JOHN WORRALL

THEORY-CONFIRMATION AND HISTORY

1. MUSGRAVE ON 'LOGICAL' VERSUS 'HISTORICAL' ACCOUNTS OF CONFIRMATION

There are very many topics in philosophy of science on which Alan Musgrave and I see eye to eye. So it has not been easy to do the decent Popperian thing and pick a (friendly) fight with him. However, thinking again about his influential (1974) paper on theory-confirmation ('Logical versus Historical Theories of Confirmation') solved my problem. Despite having some of its heart in some of the right places, both the argument of that paper and the position it ends up endorsing are, I believe, importantly off-beam. In this paper I shall explain why and clarify what I think is the correct account of the issue that he addressed. I shall finally take the opportunity to contrast my views on confirmation with those of Deborah Mayo (see in particular her 1996); Mayo was herself indebted to Alan Musgrave's paper and has developed her own influential account of the issues it raises. Although Alan's paper was published in 1974, the problem it faces has not been given a satisfactory resolution—at least not one that has met widespread acceptance. It remains very much a live issue within current philosophy of science.[1]

Musgrave begins his paper with a sharp formulation of the prediction versus accommodation issue: is there some epistemic premium on *predictive* success? That is, does a theory obtain, *ceteris paribus*, more confirmation from a piece of evidence that it correctly predicts than it does from an 'otherwise equivalent' piece of known evidence that it correctly entails?

He takes it that a 'purely logical' account of confirmation must answer this question negatively. Any such account sees confirmation as entirely based on the logical relationships between the theory, T, and the piece of evidence, e, at issue; and hence must entail, whatever the details of the logical relationships it highlights, that the question of whether or not e was already known to hold, or was already in 'background knowledge' however construed, when T was proposed is entirely

[1] For example, Musgrave's views are one of the starting points for the very recent paper on prediction and accommodation by Chris Hitchcock and Elliott Sober (2004).

C. Cheyne & J. Worrall (eds.), Rationality and Reality: Conversations with Alan Musgrave, 31–61.

irrelevant to confirmation. All logical accounts have their difficulties —in particular, in Musgrave's view, they supply no satisfactory answer to the 'paradox of confirmation.'[2]

An historical (or more accurately—as he allows—a 'logico-historical') account, on the other hand, sees confirmation as a relationship, not just between T and e, but also a third variable: 'background knowledge', b. All variants of the historical view entail that T fails to be confirmed by any e that is in b, even if T (of course, in conjunction with appropriate initial conditions and auxiliaries) entails e. All variants of this account do indeed have an historical element on Musgrave's view: the answer to the question 'does e confirm T?' may very well be different in two different historical epochs, because these will be characterised by different states of background knowledge.

But exactly which evidential results should be taken to be in 'background knowledge' and hence fail to be possible confirmers of new theories, according to this historical approach? Musgrave distinguishes three versions of the approach, characterised, as he sees it, by three different answers to this question.

According to the first answer—which produces 'the strictly temporal view'—background knowledge contains 'all the relevant experimental results, hypotheses, *etc.*, which are "known to science" when [the] theory [in question] was proposed' (*op. cit.*, p. 8). This entails that a theory T is *only* confirmed by facts that were unknown at the time of T's initial proposal and cannot be confirmed by any evidence that was already known to hold when T was first articulated. Musgrave points out *both* that this suggestion flies in the face of quite clear-cut intuitions about some particular cases (e.g. that the General Theory of Relativity (hereafter: GTR) was confirmed by getting right the already well-known details of the precession of Mercury's perihelion) *and* that it seems difficult to discern any convincing general rationale for giving such a crucial role to purely temporal considerations.

On the second version—the 'heuristic view'—the relevant background knowledge for assessing the confirmation of theory T is, Musgrave takes it, restricted to those known facts and results that *were involved in the development* of T. This gives scope for the recapture of some of the intuitive judgments about particular cases: GTR may be confirmed by the details of Mercury's orbit, for example, provided that those details played no role in the construction of GTR (as indeed they did not). However, it is not clear, suggests Musgrave, that this account has any convincing rationale, and, in any event, it is altogether too person-relative: '[i]f different scientists take different routes to the same theory, then the evidential support of that theory as proposed by one of them might be different from its evidential support as proposed by the other.' (p. 14) And he regards this—entirely reasonably, it would seem—as in effect a *reductio ad absurdum* of the account.

Musgrave is inclined to endorse the third variant ('for my money it is the best version of the historical approach to confirmation' (*op. cit.*, p. 19)) This holds that the relevant 'background knowledge' for T consists only of the 'touchstone theory'

[2] Since I am one of those (like Hempel himself) who do not believe that *there is* a 'paradox of confirmation' (red herrings do confirm 'all ravens are black' just not very strongly!), this cuts no ice with me.

for T—in effect T's most plausible current rival. A theory T is then confirmed by any correct piece of evidence e that it entails provided that e is not also entailed by its 'touchstone' T′. Clearly there will in general be two types of such evidence: evidence that contradicts the touchstone T′ and evidence on which T′ is simply silent. On this account, GTR is confirmed by getting the details of Mercury's perihelion correct, since its rival, Classical Physics (hereafter: CP), gets those details wrong. Ditto with the Special Theory (STR) and, say, the Michelson-Morley result.[3] On the other hand, neither STR nor GTR is confirmed by any correct observational result that CP already correctly entails, even if it entails that same result in as straightforward and 'natural' a way as does CP.

2. A CLARIFICATION AND A PROBLEM WITH MUSGRAVE'S CLASSIFICATION

The main purpose of this paper will be to argue that, despite its neatness and intuitive appeal, Musgrave's whole classificatory scheme is off-beam: I shall argue for what might look like a modified version of Musgrave's second variant of the historical view, but also show that the account I favour is, when properly understood, logical rather than the historical! However a couple of detailed points about Musgrave's classification should be made beforehand.

First, a clarificatory point: Musgrave's approach, along with much of the subsequent literature (including my own contributions)[4], is focussed on one particular aspect of the general issue of confirmation—the impact of general observational or experimental results on deterministic theories that entail them. Of course this does not exhaust all confirmational issues—in particular those concerning stochastic or probabilistic theories. Although investigators such as Deborah Mayo and, more recently, Christopher Hitchcock and Elliott Sober in their (2004) have developed accounts that attempt to cover both deterministic and probabilistic cases, I think that there are problems with these accounts and will continue throughout this paper to concentrate (at least very largely) on the particular type of issue outlined. Was Fresnel's theory of diffraction better confirmed by the (novel) 'white spot' result than it was by getting the already known details about straightedge diffraction right? Was Einstein's theory better confirmed by the (novel) prediction of light-bending than it was by accounting for the already known facts about Mercury's perihelion? And so on. I recognise, of course, that statistical issues lie hidden here: real experiments and observations always show a certain amount of variation and the issue of how we get from real data to general observational or experimental results of the kind we are considering itself involves statistical considerations. Nonetheless I shall ignore these issues for current purposes and just take as a starting point the fact that certain general results

[3] Of course these judgments depend on *which particular versions* of classical physics we are considering (it is for example well known that Dicke and others eventually produced a version of CP that does yield the correct account of Mercury's motion)—therein lies much of the tale that will unfold in this paper.

[4] See my (1985) and especially (2002).

have been accepted as evidence, leading on to the question of the extent to which various deterministic theories that entail that evidence are confirmed by it.

Secondly, there are immediate questions about the completeness (or perhaps aptness) of Musgrave's classification: where, in particular, does the currently most widely held account of confirmation—personalist Bayesianism—figure within his scheme? I suppose that intuitively most philosophers of science would regard the Bayesian theory as the archetypically 'logical' approach to confirmation. Musgrave sees the logical approach as an aspect of 'modern logical empiricist orthodoxy' (*op.cit.*, p. 2) and Bayesianism certainly seems to be what eventually became of that orthodoxy, even to the extent to its being explicitly adopted by Carnap in his later years.

Bayesians standardly measure the support that evidence e lends to theory T by the difference between T's 'prior' probability, p(T), and its 'posterior' probability in the light of e, p(T,e). This appears to make Bayesian confirmation a two-place relation and hence indeed to make Bayesianism a 'logical' account on Musgrave's characterisation. However, Bayesians insist that all probabilities are *implicitly* relativised to background knowledge,[5] in fact to the background 'knowledge' of a particular Bayesian agent, where background knowledge, at least in the most straightforward account, consists simply of everything that the agent takes as evidence, ahead of the time at which we are considering the question of whether, and to what extent, the particular piece of evidence e confirms the particular theory T.

This relativisation to what a particular agent takes to be background knowledge is one—comparatively under-emphasised—source of the enormous (and in my view clearly unacceptable) subjectivity in the Bayesian approach: what a Bayesian 'agent' counts as evidence and hence puts in background knowledge is purely a matter for the agent, no less than are her 'priors'. Moreover, far from a Bayesian agent being required to justify every change in her degree of belief in a theory by appeal to the principle of conditionalisation, such an agent is entirely free (so far as the constraints of rationality are concerned) to feel at any stage that the epistemological earth has moved, that her background knowledge has changed and hence that an (in principle quite unconstrained) reassignment of all probabilities is called for. So for example there is nothing in pure Bayesian theory to rule that the following scenario involves anything that is counter-rational: a 'scientific' creationist begins with a very high prior for creationism and a very low prior for Darwinism; conditionalising on the accumulating evidence in approved Bayesian manner, however, leads her posterior for creationism to become steadily smaller and her posterior for Darwinism steadily greater; next however she receives a (perhaps further and powerful but ineffable) message from God or elsewhere that leads her to revise all her erstwhile judgments and to call for a new round of assignments of priors in the light of a radically revised background 'knowledge'; this new assignment of priors sees creationism back at a very high level and Darwinism back at a very low one. Of course we would all suspect the sincerity of such a creationist and the Bayesian trades in real, rather than

[5] Most deny that this dependence should be captured by explicitly conditionalising on background knowledge hence producing an absolute confirmation measure—see below.

merely *alleged*, degrees of belief—but were there such a sincere Creationist, the Bayesian could raise no objection to her (surely in fact irrational) belief-dynamics.

However, although such a sudden change in (personal) background knowledge is permitted in theory, when Bayesianism is applied *in practice*, things are generally made to look altogether more sensible and objective: it is quietly assumed that everyone will have the same background knowledge; that this is gradually augmented with extra material that everyone regards as evidence;[6] and that no sudden shifts of the kind just envisaged in fact occur.

Bayesianism appears, then, in practice at least, to be a version of Musgrave's historical approach—judgments about the impact of evidence e on theory T are made relative to, or in the light of, background knowledge (in practice implicitly assumed to be general amongst all competent agents at any given historical epoch). And indeed, on the most straightforward, 'natural' construal Bayesianism would fall squarely into the first, 'purely temporal' camp.

On this most natural construal, the background knowledge that the Bayesian sees as relevant for assessing the impact of some result e on T would consist of everything that is accepted as evidence (that is, assigned probability one) ahead of the question being raised of what impact particular piece of evidence e has on theory T. Suppose we are, then, asking about the confirmation of T at a time when some evidence e is already known (say that we are interested in the impact of the evidence about the precession of Mercury's perihelion on the GTR when that theory was first proposed). The fact that e is already known and accepted as evidence, entails that e will already at that time be part of background knowledge and hence that $p(e) = 1$. But this in turn immediately implies—as has been heavily emphasised under the name of the 'problem of old evidence'—that e cannot Bayesian-confirm T: if $p(e) = 1$, and T entails e, then it straightforwardly follows that $p(T,e) = p(T)$ and hence that there is, on the Bayesian account, no confirmation. As previously noted, this implication flies in the face of a number of intuitively firm judgments of confirmation in particular cases (which is why, of course, it is known as the *problem* of old evidence).

However, some Bayesians, such as Colin Howson and Peter Urbach (see their 1994), have insisted that the old evidence problem is based on a misunderstanding of the approach. If e is the evidence whose confirmational impact is under consideration at time t, then, if e is already known, that is, accepted as evidence and hence as part of background knowledge B at t, the correct background against which to make the confirmational judgement is not B itself, but rather, so to speak, $B - \{e\}$: the relative complement of B with respect to $\{e\}$, that is, the background knowledge that you 'would have had at time t, had you not known e but all else remained the same.'[7] It is, as I have argued elsewhere (especially in my (2000a)), extremely tricky (to say the very least) to make coherent sense of this counterfactual judgment. For

[6] Of course on pure personalist Bayesianism, 'evidence' (really 'evidence for the agent') is anything that the agent comes to assign probability one! (This is another massive and comparatively underemphasised source of subjectivism in the account.)

[7] See also the burgeoning literature on both Bayesian and non-Bayesian 'belief revision'.

current purposes, however, let's assume that its admitted intuitive appeal should override any formal difficulties—where does this alternative construal of the relevant background place Bayesianism within Alan Musgrave's scheme?

The answer, I think, is 'outside of it'. This version of Bayesianism *is* 'historical' in the sense that historically varying background knowledge plays a role in confirmation: it makes it entirely possible that the answer to the question 'Does e support T and if so to what extent?' may be different for different historical epochs, because of the differing content of background knowledge. But Musgrave assumes that for all versions of the historical approach, this variability will rest on the question of whether or not the piece of evidence, e, at issue is itself a part of the relevant background knowledge. On this alternative construal of Bayesianism, on the contrary, the evidence whose confirmational impact we are interested in is automatically 'subtracted' from background knowledge before the Bayesian formulas are applied. (Hence Colin Howson (see, for example, his 1990) in particular believes that Bayesianism, when properly construed, makes both the issue of when some evidence was discovered, and that of whether or not it was used in the construction of some theory, entirely irrelevant to confirmation.) The historical nature of confirmation on this alternative Bayesian view depends instead on the (ineffable) way in which background knowledge informs judgments of 'prior' probability.

3. A PROBLEM WITH MUSGRAVE'S PREFERRED VERSION OF THE HISTORICAL THEORY

It is not, then, clear that Alan Musgrave's classification scheme covers all accounts of confirmation that currently deserve serious attention and more problems in this regard lie ahead. But let's return for the present to operating within his scheme, and consider the merits of his own preferred alternative version of the 'historical account'. This asserts, remember, that the relevant background knowledge, membership of which prevents an observational or experimental result from confirming some theory T, is supplied by T's 'touchstone theory'—its most plausible current rival. This 'touchstone account' implies that GTR, for example, cannot be confirmed by any empirical result that is already entailed by CP.

But this is surely an extraordinarily counterintuitive judgement and hence not one that any sensible account would, on reflection, want to endorse. Scientists will, naturally, be especially interested in the question of whether GTR, for example, is *better* confirmed, obtains *greater* empirical support from the total evidence, than CP, and this will direct particular attention to those pieces of evidence that are entailed by Einstein's theory but not also by the classical one. But this is an issue of *extra* empirical support, not empirical support *simpliciter*. Assuming at least that both theories yield some piece of data in a 'natural' (non *ad ho*c) way (as is the case, for example, with the accounts they give of the precession of the equinoxes), then surely the reasonable judgement is that *both* CP *and* GTR are confirmed by the phenomenon—this is why the precession of the equinoxes, unlike, say, the precession of Mercury's perihelion, is irrelevant to the *comparison* of the degrees of evidential support of the two theories.

I can see no general rationale for Musgrave's preferred alternative and it certainly leads to any number of intuitively extremely awkward consequences. GTR entails the correct details of the precession of the equinoxes and it does so in as straightforward, natural, non-*ad hoc* way as does CP. Why on earth, then, should it fail to provide any confirmation for GTR just because there is another theory that also gets the phenomenon correct? Or consider what the account says not about the confirmation of the newer theory in some case of inter-theoretic rivalry but about the confirmation of the *older* one. Presumably, once GTR has been articulated, the precession of the equinoxes ceases to be a possible confirmation for CP too—since GTR now becomes the classical theory's 'touchstone' no less than vice versa. This means that while Newton's theory was confirmed by the precession of the equinoxes in, say, 1900 (when its 'touchstone theory' was what? Galileo's (very partial) mechanics? or Aristotle's more comprehensive but hopeless system?), by 1914, when nothing relevant had changed either in the theory or (of course) in the phenomena, it was no longer confirmed by those phenomena because a new theory, GTR, had arisen that equally well entailed a correct description of them. (I suppose that the 'touchstone' theorist could claim, alternatively, that the right way to judge the empirical support gained by a theory is by always taking as background knowledge that theory's chief rival *at the time it was introduced*. But this alternative is worse, *much* worse than the original. For one thing, it would disqualify the account as an historical one on Musgrave's terms—the question of a theory's confirmation by e eternally carries with it the historical context of that theory's initial articulation and hence the question becomes ahistorical! More importantly, the alternative yields even more counterintuitive results than the initial suggestion. Admittedly the alternative would have the intuitively pleasing consequence of allowing Newton's theory to retain its support from phenomena such as the precession of the equinoxes or the existence of Neptune even after the articulation of a rival that equally adequately explains them. But at the same time it would of course, all too readily, yield the judgment that, as well as there being empirical phenomena (like the Michelson-Morley experiment, and the precession of Mercury's perihelion) that support GTR but not CP, there are also phenomena (like the precession of the equinoxes or the existence of Neptune) that support CP but not GTR (because CP supplies a more demanding 'touchstone' for GTR than it itself had faced when first articulated). And all this, despite the fact that GTR entails correct descriptions of these phenomena too in a 'natural' non *ad hoc* way just as CP does![8] Surely the right judgment, as I suggested earlier, is that these phenomena support *both* theories and hence drop out of the equation when it comes to *comparing* the empirical supports enjoyed by the two theories.)

[8] Of course, as Kuhn liked to emphasise (see in particular his 1977), it does sometimes happen—especially early on in the development of some new theory—that different pieces of evidence point in different directions: e better supporting the new theory T′, while e′ better supports the older theory T. But it *clearly* cannot be the right judgment in general, that once some evidence has (fully) confirmed T, it always supplies a reason for preferring it.

We saw that Musgrave castigates the other two alternative construals of the historical approach as lacking any obvious rationale, but, as we have just now seen, his own preferred version certainly does no better. This surely should make us reflect again on the general underlying claim: Why should 'background knowledge' *in any form* be a factor in empirical support? There is of course, as just remarked, an obvious rationale for taking background knowledge to be a factor in *increased* support: if we are interested in why new theory T′ is *better* supported than its earlier rival T, then results that T has already either predicted or adequately explained will be in background knowledge and hence *may* drop out of the equation [9]—if e already confirms T then the fact that it also confirms T′ *may* provide no reason to prefer T′ over T. But it seems difficult to see why the fact that an empirical result e is already in background knowledge *in any sense* should by itself totally rule out e as support for some newly proposed theory T′, in the *non-comparative* sense of support.

Alan Musgrave's residual Popperianism leads him to claim that a justification for giving background knowledge this central role might be developed by considering which bits of evidence do or do not supply a proper *test* of the theory concerned. The suggestion is that for some reason results already in background knowledge at the time of T's proposal cannot provide a test of T. But why should this latter claim be true? If we already know that e holds rather than some alternative result of the experiment or observation it describes, then of course the fact that it turns out that some new theory T entails e rather than any of the alternatives will not have us on the edges of our seats wondering if the theory might turn out to be refuted by this particular experiment or observation. In that sense there is no test from old data. But why should that sense have the slightest epistemic relevance? The new theory is by no means *a priori* guaranteed to entail correct descriptions of all the phenomena equally well dealt with by its predecessor. (Indeed if Popper's account of new theories as 'bold conjectures' were true, it would be a miracle if this happened in a field where the old theory had had any considerable degree of empirical success). Still less is there an *a priori* guarantee that the new theory will get right all known phenomena, whether or not dealt with successfully by its predecessor. And indeed few, if any, theories do get *all* known phenomena correct (at least when first proposed). There is, then, a clear sense in which such a theory *was* tested by the already known data: it might have entailed different data that contradicts that actually recorded, but in fact it did not. If a theory might perfectly well have got some already known phenomenon wrong, but in fact got it right, then it seems perverse to rule ahead of time that this success fails to count as surviving a 'test', and so cannot yield any degree of empirical support for that theory.

Alan Musgrave's preferred solution of the prediction versus accommodation problem is, I claim, wrong; and, as so often happens in philosophy, this is because he has got the problem wrong.

[9] The fact that it only *may* drop out of the equation is important: if T provided only an *ad hoc* accommodation of e, while T′ genuinely predicts e (in the non-temporal sense, see *below*) then, on the account that I favour, e may, on the contrary and far from dropping out of the equation provide an important reason for preferring T′ over T.

4. THE REAL PROBLEM: PREDICTION VERSUS ACCOMMODATION

The problem is not whether new evidence counts more than old—it doesn't (at any rate it doesn't just because it's new). The problem is *adhoc*ness (indeed the real problem is perhaps seeing that the *adhoc*ness problem is the *only* problem in this area).

In the early 19th Century, the classical wave theory of light predicted the results of various diffraction or interference experiments. Intuitively these results told very strongly in favour of this theory against its then rival—the emission or corpuscular theory of light. Yet, as we would expect on Duhemian grounds, the emissionists by no means immediately surrendered. Duhem emphasised that single 'isolated' theories such as the corpuscular theory have no empirical consequences of their own, but achieve them only when conjoined both with specific assumptions (answering the questions: what velocities do the light-corpuscles have? and what masses? most importantly, what forces are they subjected to in particular circumstances?) and with further auxiliary and instrumental assumptions. It follows that there is always logical leeway for holding onto the central theory in the light of experimental 'anomalies' and looking to modify either a specific or auxiliary assumption. 18th and 19th century corpuscularists duly obliged—some postulated, for example, a force of diffraction, exercised on the light-corpuscles as they pass the edges of any 'gross' opaque object; others considered the possibility that the fringe phenomena that wave theorists attributed to interference and/or diffraction were in fact physiological phenomena. Although in this case it was never achieved, it clearly has to be possible in principle for the emissionists to have given themselves an expression for the 'force of diffraction' with so many (initially free) parameters that any given particular fringe phenomenon (or finite set of such phenomena) could have been accommodated. Certainly by appealing to (unknown) physiological facts about vision an entirely cheap corpuscularist 'explanation' was suggested at the time and could have been developed in some detail.

Or consider another case where this sort of dodge definitely works. ('Works' in the sense that it produces a theory that yields the accommodated data, not of course in the sense that it produces a scientifically respectable theory that does so.) The fossil record looks like strong confirmation of the Darwinian theory of evolution. (Of course the situation is less straightforward in this example because that theory does not actually deductively entail any particular aspect of the fossil record, but this is inessential to the point at issue.) As is well known, however, it is trivially easy for the 'scientific' creationist to 'match' this success. All that she needs to do is follow Gosse and assert that God decided, when creating the Universe in 4004 BC, to include some pretty pictures in some rocks that look awfully like the marks of the skeletons of now extinct organisms but are in fact *just* pretty pictures, and to include some buried bone-*like* objects that seem to fit together to form the skeletons of impressive and now extinct creatures but are in fact just artefacts, and so on. She will thus create a version of 'scientific' creationism that entails the correct facts about the (now alleged) 'fossil record', but clearly it would be absurd to hold that this requires us to abandon the view that this record supports the Darwinian theory over its rival.

There is a long tradition in science of deeply engrained distrust of such *ad hoc* moves. We surely require an account of the confirmation of theories by evidence that underwrites the judgement that the interference effects continued to give more empirical support to the wave theory in the early 19th century even once it had been indicated that emissionist accounts of those effects could be constructed, and similarly underwrites the judgment that the fossil record continues to give good empirical reason to prefer the Darwinian theory even after creationists have availed themselves of the 'Gosse dodge'. But how *exactly* are we to capture these judgments within a generally defensible account of confirmation?

The obvious initial suggestion is to say that no theory can be confirmed by evidence that it has simply accommodated in this *ad hoc* way, where the advocates of the theory have taken the evidence at issue as given and *used it* to produce a specific version of their favoured theory that yields that evidence. At least when the notion is used liberally, these are all exercises in parameter-fitting. The idea behind the 'diffracting force' emissionist account of fringe-phenomena was to start from a very complicated expression for that force as a function of the distance from the diffracting object (allowing this to be attractive at some distances and repulsive at others) and then use particular fringe measurements to fix those parameter values so that the required phenomena are entailed. Similarly, the Creationist's general theory—that God created the Universe in 4004 B.C. 'essentially' as it now is—effectively gives the Creationist a whole series of 'free parameters' that specify how *exactly* it was that God chose to create the universe: if you observe particular patterns in some rocks, then that specifies one part of God's creation, this 'parameter' value is tied down on the basis of the observation and this, unsurprisingly, produces a specific theory that entails the observed data—the theory being of course that God created the Universe, not just any old how, but in particular with these patterns in these rocks.

The positive side of the account would then be that a theory is confirmed by any piece of data a correct description of which it entails, provided that the evidence was not used in the construction of the specific version of the theory that entails it, *whether or not* the data was already known. There appears to be, then, an important methodological distinction between accommodation and prediction in the general sense in which it is generally used in science (meaning simply that some evidence follows from a theory without having needed to be accommodated within it)[10].

The most straightforward way to capture this difference would, of course, be by ruling that theories are confirmed only by predictions (understood as not requiring novelty) and not at all by accommodations. This amounts, it would seem, to the

[10] Here for example is an especially clear passage from French's excellent textbook on Newtonian Mechanics: '[L]ike every other good theory in physics, [the theory of universal gravitation] 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 recognising 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.' (French 1971, pp. 5-6)

'heuristic account' as Musgrave characterises it—namely, the version of the historical approach to confirmation which identifies those results that belong in background knowledge and hence cannot confirm the theory as those that have been used in the construction of the theory. The 'heuristic account', as so construed, is also sometimes known as the 'no double use' or 'use novelty' rule.[11]

I shall in fact argue, first, that this 'most straightforward way' of underwriting the prediction/accommodation distinction is altogether too straightforward to be true; and secondly, that the correct way to underwrite the distinction and hence arrive at the correct account of confirmation of the sort here at issue produces a view that cannot properly be regarded as a version of Musgrave's historical approach. However, it should be noted that even in its most straightforward form, the 'no double use' account seems to have some immediate attractions. *First,* it accords with a range of intuitive judgments about particular cases (one such is the precession of Mercury's perihelion and the GTR) where 'old evidence' is taken to provide strong support for a theory: provided that the facts about Mercury's orbit were not involved in the construction of its explanation within GTR, then there is no reason, on this account, to deny that those facts support GTR. And *secondly* (and of course relatedly) the account relegates the time-order of theory and evidence *in itself* to what it should be—namely, a complete historical irrelevance (what possible *general* justification could there be for old evidence always to count less? why give such an epistemic role to what may have been a mere historical accident?).

However, despite these attractions, the 'no double use' rule has been alleged to face at least two fundamental objections of its own. The objection that Musgrave himself cites, as we already noted, concerns the fact that it seems to make theory-confirmation an unacceptably relativistic (enquirer-relative) affair:

> If different scientists take different heuristic routes to the same theory, then the evidential support of that theory as proposed by one of them might be different from its evidential support as proposed by the other. In short, Zahar's ['heuristic' or 'no double use'] view makes confirmation a person-relative affair. (*op. cit.*, p. 14)

An even more frequently voiced criticism of this view is that, just like the purely temporal view that it attempts to replace, it flies in the face of deeply held intuitions about particular cases. Nickles, Mayo, Howson and others[12] have all pointed to cases in which evidence e was used in the construction of some theory T and yet where e was taken to provide (strong) support for T. As Colin Howson, for example, claimed (*op. cit.*, p. 231) 'counterexamples abound to' the idea that evidence used in the construction of a theory cannot be used in its support, and indeed 'can be invented ad lib'. In the next section, I address this second objection—therefore cunningly renamed 'objection one'. I then show how to develop the idea underlying the no double use rule so to produce an account that escapes objection one, and then I will show how the developed view also overcomes Musgrave's objection (now 'objection

[11] See for example Nickles (1987)
[12] Nickles *op.cit,* Mayo (1996) and Howson (1990)

two').[13] Finally, I will show how this developed view is not properly regarded as a version of the historical approach as Alan Musgrave construes it.

5. OBJECTION ONE AGAINST THE 'HEURISTIC' VIEW: USED DATA SOMETIMES (STRONGLY) CONFIRMS

Allan Franklin once gave a seminar talk at the LSE under the title '*Ad hoc* is not a four letter word'. Underneath the (multiple) surface correctness of this title, there lies a somewhat deeper but no less correct point: scientists entirely legitimately use data all the time in the construction of their theories. If, to take the most clear-cut case, general theoretical considerations leave open the value of some important parameter, then how else would a scientist tie down that parameter's value *except* by using data? The only other alternative that seems open would be to conjecture a value and then test (and then re-conjecture when the test is failed as it almost inevitably will be, and then re-test…)—but this attempt to find a needle in a (generally infinitely large) haystack would be madness. Here is one extremely simple but canonical instance of the systematic use of data in theory-construction.

Suppose a mid-19th Century scientist already accepted the *general* wave theory of light—the theory that light from any particular source consists of waves of some wavelength or other transmitted through the luminiferous aether. This general theory does not specify the wavelength of any particular kind of monochromatic light—say light from a sodium arc. The scientist would like a more detailed theory that *does* specify that wavelength. Rather than attempt to conjecture a value, she would 'deduce' the specific theory, involving the specific value of the wavelength, 'from the phenomena'. She would look for some consequence, e, of her general theory T, where e characterises some observable magnitude (fringe separation in some particular experiment, say) as a one-to-one function of the wavelength. She would perform the experiment using light from a sodium arc, measure the magnitude at issue—here, the fringe separation (call the result of this measurement e′)—and infer to a more specific theory T′. So, for example, subject to a couple of idealisations, it follows from the general wave theory that, in the case of the famous two-slit experiment, the (observable) distance X from the fringe at the centre of the pattern to the first fringe on either side is related to (theoretical) wavelength λ, via the equation $X/(X^2 + D^2)^{1/2} = \lambda/d$ (where d is the distance between the two slits and D the distance from the two-slit screen to the observation screen—both of course observable quantities). It follows analytically that $\lambda = dX/(X^2 + D^2)^{1/2}$. But all the terms on the right hand side of this last equation are measurable. Hence particular observed values will determine the wavelength (within of course some small margin of experimental error), and so determine the more specific theory T′, with the parameter that had been free in T now given a definite value—again within a margin of error. Far from being scientifically questionable, this is, to repeat, entirely standard (and patently legitimate) scientific procedure.

[13] My treatment here follows and builds upon that given in my (2002)—actually written for a conference in 1999.

Several of the most celebrated episodes from the history of science involve using data (often anomalous data for an earlier theory) to construct a new theory. For example, Adams and Leverrier used the data from Uranus's orbit that had proved inconsistent with the initial Newtonian account essentially as follows. They took it that the basic Newtonian theory (of mechanics plus universal gravitation) was correct, and then worked backwards from the initially anomalous Uranian data to figure out what assumptions would have to be made about a further trans-Uranian planet, such that, when that further planet's gravitational interaction with Uranus was taken into account (along of course with the gravitational interaction with the sun and the other, already known planets), the overall Newtonian theory would ascribe the correct orbit to Uranus. This manoeuvre, as is well known, led to the discovery of Neptune—one of Newtonian theory's greatest successes and indeed one of the most impressive confirmations of any theory in the history of science.

So how, in the light of facts like these, could anyone have defended the 'heuristic account' of confirmation, committed, as it seems to be, to the view that evidence used in the construction of a theory *can never* confirm it? In the specific case from optics that I just sketched, there is a very clear sense in which e, the fringe data used in the construction of the more specific wave theory T′ supports that theory: *given that* the general theory T has already been accepted, e *deductively entails* T′, and what better support could there be than deductive entailment?

Colin Howson likes to emphasise a still more general sort of case—standard statistical examples such as the following (see again his 1990). We are given that an urn contains only black and white balls though in an unknown (but fixed) proportion; we are prevented from looking inside the urn but can draw balls one at a time from it. Suppose that a sample of size n has been taken (with replacement) of which k have been found to be white. Standard statistical estimation theory then suggests the hypothesis that the proportion of white balls in the urn is $k/n \pm \varepsilon$, where ε is calculated as a function of n by standard confidence-interval techniques. The sample evidence is the basis here of the construction of the particular hypothesis, and surely also supports that particular hypothesis at least to some (good) degree—the evidence for the hypothesis just *is* that a proportion k/n of the balls drawn were white.

Deborah Mayo cites and analyses in more detail the same case and also cites the following 'trivial but instructive example' (1996, p. 271). Suppose one wanted to arrive at what she characterises as 'a hypothesis H' about the average SAT score of the students in her logic class. She points out that the 'obvious' (in fact uniquely sensible) way to arrive at H is by summing all the individual scores of the n students in the class and dividing the result by n. The 'hypothesis' arrived at in this way would clearly be 'use-constructed'. Suppose the constructed 'hypothesis' is that the average SAT score for these students is 1121. It would clearly be madness to suppose that the data used in the construction of the 'hypothesis' that the average SAT score is 1121 fails to support that hypothesis. On the contrary, as she writes (*ibid.*):

> Surely the data on my students are excellent grounds for my hypothesis about their average SAT scores. It would be absurd to suppose that further tests would give better support.

Exactly so: the data provide not just excellent, but, short of some trivial error, entirely *conclusive* grounds for the 'hypothesis'—further tests are entirely irrelevant. (This is precisely why it seems extremely odd to talk of a 'hypothesis' at all in these circumstances—a point to which I will return *below* in my more extensive consideration of Mayo's views.)

6. RESPONSE TO OBJECTION ONE: TWO SORTS OF CONFIRMATION

Does the admission that these sorts of 'deductions from the phenomena' (such as the deduction of the specific version of the wave theory T′ from the general wave theory T plus fringe data e) provide clear-cut cases of theories that *are* supported by data used in their construction spell the end for the heuristic account of confirmation?

To start to see that the answer is 'no', consider again the 'Gosse dodge' within 'scientific' Creationism, or indeed any of the other standard cases of blatantly *ad hoc* moves in defence of a theory that have been cited in the literature.[14] In all these cases, the specific theory is 'deduced from the phenomena'—meaning, as always, of course deduced from the phenomena *plus already accepted general principles*.[15] 'Deduction from the phenomena' is a very powerful technique in the case where the necessary general principles are indeed *generally* accepted and therefore, presumably, themselves have strong evidence in their favour. But what if, on the contrary, the necessary 'background principles' are not universally accepted as based on sound evidence, but instead accepted only by some group or other, one with its own particular axe to grind?

If you *were already convinced* of the general Creationist claim that God created the Universe 'essentially' as it now is in 4004 B.C., then the data that your irritating Darwinian supporters insist on calling the 'fossil record' do of course deductively Those data thus give you not only good but 'essentially' *conclusiv*e reason to accept this particular version of the general theory that you already accepted on other entail[16] the more specific version of your theory that says that part of God's creation was some pretty pictures in the rocks and buried bone-like artefacts, and so on.

[14] Another favourite example that I and others have used elsewhere is provided by Immanuel Velikovsky's famous theory that a large chunk of Jupiter broke away and careered towards the Earth, orbiting it on a series of occasions before (somehow or other) settling down to a quieter life as the planet Venus. Velikovsky saw these close encounters with this 'comet' as the explanation for 'events' 'recorded' in the Old Testament—such as the parting of the Red Sea and the fall of walls of Jericho. Velikovsky recognised that other contemporary record-keeping cultures ought, in that case, to have recorded cataclysms on a similar scale, since such amazing effects of the 'comet' were unlikely to have been confined to the particular area of the Middle East covered by the Old Testament scribes. He found one or two (arguable) confirmations, but several altogether more clear-cut apparent refutations. But Velikovsky rose to the task, arguing that in the cultures that otherwise kept records the cataclysmic events associated with the 'comet' had proved *so* cataclysmic that 'collective amnesia' had set in there. Of course he read off which particular record-keeping cultures had suffered from this unfortunate complaint precisely by noting which ones had no records of suitable cataclysms.

[15] 'Deduction from the phenomena' was of course Newton's preferred method of theory-construction (and of avoiding 'hypotheses'). For Newton on deduction from the phenomena and references to the literature, see my (2000b).

[16] It is admittedly only a more or less deduction—it would be a valid deduction only if the Creationists assumed that the world now is exactly as it was when god created it, but of course even they have to admit that there has been some change (hence the 'essentially' as it now is).

grounds. In this regard the case is surely no different from the (intuitively more scientifically respectable) case of the early 19th Century optical scientist, who, being already convinced of the general wave theory, deduces from the phenomena the more specific version with specific wavelengths for light from particular monochromatic sources: in this latter case too, *given* that she accepts the general wave theory, T, the fringe data, e, give her (in this case entirely) conclusive reason to accept the particular version of the theory T′, involving a now fixed value of an initially free parameter.

But the natural reaction to the Creationist/Gosse dodge case is surely that while the 'fossil record' data may indeed give you reason, let's say conclusive reason, to adopt the particular Gossefied version of Creationism, this is an ineliminably conditional judgment—the evidence gives you absolutely no reason to have adopted the general Creationist view in the first place. If you are going to be any sort of Creationist at all, then this data gives you as solid a reason as could be for being a Gosse-dodge-Creationist, but it gives you absolutely no reason to be any sort of Creationist at all! There is no reason to think that the general underlying theory itself obtains any empirical support just because the specific version of it entails the correct empirical data.

What is sauce for the goose is sauce for the gander. Exactly the same judgment is valid in the (intuitively scientifically respectable) wave theory case: the fringe data, e, give you solid (indeed conclusive) reason to believe T′ (the wave theory with a specified wavelength for monochromatic light from a sodium arc), *provided* that you have already accepted the general wave theory (with free parameter), but give you absolutely no reason to accept the general wave theory in the first place. Both in this—seemingly legitimate—case and in the, apparently illegitimate case of the Gosse dodge, the correct judgment seems, then, to be twofold: *first* that, if the general underlying theory is taken as given, then if e is used in the construction of a specific version of that general underlying idea, e gives very strong (perhaps conclusive) support for the specific theory; however, *secondly*, there is no support from that evidence for the general, underlying theory itself.[17]

The difference between the two cases seems clearly to be that while there were *other*, *independent* empirical reasons for taking the general wave theory of light seriously, there are no such reasons in the case of 'scientific' creationism. There was already good reason to accept the general wave theory with the free parameter,

[17] A similar remark also applies to Colin Howson's statistical examples: so long as the basic theory or 'model' is given (basically in his urn case, that we are dealing with a 'Bernouilli process' with fixed, but unknown parameter p (the proportion of white balls in the urn)), then the evidence that k/n of the sampled balls were white gives support (in this case of course not conclusive) for the specific theory that estimates p as lying in the interval $k/n \pm \varepsilon$. But that data gives no conceivable reason for having greater faith in the idea that this is the correct model. (Indeed this is not an issue that would normally even arise in that case.)

ahead of any measurement of fringe distances with light from the sodium source. Hence, when evidence e turns out deductively to entail the specific theory T′ (complete with filled-in value for the wavelength of light from the sodium arc) *given* T, we can 'discharge the antecedent' and infer that e gives us (of course some, defeasible) reason to accept T′ *full stop*. In contrast, in the Gosse dodge case, exactly because there is no independent reason to accept the underlying general Creationist account, the fact that the fossil record entails the Gosse dodge variant of Creationism, justifies *only* the *conditional* judgment that e gives us reason to accept the Gosse dodge variant to the extent that (but only to the extent that) we already have reason to accept the general theory.

But how exactly can these general underlying theories earn their independent empirical support, as, if the line I am defending is correct, in some cases they must do? After all, the Duhem problem is exactly posed by the fact that such general theories do not have directly checkable empirical consequences of their own. All empirical tests of the wave theory of light, for example, are tests of the general wave theory *plus* particular assumptions. It seems, then, that if the dual approach to confirmation that I am outlining is to be at all coherent, there must be a contrast class to the sorts of cases we have considered so far. That is, there must be some empirical tests, the results of which not only confirm the specific version of the theory that entails the results of those tests, *but also* confirm the underlying general theory (despite the fact that that general theory does not entail those results on its own). It must be the case, in other words, that scientists do sometimes take it that the empirical success of some particular version of a general theory gives good reason to accept the general theory itself —and in particular good reason to seek to develop another specific theory for a different field of phenomena based on that same general theory.[18]

Certainly this seems to be an actual feature of scientific practice: for example, the discovery of Neptune seems to have been regarded as a success not just for the particular Newtonian model of the universe (now involving Neptune), but also for Newtonian gravitational theory itself. Similarly, returning to optics, both the (new) white spot result and the (long known) straightedge diffraction experiments were taken to support not only Fresnel's specific wave theory of diffraction that entailed them, but also the general theory of light as waves in an elastic medium on which it was based. Hence these phenomena, although following only from the specific wave theory of diffraction, were taken as providing good reason to develop another specific theory based on the same general elastic medium wave theory to deal with the quite separate phenomena of polarisation and crystal optics. (See my 1989.)

My claim, then, is that scientists do not restrict themselves simply to judgments of the conditional kind that we highlighted—judgments to the effect that, *against the given background* of some general framework theory, some piece of evidence e gives strong support to some specific version of the general theory. They *also*

[18] This in practical terms seems to me the main work that confirmational judgments do in the development of science.

sometimes see the general framework theory as empirically supported. Yet, as Duhem showed us, such support must always be achieved, not directly, but *via* specific versions of the general theory (*i.e.* not the general theory alone but that theory plus some further assumptions). Some, but not all, types of empirical success must somehow spread from the particular theory that directly enjoys them to the underlying general theory.

What kinds of empirical success turn this second and stronger confirmational trick? The answer, I think, is two kinds, of which the more straightforward is the following. A scientist starts with some general theory T, uses e to fix some parameter in T and thus creates (by 'deduction from the phenomena') the more specific theory T′; T′then goes on to make some *further* independent prediction e′. If e′ is experimentally verified then this confirms not only the specific theory T′ but also the underlying more general theory T. This is exactly what happens in our first wave theory case: once the parameter corresponding to the wavelength of light from a sodium arc has been fixed using the fringe distances in the two-slit experiment, the more specific theory thus created can then go on to be directly tested in *other* experiments using light from the same source (notably the single slit diffraction experiment). (It is standard to talk of 'overdetermination' of parameter-values in such cases: the initially free parameter could be fixed using *any one* of a range of experimental results and the specific theory with fixed parameter would then proceed to entail the rest of that range of results.)[19]

A more significant episode in the history of the wave theory illustrates the same lesson. The result of the experiment of Fresnel and Arago—that the interference fringes in the two slit experiment disappear when the light from the two slits is oppositely polarised through the interposition of suitably oriented quartz plates—more or less forces the wave theorist to adopt the view that the wave motion in light occurs *at right angles* to the direction of propagation, rather than *along* the direction of propagation, as previously believed. ('More or less' because you can deduce the specific tranverse wave theory from the general theory (light is some sort of wave in a medium) plus the Fresnel-Arago result, only if you add some further extra assumptions, that are, however, entirely 'natural'.[20]) The Fresnel-Arago result then very strongly confirms the tranverse version of the wave theory in the first (conditional) sense—if you have already accepted the general wave theory then the

[19] Alan Musgrave too highlights the importance of independent testability and independent evidence (*op. cit.,* p. 6) But he takes it that the idea that scientific theories require not just testability, but *independent* testability to be accepted is captured by his favoured third variant of the historical approach: T is independently testable through any of its empirically checkable consequences that are not also consequences of its 'touchstone' T′. But as we are now seeing the really important idea is not one involving a comparison between theories, instead a single theory is independently tested by any piece of evidence that it makes a prediction about, provided that evidence was not 'written into' the theory in advance.

[20] Light waves could for instance in principle have *both* a transverse *and* a longitudinal component. However the fact that this Fresnel-Arago result (along with others) shows that any longitudinal component could have no observable effect means that simplicity dictates it be rejected. (It is in this particular sense, rather than any nebulous general way, that simplicity judgments play an important role in science.)

result shows (pretty well) that tranverse waves are what you *must* plump for. However, the fact that—having in effect deduced the transverse wave version from the Fresnel-Arago result—that experimental result can in turn be deduced from the transverse version of the theory would clearly give anyone unconvinced of the general wave theory no further reason to adopt it. But Hamilton saw that the transverse wave theory made predictions about the wave surface in particular types of birefringent crystal and hence about certain phenomena in crystal optics that are quite independent of the initial Fresnel-Arago result; and these predictions were successfully tested by Lloyd. These crystal optics results represent exactly the sort of *independent* evidence that, unlike the Fresnel-Arago result, *does* support not only the particular theory that entails it but also the general underlying approach—they *do* give the unbeliever extra reason to adopt the wave theory approach in general.

Finally, in the famous Newtonian case, using the (initially anomalous) data from Uranus's orbit to fix (in fact, in this case, re-jig) a parameter about the number of other planets affecting that orbit produces a theory that turns out to entail an independently checkable prediction about the existence of a further (and hitherto unrecognised) planet. Confirmation of this prediction in the form of the discovery of Neptune supports not only the specific version of Newtonian theory, partially created from the Uranian data, but also the general Newtonian theory itself. So the 'prediction' of the Uranian data gives only the first, conditional sort of support for the specific Newtonian model, while observations of Neptune yield the stronger kind of support that reaches the general theory by 'confirmational osmosis'.

These cases, then, exhibit the *first* type of stronger confirmation—*independent* evidence. The second type is equally important. This sort of confirmation (again: of the general underlying theory, rather than of some specific theory, *given* the general underlying theory) is provided in cases in which, roughly speaking, some prediction 'drops out of' the basic idea of the theory. Here's an example.

The explanation of the phenomena of planetary stations and retrogressions within the Ptolemaic geocentric theory is often cited as a classic case of an *ad hoc* move. The initial geocentric model of a planet, Mars, say, travelling on a single circular orbit around a stationary Earth, predicts that we will observe constant eastward motion of the planet around the sky (superimposed, of course, on a constant apparent diurnal westward rotation with the fixed stars); this is directly refuted by the fact that Mars' generally eastward (apparent) motion is periodically interrupted by occasions when its gradually slows to a momentary halt and then begins briefly to move 'backwards' in a westward direction, before again slowing and turning back towards the east. The introduction of an epicycle of suitable size and the assumption that Mars moves around the centre of that epicycle at a suitable velocity while the whole epicycle itself is carried around the main circular orbit (now called the deferent) leads to the correct prediction that Mars will exhibit these stations and retrogressions. Although not as straightforward as normally thought, this case surely is one that fits our first, entirely conditional, kind of confirmation—if you *already accept* the general geocentric view, then the phenomena of stations and retrogressions give you very good reason to accept (and in that sense they strongly confirm) the particular version of geocentricism involving the epicycles.[21] However the fact that stations and retrogressions are 'predicted' (better: entailed) by the specific version of geocentricism with suitable epicyclic assumptions gives absolutely no further reason to accept (and so no support for, or confirmation of) the underlying basic geocentric (geostatic) claim.

The situation with Copernican heliocentric (or again, better, heliostatic) theory and planetary stations and retrogressions is, I suggest, entirely different. According to the Copernican theory we are, of course, making our observations from a moving observatory. As the Earth and Mars both proceed steadily eastward around the sun, the Earth, moving relatively quickly round its smaller orbit, will periodically overtake Mars. At the point of overtaking, although both are in fact moving consistently eastward around the sun, Mars will naturally *appear*, as observed from the Earth, to move backwards against the background of the fixed stars. Planetary stations and retrogressions rather than needing to be explained *via* specially tailored assumptions ('having to be put in by hand' as scientists sometimes say), drop out naturally from the heliocentric hypothesis. Copernican theory, in my view, genuinely *predicts* stations and retrogressions even though the phenomena had been known for centuries before Copernicus developed his theory. (I am talking here about the qualitative phenomenon not the quantitative details which, as is well known, need to a large extent to be 'put in by hand' by both theories—and courtesy of multiple epicycles in Copernicus no less than in Ptolemy.[22])

The way that Copernicus's theory yields stations and retrogressions may, indeed, seem to be *so* direct that it challenges Duhem's thesis: doesn't the basic heliocentric hypothesis on its own, 'in isolation', entail those phenomena? This is a general feature of the sort of case I am trying to characterise: the way that the confirming phenomenon 'drops out' of the basic theory appears to be so direct that scientists are inclined to talk of it as a direct test of just the basic theory, in contradiction to Duhem's thesis. But we can see that, however tempting this judgment might seem, it cannot be literally correct.

No theory T, taken 'in isolation', can deductively entail any result e, if there is any assumption A which is both self-consistent and consistent with T and yet which together with T entails not-e. So in the case we are considering, if the basic Copernican theory alone entailed stations and retrogressions, then there would have to be no possible assumption consistent with that basic heliocentric claim that, together with it, entailed that there would be no stations and retrogressions. But

[21] This is often thought of as the archetypically *ad hoc* move (epicycles are almost synonymous with ad hoccery). However the Ptolemaic move does produce an independent test (and indeed an independent confirmation) but not one that, so far as I can tell, was ever recognised by any Ptolemaist. It follows from the epicycle-deferent construction that the planet must be at the 'bottom' of its epicycle and hence at its closest point to the Earth exactly at retrogression. But this, with other natural assumptions, entails that the planet will be at its brightest at retrogression—a real fact, that can be reasonably confirmed for some planets with the naked eye. (Of course even had it been recognised, this test would not have been reason to continue to prefer Copernicus over Ptolemy, since, as will immediately become apparent, the former too entails—in an entirely non *adhoc*—way that the planet is at its nearest point to the Earth at retrogression.)

[22] See, for example, Kuhn (1957)

there *are* such possible assumptions. Suppose for example that the earth and Mars are orbiting the Sun in accordance with Copernicus's basic theory. Mars happens, though, to 'sit' on an epicycle, but only starts to move around on that epicycle when the Earth is overtaking Mars and does so in such a way as exactly to cancel out what would otherwise be the effects of the overtaking (that is, the station and retrogression). Of course this is a monstrous assumption—but it is both internally consistent and consistent with the basic heliocentric view. The existence of this assumption implies that, contrary to first impressions, Duhem's thesis is not refuted in this case: the heliocentric hypothesis *alone* does not entail the phenomena.

However those first impressions and the monstrousness of the auxiliary necessary to 'prevent' the entailment of stations and retrogressions both reflect just how 'natural' the extra assumptions are that are necessary for heliocentricism to entail the phenomena. All that needs to be assumed, in addition to the basic idea that Mars and the Earth are both orbiting the sun, is that the Earth (which has an observably smaller average period) moves relatively quickly round its smaller orbit and hence periodically 'laps' Mars. (Many philosophers—including both Duhem and Quine themselves—have been overimpressed by Duhem's arguments. There is nothing in those arguments that favours 'holism' in any serious sense, nor that contradicts the idea that some predictions require *fewer* auxiliary assumptions than others.)

A similar case is again provided by the classical wave theory of light. Fresnel's account of diffraction is so natural within the context of the general idea that light consists of periodic motions transmitted through an elastic medium, that he was led to suggest that no auxiliary assumptions are involved:

> I am … going to show that one can give … a general theory [of diffraction] within the system of waves *without the aid of any secondary hypothesis,* by depending on the Huygens principle and that of interferences, which are one and the other consequences of the fundamental hypothesis. (Fresnel 1819, pp. 282-3; my translation and emphasis.)

However, without going into the details, the same message applies here as in the heliocentric theory. The 'direct test' or 'no auxiliary needed' view cannot be literally correct but it is easy to see why Fresnel claimed it was—the great plausibility of the claim reflects the naturalness of the auxiliary assumptions that were in fact necessary.

So in summary, the *real* heuristic view of confirmation that emerges from this consideration of objection one and that I want to defend is as follows:

Two types of confirmation need to be distinguished. *First* a purely intra-paradigm or intra-research programme judgement—e supports specific theory, T′, *relative to* a given general theoretical background T. The most straightforward case is where e, in conjunction with T, deductively entails T′. Even manoeuvres that are patently *ad hoc* (in the pejorative sense) produce specific theories that are confirmed in this (ineliminably) conditional sense. The *second* type of confirmation, unlike the first, produces support not only for the specific theory that entails the phenomenon at issue, but also for the general underlying theory which does not. There are in turn two cases in which this second type—call it 'unconditional support'—is produced: (i) cases of independent evidence (e entails T′ modulo T, but then T′ turns out also to predict e′ which is experimentally verified) *and* (ii) cases where e 'drops naturally out' of T (or, if you like, where the T′ that really entails e is the 'natural version' or 'natural extension' of the underlying general theory T).

I am confident that this dual account of confirmation captures all the intuitive judgments that have been cited in this debate, *both* those used to support the heuristic account or 'no double use' rule in its original formulation *and* those used by critics of that view as originally formulated. Is that all that can be said in its favour or can the heuristic account also be given a plausible general rationale?

The justification of the first (conditional) sort of judgement of confirmation is surely straightforward. If e deductively entails the specific theory T′, given the more general theory T, then e confirms T′ for anyone who already accepts T in the clear sense that it supplies conclusive reason for also accepting the more specific theory T′ (and, in cases of 'near deduction', e supplies a very strong reason for accepting T′, given that the background general theory T is already accepted). This first sort of confirmation in a clear-cut way 'passes the confirmational buck': e, in these cases, demonstrates that you ought to have exactly as much (or, in the 'near deduction' case, almost as much) confidence in T′ as you have in T (despite, of course, the greater content of T′). From outside the 'paradigm', this sort of confirmation shows that T and T′ are, given the evidence e, epistemically inseparable—they stand or fall together.

As for the justification of the second, unconditional and hence more powerful, sort of confirmation, here, for all philosophers' fancy talk, we are, I think, just thrown back on the basic, intuitive 'no miracles' consideration (despite feeling its force, I have always thought that 'no miracles *argument*' was an overly flattering description). The two types of case—of independent evidence and evidence that 'drops out of the basic idea'—that are identified by my account as producing this type of confirmation are exactly the sorts of case that elicit the no-miracles response: 'surely it would be a miracle if the theory could have such evidence in its favour and yet be somehow entirely off-beam?' We are, of course, from the point of view of deductive logic, as always with ampliative inference, committing some version of the fallacy of affirming the consequent. That is why we need to be circumspect about the conclusion to be drawn. This conclusion should not, of course, be that the theory is true (the history of science would soon put paid to that conclusion), nor yet I think, even in an intuitive sense, 'approximately true', but rather 'along the right lines'—probably destined to have its structure preserved, perhaps in approximate or limiting case form, in later successful theories. I do not claim that this is much of a justification; I do believe that it is the only justification we can ultimately give for any account of the confirmation of theories by evidence.

7. OBJECTION TWO AND THE RESPONSE TO IT: SAME THEORY, SAME EVIDENCE, SAME CONFIRMATION

The objection to the 'heuristic' account raised by Alan Musgrave himself—now 'objection two'—was, remember, that the account is unacceptably investigator-relative.

Reformulated to take account of the distinction that I have just now emphasised, the objection goes as follows. Two scientists, A and B, employ two different methods of construction—A uses evidence e, B does not; nonetheless A and B still arrive at the very same theory T; when that theory 'turns out' to entail e, e will confirm—in the strong unconditional sense—theory T as constructed by scientist B, but not as constructed by scientist A (who will, instead, obtain only the conditional sort of confirmation from e). But this, so the objection goes, is surely ridiculous—if they arrive at the very same theory then surely that theory ought to receive the same confirmation from any piece of evidence including e, independently of the way the theory was arrived at. Hence, since the account has a ridiculous consequence, it cannot be correct.

How could the two-scientist story that underlies the objection ever in fact be realised? It cannot be emphasised sufficiently that 'means of construction' is, in the mature sciences at least, *not* a personal notion—finding out about it does *not* require combing through a scientist's personal diaries and the like. It depends instead on the research programmes involved. And these programmes can be articulated and objectively assessed.

The most straightforward way in which two different scientists might take different routes in trying to develop a theory for the same field of phenomena is in fact by pursuing two different research programmes: Biot tried to develop a corpuscularist account of diffraction, Fresnel a wave account; the Ptolemaists tried to develop a geostatic account of observed planetary motions, Copernicus a heliostatic one; and so on. But of course no pair of scientists can possibly arrive at the *very same* theory in such ways (though they might very well, of course, arrive at two different but empirically equivalent theories). The specific theory that scientist A arrives at will of necessity entail the general, 'hard core' theory underlying her research programme, while the specific theory that scientist B arrives at will equally entail the general hard core theory underlying his different research programme—the hard cores of rival programmes are, by definition, inconsistent and so, therefore, are the two specific theories.

This 'two scientist' story, then, can only start to make sense if A and B are working within the same research programme. Again it is important to realise that there is no significant subjective element here: a research programme either supplies a theoretical reason for parameter to have a particular value or it does not. It is, for instance, just a fact that the wave optics programme, to take again my favourite example, supplies no general theoretical reason to fix the parameter corresponding to the wavelength of light from a sodium arc at any particular value (at least within a wide range). A more extensive (but in the 19th century, of course, unavailable) theory involving that wave theory but also an account of the radiation of light from particular sources with particular chemical constitutions, and subject to particular inputs of energy, might conceivably have done so, but the wave theory of light itself—objectively—just does not. Hence no 19th century scientist could see a theoretical justification for taking some particular value of the parameter, and all such scientists needed instead to use the results of experiments to fix the value (or take a blind guess—see *below*). Such a scientist could not have seen a theoretical justification for a particular value of that parameter, because there was no theoretical justification to be seen.

If both scientist A and scientist B work systematically in such a case, then both would need to use data in order to arrive at their more specific theory—it couldn't be that A, say, used data but B purely theoretical considerations in arriving at the same theory, since there are no such theoretical considerations to be considered. The only way that the two-scientist story could get going in such a case would be if one of them, A say, made a blind guess at the value of the parameter left free by theoretical considerations and yet happened, by simple good fortune, to hit on the very same value that B arrived at systematically by using data e. Each starts from the same general theory T, each arrives at the same more specific theory T′, though by different routes. Is not the 'heuristic' approach then forced into the absurdity that T′ as arrived at by systematic scientist B fails to be supported in the stronger sense by the evidence e, while that very same theory as arrived at by unbelievably lucky A *is* supported in that stronger sense by e?

This, admittedly wild, possibility is one that used to exercise Peter Urbach.[23] Once it is realised, however, that we are not appraising scientists but rather theories-in-the-context-of-research-programmes, then any apparent awkwardness here evaporates. The stronger sort of confirmation that I have highlighted is the sort that spills over from the specific theory that entails the relevant data to the underlying general theory or programme. The chief practical impact of such confirmation is to supply confidence in the successful extendability of that same general idea to a different (sub-)field (which will of course mean constructing a different specific theory T′′). Clearly lucky A in the above case has not shown anything relevant in this regard. She has *not* shown that the underlying general idea deserves this sort of support from e, since she has not shown that there are theoretical considerations attached to that general idea tying down the relevant parameter to the particular value, that she merely (and with quite incredible good luck) conjectured. The correct judgement is surely the one supplied by my dual account: (i) that B has shown, while A has not, that T′ is maximally confirmed by e in the conditional sense: B has shown that, since e entails T′, modulo T, if you accept the general theory T you must accept the more specific theory T′; while (ii) A has *not* shown that e supplies 'unconditional' 'stronger' support for T′ in the sense that would spill over to the underlying T. Of course, if it turns out that T′ also yields further so far unconsidered (though actual) data e′, then e′ (unlike e) does provide this stronger unconditional sort of support and it supplies it for T′ as proposed by either A or B. Of course it does: A and B have proposed the same theory!

So far we have considered the case where the underlying research programme gives the theorist no reason why a particular parameter should have a particular value and hence she needs, if she is to work systematically at all, to invoke data. Suppose now, to the contrary, that there *is* a theoretical justification, provided by the research programme concerned, why some parameter λ should have a particular value, but scientist A fails to see that reason. A instead uses evidence e to tie down

[23] See his (1978).

λ, at, say, the value λ_0; thus producing a theory, T(λ_0), that, in turn, entails some further, initially unconsidered, evidence e′. Scientist B, on the other hand, sees that her research programme already supplies a theoretical reason why λ should have the value λ_0 and goes *directly* to T(λ_0), pointing out that it entails *both* e *and* e′. I cannot see any reason why this scenario couldn't be realised, though I am doubtful that there are any real historical examples. However, if there are such examples, what we ought to say about them again seems entirely straightforward: *not* that the theory as proposed by A is supported (in the stronger, general theory or research programme supporting sense) only by e′, while the very same theory as proposed by B is supported (again in that stronger sense) by both e and e′; instead we would say that scientist B has *shown*, what A simply subjectively failed to recognise, that the theory is supported in this stronger sense by both e and e′. Once it has been realised what the different support judgments I have highlighted are doing—giving merely conditional support against the background of a presupposed general theory or, more interestingly, giving support to that more general theory itself (though *via* a specific representative of it)—then any apparent mystery in this sort of case too disappears.

Could there be, objectively speaking and laying aside random guessing, more than one route within a research programme to the same theory? And would the existence of such multiple routes pose any threat to the theory of support that I have outlined?

I can think of only one such way. And this is an entirely benign case that has already in fact been mentioned. Quite often with powerful scientific theories (as, for example, in the simple wave-theoretic case I sketched above involving the determination of the wavelength of monochromatic light from a particular source) experiments *overdetermine* the value of that parameter in the following sense. The general theory, in this example the general wave theory in which the values of all wavelengths of monochromatic light-sources are free parameters, entails not just one, but a range of formulas, involving the wavelength and measurable quantities *in different experiments*. So for example, alongside the equation cited above linking the wavelength to measurable fringe-distances in the two-slit experiment, the general theory entails another equation linking that wavelength to measurable fringe-distances in the one-slit diffraction experiment. This *does*, then, admit a genuine scientist A/scientist B scenario: A might produce T′ out of T in the way outlined earlier, using the result of the two slit experiment with monochromatic light from a sodium arc to fix the value of the wavelength, and then use T′ to predict the exact outcome of the one slit experiment with light from the same source in quantitative, rather than merely qualitative terms; while Scientist B on the other hand might produce what turns out to be the very same theory T′ on the basis of the result of the one-slit experiment with light from a sodium arc and then use it to predict the quantitative details of the two-slit experiment. (Of course the fact that it turns to be the very same theory is a contingent fact reflecting the predictive power of the wave theory. Scientist A using the two-slit data might have produced T′, while B's use of the one-slit data led to the different T′′ (in fact inconsistent with T′). This would mean that A's theory failed the one slit diffraction test, while B's failed the two-slit test.)

Is this really a problem for the dual account of confirmation I have sketched? Let's call the two-slit fringe data with light from a sodium arc e_1 and the one-slit fringe data using the same light-source e_2. Telling it from the point of view of scientist A, e_1 confirms T′ in the conditional sense (it entails T′ given the general theory T), while e_2 confirms T′ in the stronger sense that spills over to T; from the point of view of scientist B, on the other hand, the roles of e_1 and e_2 are reversed: e_2 confirms T′ conditionally, while e_1 supplies the stronger T-involving confirmation. These may be strictly different accounts but they are surely equivalent *modulo* any genuine interest that we would have in making confirmation judgments: each of A and B has shown that the general theory needs to fill in one parameter value on the basis of one piece of data, thus producing a specific theory that gains genuine empirical success from the other piece of data (at least—there may of course be other results that specific theory also correctly predicts). So each scientist shows that there is, so to speak, one unit of genuine, unconditional, general-theory-involving data and hence delivers the judgment that that general theory is ahead in terms of empirical support of any theory (such as the rival emissionist theory in the early 19th century) that *merely* accommodates both pieces of data. (On the other hand if, as was not the case historically, there were still a third theory which, without needing either e_1 or e_2 to fix parameters, entailed both of them 'naturally'—in the way that Copernican theory entails planetary stations and retrogressions—then that third theory would be even better confirmed. Intuitively we would want to say that the score—relative of course to just e_1 and e_2 (judgments might be different in view of the *total* evidence)—would be 'Imaginary theory 2, wave theory 1, emission theory 0'; and this is exactly the score that is delivered by my dual account of confirmation.)

In sum, then, objection two fails: contrary to Alan Musgrave's claim, confirmation is not unacceptably inquirer-relative on the approach that I endorse. It is clearly a desideratum on any account of confirmation that it underwrite the judgement 'same evidence, same theory, same confirmation' and my account underwrites exactly this judgment. (This is in contrast, of course, to the judgement 'same evidence, two *rival but empirically equivalent* theories, same confirmation', on the denial of which this whole approach is based.)

Is the 'heuristic' account, when properly understood, a version of the historical approach?

According to Musgrave's classification, all the accounts that make confirmation dependent, not only on theory and evidence, but also on background knowledge, for that reason make confirmation (at least partly) 'historical'. This is because background knowledge may change over time and so the answer to the (in fact elliptical) question 'does evidence e support theory T?' may be different in different eras—eras that are characterised by different states of background knowledge.

When properly understood, however, the 'heuristic' view I advocate does not have this historical character. It does, certainly, make confirmation a three-, rather than two-place relation. But, although describable in a loose way as making confirmation dependent on background knowledge, in fact this account makes confirmation (or rather *both kinds of confirmation*) depend on evidence e, specific theory T′, and the underlying general theory T. It is not a history-dependent, but rather a research programme-dependent account:

(1) Evidence e confirms$_1$ T′ in the context of the general underlying theory T if the conjunction of e and T entails T′ (or more generally to the extent that e and T entails T′);

while:

(2) Evidence e confirms$_2$ T′ in the context of the general underlying theory T if (i) T′ entails e, and (ii) T′ has been developed out of T in a way independent of e.

There is no question, then, of historical variability in either of the types of confirmation-judgment. Fresnel's wave theory of diffraction was, is, and forever shall be, confirmed (confirmed$_2$) by the 'white spot' result—this result follows from that wave theory of diffraction and gives support to the whole wave programme. Fresnel's specific claim that light waves are transverse rather than longitudinal is, was, and forever shall be, confirmed (confirmed$_1$) by the disappearance of the fringes in the two slit experiment when the two beams are oppositely polarised—this result did, does and ever more shall entail the specific transverse claim given the general idea that light is a wave in an elastic medium.

Of course the historical context changes, because other theories are articulated. Hence the question of whether only one theory of light is confirmed$_1$ by the white spot result may (and indeed of course did) have one answer in 1819 ('yes, Fresnel's') and another answer ('no') in the 1860s, once Maxwell's electromagnetic theory had been formulated. Hence the issue of whether some result provides grounds for accepting a theory as the currently best available in its field quite properly, and obviously, has an historical dimension. But, as I argued earlier in considering Alan Musgrave's own preferred version of the historical approach, it would surely be a mistake to confuse these patently historical issues with the ahistorical one of whether some theory is confirmed by some piece of evidence. The main conclusion of this paper is that there are two types of confirmation—both of them (three-place) 'logical'. These confirmation judgments then feed into the clearly historical issue of which currently available theory is *best* confirmed by the currently available evidence.

8. DEBORAH MAYO AND CONFIRMATION VIA 'SEVERE TESTS'

Finally, I want to try to work out an issue that has troubled me for some time—namely, the relationship between my account and the much-discussed views of Deborah Mayo. She—again taking 'Musgrave's neat analysis of the situation' (*op.cit*, p. 255) as one of her starting points—has developed a theory of confirmation that, amongst other things, claims to account both for the cases in which the heuristic account of confirmation accords with our intuitive judgments about particular cases of confirmation and for those cases in which the heuristic account conflicts with those judgments. Mayo in effect takes the heuristic account to be captured by the 'use novelty' or 'no double use' slogan: 'You can't use the same fact twice, once in its construction and then again in its support.' The view that I have been developing here, as we saw, also rejects this slogan in its straightforward interpretation: claiming instead that you can indeed both use a datum in the construction of a theory and use it to support the constructed theory in the sense of support that is conditional on pre-acceptance of the underlying general theory, but that used data cannot support in the stronger, unconditional sense that spreads from the specific theory that entails the data to general theory underlying that specific theory. Intuitive judgments that were in conflict with the 'no double use rule' are in fact judgments of conditional support; intuitive judgments that conform to that rule are judgments of the stronger, unconditional kind of support.

Both Mayo and I, then, claim to capture the intuitively underwritten judgments in all particular cases of confirmation or support—both those that have been cited in favour of the 'no double use' rule and those that have been cited as refutations of that rule. What, then, is the relationship between Mayo's account and my own: which is better, or are they perhaps just two different ways of saying the same thing?

Mayo's basic line of reasoning is very simple. Hypotheses should gain empirical credit only from passing genuine tests; and the more severe the test, the higher the confirmation or support, if the theory passes it. The defenders of the use-novelty account hold that evidence used in the construction of a hypothesis cannot provide a genuine test of it and hence cannot supply genuine confirmation. Underlying their view, on Mayo's analysis, is the claim that a severe test is one that a theory has a high probability of failing; and hence, since a theory constructed with the help of evidence has no chance of failing the 'test' supplied by e, that the use-novelty view is correct. However plausible this may sound, argues Mayo, it in fact misidentifies the probability that we should be concerned to maximise: a non-severe test is *not* one that has a high probability of being passed by a theory, but rather one that has a high probability of being passed by the theory, *even though the theory is false*. As she puts it 'what matters is not whether passing is assured but whether erroneous passing is' (*op. cit*., pp. 274-5).

In cases where the heuristic view as originally formulated goes wrong—such as her SAT score 'hypothesis' and standard statistical estimation (where we use, for example, the evidence of the observed relative frequency of white balls in a sample to arrive at a theory about the unknown frequency in the urn)—there is indeed no chance of the (constructed) hypothesis failing the 'test'. However, the chance of the hypothesis having passed the 'test' *if it were false*, is zero (in the case of the SAT score example) or very small (in the confidence-interval case). Concentrating on the more straightforward SAT score case, the 'test' of the 'hypothesis' that the average SAT score of her logic students is 1121—the "test" consisting of taking the individual SAT scores and dividing by the number, n, of students (thus producing evidence e)—is in fact *maximally severe*, according to Mayo: 'since there is no way that such a result can lead to passing H erroneously, H passes a maximally severe test with e.' (*op.cit*., p. 271)

Part of my response will involve justifying the scare quotes I have placed around 'hypothesis' and especially around 'test' in outlining her view; this will in turn lead to the criticism (which Deborah Mayo herself cites and tries—unsuccessfully—to meet) that the SAT score and statistical estimation cases are not representative of the interesting cases from science. While there is no doubt that Mayo's account and my own are based on a number of shared views and intuitions, my account gets the situation straight whereas her own is somewhat skewed.

As already remarked, it does seem extraordinary to call the assertion arrived at about the average SAT score of Mayo's students an 'hypothesis', and at least equally extraordinary to call the process of adding the individual scores and dividing by the number of students a 'test' of that claim. Of course had someone made a 'bold conjecture' about the average score, then one might talk of the systematic process of working out the real average as a test of that conjecture. But boldly conjecturing would clearly be a silly way to proceed in this case, and, as already remarked, not one that would ever be used in science. As it is, the process of adding the individual scores and dividing by the number of students surely is a *demonstration that* the average score is 1121, not a *'test' of* the 'hypothesis' that this is the average score.

This also points to a real problem in applying Mayo's central justification for all confirmation judgments to this particular case. In the circumstances (and assuming that both the data on the individual students and the arithmetic have been carefully checked) there is *no* chance that the average SAT score is *not* 1121. So we are being asked to make sense of a conditional probability—the probability that the claim about the average score would have passed the test, had it been false—where the conditioning event (the claim's being false) has probability zero; and indeed asked not only to make sense of it but to agree that the conditional probability at issue is itself zero. It is well known, however, that—at any rate in all standard systems—$p(A/B)$ is not defined when $p(B) = 0$. It is true that Mayo wants us to concentrate primarily on intuitive judgments about 'probability' and not on what can formally be justified as genuine probabilities. However I confess that I have no idea what it means in this case, even 'intuitively', to imagine that the average score is *not* 1121, when the individual scores have been added and divided by n and the result *is* 1121!

There is not the same formal difficulty of course in the statistical estimation case, where we can readily make sense of the probability that the estimate is wrong (that is, that the interval systematically arrived at on the basis of the sample data does not in fact include the real population value of interest). However it is intuitively quite wrong to talk of 'tests' in this case too. In the deterministic case, we *measure* a parameter (or demonstrate that that parameter has a certain value); in the stochastic case we *estimate* a parameter. Although apostate, I remain enough of a Popperian to put very little weight, in general, on how we happen to talk, but there seems to be in this particular case a very good reason why we do not talk of tests: despite Mayo's claims, a test of a theory surely must have a possible outcome that is inconsistent with the theory— neither the SAT score process nor the confidence-interval technique could possibly refute the 'theory' that we end up with.

As should be clear from my own positive account, I am far from disputing Deborah Mayo's claims that both measurements and estimations of the value of parameters form important aspects of scientific reasoning. I also agree with her in particular that statistical inference from actual experimental data to the claim that we normally regard as 'the' (generalised) result of the experiment is an important, and relatively underexplored, aspect of the logic of science. However these are *not* aspects of any testing-process in science. The lesson to learn, contrary to Mayo's general view, is that science is not all about tests of theories and so not all about attempts to detect error; some of the important logical relationships between evidence and theory are of a quite different nature. Mayo gets herself into trouble by attempting to produce a 'one size fits all' account—all (let's say) *accreditations* of theory by data are, she claims, the results of tests, once tests are properly construed (that is, construed in line with her account, of course).

The problems that this approach leads to are made still clearer when we consider cases that are more representative of reasoning in science. The feature of the SAT score case that makes it unrepresentative, as already indicated, is that there just is no genuine theory around—the framework is simply given, in particular the relationship between the individual scores and the average score is analytic. In the statistical estimation case, there *is* an underlying 'model'—when drawing balls from the urn, for example, we are assuming that it is Bernouilli process with underlying fixed population parameter p (the fixed proportion of white balls in the urn), but this underlying model is not itself usually thought of as at all conjectural. (We can't see inside the urn, so it *might be* that some demon is constantly changing the proportion of white and black balls—but we just assume that this is not the case and don't look for any experimental confirmation of this assumption.)

The interesting scientific cases of 'deduction from the phenomena', as indicated earlier, on the contrary, all involve a general underlying *theory*. In the simplest case, a specific value of a parameter is deduced from the data, but only *given* an underlying general theory that yields, without any experimental input, some functional relationship between the free version of that parameter and experimental results. This underlying theory, although it may be assumed as 'given' for the then current purposes, is itself clearly a substantive and defeasible assumption and as such stands in need of confirmation from evidence no less than the specific theory deduced from it and the phenomena. The (general) wave theory of light replaced the (general) emission theory and was itself then replaced by the electromagnetic, and later photon theory. The general theory's fortunes are subject to the changing verdict of ever accumulating evidence—we need to take its defeasibility into account.

Consider the case, analysed in detail above, where a specific version of the wave theory T′ with a definite value for the wavelength of light from a sodium arc is deduced from the general wave theory T using evidence e from slit- and fringe-distances in the two slit experiment with light from that source. My account entails that e does confirm T′ in this case—strongly, but in the conditional sense. What does Deborah Mayo's account say? In line with her claims about the SAT score case she too will want to say that e gives some good degree of confirmation to T′ (these are exactly the sorts of *real* scientific cases where the unmodified 'no double-use' rule goes wrong). She will be forced to say that this is because e constitutes a pass for T′ in a test that had relatively little chance of passing it if it were false. However, it is surely clear that this 'test' in fact had every chance of passing T′, *whether it is true or not*. Whatever general theory of light is true, that is, whether or not T is true (and T′ can't be true if T is not) T′ was indeed *bound to* get the fringe distances in the two-slit experiment with sodium light correct—exactly because T′ was fitted to e!

The correct judgment is surely the one delivered by my account: e constitutes no test of T′ but it does tell us something positive about it—namely, that it is *the* specific representative of T (so far of course as this particular detail is concerned). If T′ is not correct (and we have, to repeat, no chance of finding that out from the two slit experimental result though we might, as explained, from other experiments), then neither can the general theory T be correct. These further experimental results (such as the prediction of the outcome of the one-slit experiment with light from the same source) *are,* on the contrary, genuine tests. Genuine tests produce the stronger, non-relativised type of confirmation my account talks about (and it is here that Mayo's 'error probability' intuitions —about it being improbable that the theory should get such a test result right if it is, not false, but something weaker like 'structurally off-beam'—come genuinely into play). Parameter-fixing exercises and other such inferences, on the contrary, are important (indeed crucial) aspects of the scientific endeavour and they carry important information about theories, but they are not *tests* of those theories.

There are hints of possible responses to this criticism in Mayo's book (especially when she discusses the problems posed by the possibility of alternative theories, really alternative theoretical frameworks). But even without going into details it seems clear that her approach is in general barking up the wrong tree. We have here two *quite different* roles for evidence *vis à vis* theory, just as my approach implies. This is exactly why my approach yields two quite different notions: confirmation$_1$ and confirmation$_2$. They are not, as Mayo is trying to make them out to be, simply two different aspects of the one drive—to test theories in, and hence to eliminate error from, science.

REFERENCES

French, A. (1971) *Newtonian Mechanics*. Cambridge, Mass: M.I.T. Press.

Fresnel, A.J. (1819) 'Mémoire sur la diffraction.' reprinted in E.Verdet et al. (eds) (1865).

Hitchcock C. and Sober, E. (2004) 'Prediction Versus Accommodation and the Risk of Overfitting.' *British Journal for the Philosophy of Science* 55: 1-34.

Howson, C. (1990) 'Fitting Theory to the Facts: Probably Not Such a Bad Idea After All.' in C. Wade Savage (ed) *Scientific Theories*. Minneapolis: University of Minnesota Press.

Howson, C. and Urbach P. (1994) *Scientific Reasoning: the Bayesian Approach,* 2nd Edition. Chicago and La Salle: Open Court.

Kuhn, T.S. (1977) *The Essential Tension*. Chicago: University of Chicago Press.

Mayo, D. (1996) *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Musgrave, A. E. (1974) 'Logical versus Historical Theories of Confirmation.' *British Journal for the Philosophy of Science* 25: 1-23.

Nickles, T. (1987) 'Lakatosian Heuristics and Epistemic Support.' *British Journal for the Philosophy of Science* 38: 181-205.

Urbach, P. (1978) 'The Objective Promise of a Research Programme.' in Radnitzsky and G. Anderson (eds) *Progress and Rationality in Science*. Dordrecht: Reidel.

Verdet. E. & Senaramont H. (eds) (1865) *Oeuvres Complètes d'Augustin Fresnel.*

Worrall, J. (1985) 'Scientific Discovery and Theory-Confirmation.' in J.C. Pitt (ed): *Change and Progress in Modern Science*. Dordrecht: Kluwer.

Worrall, J. (1989) 'Fresnel, Poisson and the White Spot: the Role of Successful Prediction in Theory-Acceptance.' in D. Gooding et al (eds) *The Uses of Experiment—Studies of Experimentation in Natural Science*. Cambridge: Cambridge University Press.

Worrall, J. (2000a) 'Kuhn, Bayes and "Theory-Choice".' in R. Nola and H. Sankey (eds) *After Popper, Kuhn and Feyerabend*. Dordrecht: Kluwer.

Worrall, J. (2000b) 'The Scope, Limits and Distinctiveness of Newton's Method of "Deduction from the Phenomena".' *British Journal for the Philosophy of Science* 51: 45-80.

Worrall, J. (2002) 'New Evidence for Old' in P.Gardenførs et al (eds) *In the Scope of Logic, Methodology and Philosophy of Science*. Dordrecht: Kluwer.