rest then become genuine, explicit predictions of the theory.

Generally, therefore, the question is not does *e* support *T*? But rather how much support does (*e* , *e* , ...,e) lend to *T*? My claim is then that the size of the smallest subset of this evidence set needed to tie down initially free parameters in *T* must be taken into account in answering this question. For clearly if *T* entails every element of (*e* , *e* ,......,*e*) without having used any of them, while *T*' needed *n*-1 of these results to fix parameters and then predicts the *n*th, then *T* is considerably better supported by (*e* , *e* ,....,*e*) than is *T*'. (For detailed consideration of an example see (Elie Zahar 1978).)

[15] None of this is said in an attempt to belittle Newton's achievement – indeed in my opinion it enhances it. It is meant only to show that scientific creativity is rather different from artistic, and much closer to *mathematical*, creativity. The framework within which creativity is possible is much more tightly constrained than might be imagined – see (Duhem 1904).

[16] The above account of Newton's inference to universal gravitation is of course oversimplified – for a fuller, but

still somewhat simplified account see (Zahar 1983). I am heavily indebted to Zahar's account. For a better worked-out treatment of a rather simpler case of "inference" to a general theory, see (Worrall 1983).

[17] Indeed the way that a scientist really arrived at the theory and the way the theory was constructed out of available materials may be quite different. For more details see (Worrall 1983).

[18] This is shorthand for 'whether or not we regard the empirical results which supported Kepler's laws as in turn supporting Newton's theory'.

[19] One example of such a pair of theories is the Poincare-Lorentz modified classical physics and the special theory of relativity.

[20] Peter Urbach raised this objection in response to an earlier version of my account.

[21] Whewell's ideas on 'consilience' are more complicated than I suggest here – but underlying them is the thesis that there is a great difference between genuine predictions and results which have been 'written into' a theory.

[22] For a more detailed defense of the position outlined in this paragraph see Worrall (1982a).

REFERENCES

Duhem, P. (1904) *The Aim and Structure of Physical Theory*.

French, A. (1971) *Newtonian Mechanics*. Cambridge: MIT Press.

Gardner, M. (1982) "Predicting Novel Facts", *British Journal for the Philosophy of science, 33*.

Giere, R. (1983) "Testing Theoretical Hypotheses", *Testing Scientific Theories*. ed. J. Earman. Minnesota Studies in the Philosophy Science, vol. X. Minneapolis: University of Minnesota Press.

Glymour, C. (1980) *Theory and Evidence*. *Princeton*: *Princeton* University Press.

Keynes, J. M. (1921) *A Treatise on Probability*. Cambridge: Cambridge University Press.

Lakatos, I. (1976) *Proofs and Refutations*. Cambridge: Cambridge University Press.

__________ (1978) "Falsification and the Methodology of Scientific Research Programmes", *The Methodology of Scientific Research Programmes: Philosophical Papers. Volume I*. Cambridge: Cambridge University Press.

Laudan, L. (1980) "Why was the Logic of Discovery Abandoned", *Scientific Discovery, Logic and Rationality*. ed. T. Nickles. Dordrecht: D. Reidel.

__________ (1981) "A Study of Some Philosophical Controversies about Ether", *Conceptions of Ether*. eds. Cantor and Hodge. Cambridge: Cambridge University Press.

Mill, J. (1843) *A System of Logic* 1st ed.

Musgrave, A. (1974) "Logical versus Historical Theories of Confirmation", *British Journal for the Philosophy of Science, 25*.

__________ (1978) "Evidential Support, Falsification, Heuristics and Anarchism", *Progress and Rationality in Science*. eds. Radnitzky and Andersson. Dordrecht: D. Reidel.

Popper, K. (1979) *Die beiden Grundproblem der Erkenntnistheorie. Mohr-Siebeck*

__________ (1973) "The Aim of Science", *Objective Knowledge*. Oxford: Oxford University Press.

Whewell, W. (1837) *History of the Inductive Sciences*. 3 vols., 1st ed.

__________ (1858) *Philosophy of Discovery*.

Worrall, J. (1976) "Thomas Young and the 'Refutation' of Newtionian Optics", *Method and Appraisal in the Physical Sciences*. ed. Howson. Cambridge: Cambridge University Press.

__________ (1978) "The Ways in Which the Methodology of Scientific Research Programmes Improves on Popper's Methodology", *Progress and Rationality in Science*. eds. Radnitzky and Anderson. Dordrecht: D. Reidel.

__________ (1982) "Broken Bootstraps", *Erkenntnis, 18*.

__________ (1982a) "Scientific Realism and Scientific Change", *Philosophical Quarterly*, 32.

__________ (1983) "Hypotheses and Mr. Newton", *Boston Studies*. Forthcoming.

Zahar, E. (1973) "Why did Einstein's Programme Supersede Lorentz's?", *British Journal for the Philosophy of Science, 24*.

__________ (1978) "'Crucial' Experiments: A Case Study", *Progress and Rationality in Science*. eds. Radnitzky and Anderson. Dordrecht: D. Reidel.

__________ (1983) "Logic of Discovery or Psychology of Invention", *British Journal for the Philosophy of Science,* forthcoming.