JOHN WORRALL

KUHN, BAYES AND 'THEORY-CHOICE': HOW REVOLUTIONARY IS KUHN'S ACCOUNT OF THEORETICAL CHANGE?*

1. INTRODUCTION: KUHN AND THE HOLD-OUT

Book reviews supply a rich source of sharp, sardonic humour. Probably my favourite remark about reviewing was by the wonderfully droll Reverend Sidney Smith, who opined 'I never read a book before reviewing it – it prejudices a man so!' Another favourite – in similar (though strictly speaking contrary) vein – is from a psychologist friend of Wesley Salmon's, who, when asked for his opinion of the latest Dianetics tosh by Lafayette Ron Hubbard, apparently remarked 'I cannot condemn a book before reading it; but after reading it, I shall'. My favourite remark, though, from within a book review is probably: 'This book fills a much-needed gap in the literature.'

No one could, of course, seriously hold that this last remark applies to Thomas Kuhn's *The Structure of Scientific Revolutions* nor to his earlier wonderful book on *The Copernican Revolution*. Indeed the latter is an outstanding example of proper, largely internal history of science, while the former is one of the most influential and discussed, quoted and misquoted books of our time. But as for the whole secondary literature on what Kuhn did and did not *really* mean – a literature to which Kuhn himself contributed rather generously – I think that one could argue quite plausibly that *it* fills a much-needed gap.

Surely the sincerest tribute to an investigator is not endlessly and scholastically to interpret and reinterpret his or her writings, but rather to try to make progress towards solving the problems that he or she raised. At any rate, I shall try in this essay to arrive quickly at first-level concerns about the rationality of science – and especially the rationality of theory-change in science – using problems raised by Kuhn, and criticising claims that Kuhn seems to have made, without worrying too much about whether they express his 'real view', if indeed there is such a single unified entity.

The chief target of the critical fire from those who felt Kuhn challenged the whole idea of science as a rational process was always his apparent views about the process of paradigm *change*. (Lakatos, for example, notoriously claimed that Kuhn's views made theory-change in science a matter of 'mob psychology'.) Most of what his critics found objectionable in Kuhn's account of theory-change

Robert Nola and Howard Sankey (eds.), After Popper, Kuhn and Feyerabend, 125–151.

is reflected in his remarks about 'hold-outs' to 'scientific revolutions'. He claimed that if we look back at any case of a change in fundamental theory in science we shall always find eminent scientists who resisted the switch to the new 'paradigm' long after most of their colleagues shifted. These 'hold-outs' – Priestley defending phlogiston against Lavoisierian chemistry is a celebrated example – are often (though by no means invariably) elderly scientists who have made significant contributions to the entrenched paradigm. Kuhn added to this interesting but relatively uncontroversial descriptive claim the challenging *normative* assertion that these 'elderly hold-outs' were no less justified than their more mobile contemporaries: not only did they, as a matter of fact, stick to the older paradigm, they were also, if not exactly right, then at least not wrong to do so. On Kuhn's view, 'neither proof nor error is at issue' in these cases, there being 'always some good reasons for every possible choice' – that is, both for switching to the revolutionary new paradigm and for sticking to the old. Hence the hold-outs cannot, on his view, be condemned as 'illogical or unscientific'. But neither of course can those who switch to the new paradigm be so condemned. In one sense, then, it is easy to see why Kuhn expressed mystification over the claim that he made the history of science an irrational affair: in Kuhn's cosy world, *everyone* is rational – revolutionary and reactionary alike. But a genuinely 'rationalist' account surely needs losers as well as winners: rationalists seek general rules of theory-appraisal which presumably will show that the hold-outs were, in some important sense, simply mistaken.

Discussions of this issue are likely to become overly-abstract and the significant questions missed unless real historical examples are investigated in some detail. In the next section, therefore, I outline the views of one hold-out (a not so elderly one as a matter of fact). This is a case I have discussed elsewhere,[1] so I shall be very brief. Having resketched the historical details, I extend – and I believe, improve on – my earlier attempt to use those details to illustrate some general methodological morals. In particular I shall draw on the case-study to provide what I believe is a much improved account of the relationship between Kuhn's views and those of contemporary personalist Bayesians. This improvement, which is partly inspired by a paper of John Earman's (Earman 1993), involves looking again at the issue of how far, and in which respects, Kuhn's views can be reconciled with personalist Bayesianism, and in particular investigating one point where the two positions seem radically at odds. Roughly speaking, I shall argue that Kuhn's account is inadequate *both* where it agrees with the Bayesians, *and* where it disagrees with them.

2. BREWSTER AND THE WAVE THEORY

The early nineteenth century 'revolution' in optics saw Fresnel's classical wave theory of light triumph over the material corpuscular theory of light, generally attributed to Newton. Although this episode's impact on man's whole worldview cannot match that of, say, the Copernican or Darwinian revolutions, it did involve a sharp change in accepted theory and is explicitly cited by Kuhn as counting as a revolution in his terms. Indeed, partly because it is a narrowly scientific affair, this particular theory-change provides, I believe, an especially clear-cut instance against which to test general methodological claims. The most significant hold-out to this revolution, from Britain at least, was Sir David Brewster.

Although perhaps chiefly remembered nowadays for his biography of the great Sir Isaac, Brewster was an important optical scientist in his own right. He was the discoverer of many of the properties of polarised light; he discovered 'Brewster's law' relating the polarising angle and refractive index of transparent substances; he discovered a whole new class of doubly refracting crystals – the biaxial crystals – which soon proved to have great theoretical significance; he discovered that ordinary unirefringent transparent media can be made birefringent by the application of mechanical pressure; and he discovered the hitherto unknown general phenomenon of selective absorption.

Brewster was certainly some sort of hold-out. In 1831 a fellow knight of the realm, Sir George Biddel Airy, published a *Mathematical Tract on the Undulatory Theory of Light* which begins:

> The Undulatory Theory of Optics is presented to the reader as having the same claims to his attention as the Theory of Gravitation, namely that it is certainly true .... (Airy 1831, p. vii)

This would, I think, have been regarded at the time as a rather extreme expression of what was, however, definitely the majority view among the (admittedly small) group of those qualified in optics. Certainly the great majority of that group felt that the corpuscular approach had been definitively superseded by Fresnel's ether-based approach. Brewster held some form of minority view. In the same year that Airy published his *Mathematical Tract*, Brewster presented a 'Report on the Present State of Physical Optics' to the British Association for the Advancement of Science in which he asserted that the undulatory theory was 'still burthened with difficulties and [so] cannot claim our implicit assent' (1883a, p. 318). Two years later he reported:

> I have not yet ventured to kneel at the new shrine [that is, the shrine of the wave theory] and I must acknowledge myself subject to the national weakness which urges me to venerate, and even to support the falling temple in which Newton once worshipped. (1833b, p. 361)

This rhetorical flourish notwithstanding, Brewster was no mere irrational, 'Newton-worshipping' reactionary. He produced some sensible and challenging arguments for the *ancien regime*. There are in fact, as I see it, three main elements in Brewster's views about the then current state of play between the wave and emissionist theories.

(*i*) *Brewster accepted – fairly unambiguously – that as things stood the wave theory had proved to be empirically the more successful.*

He frequently expressed great admiration for the wave theory and fully acknowledged that it had enjoyed outstanding explanatory, and especially *predictive* success. For example, he said:

> I have long been an admirer of the *singular* power of this theory to explain some of the most perplexing phenomena of optics; and the recent discoveries of Professor Airy, Mr. Hamilton and Mr. Lloyd afford the finest examples of its influence in predicting new phenomena. (1833b, p. 360; emphasis supplied)

(Here Brewster has primarily in mind the prediction drawn by Hamilton from Fresnel's theory of the existence of both internal and external conical refraction. Since they involve directing a narrow ray of light along a very precisely characterised path through crystals of a very particular sort, cut in a very particular way, these represent exactly the sort of phenomenon that could realistically only be discovered as the result of testing some precise predictions of some powerful theory. Humphrey Lloyd confirmed Hamilton's predictions experimentally in 1833.)

(*ii*) *Brewster believed that the wave theory – for 'all its power and all its beauty' – could not be true.*

He produced two main arguments for this belief. The first was of a general methodological kind, related to the recently fashionable thesis of underdetermination of theory by evidence. Brewster pointed out that the fact that a theory had enjoyed explanatory, and even predictive, success does not of course deductively entail that it is true. Instead:

> Twenty theories may all enjoy the merit of accounting for a certain class of facts, provided they have all contrived to interweave some common principle to which these facts are actually related. (1833b, p. 360)

He did allow that the wave theory's predictive success implies that 'it must contain among its assumptions some principle which is inherent in ... the real producing cause of the phenomena of light' (1838, p. 306). However, *first* other theories – notably the Newtonian one – might well be able to incorporate such a principle; and *secondly* there was no doubt in Brewster's mind that, despite itself incorporating such an assumption, the wave theory, considered as a fully realistically interpreted claim about the universe, had to be false. In particular, a fully realistically interpreted wave theory was committed to the existence of what Brewster himself described as 'an ether, invisible, intangible, imponderable, inseparable from all bodies and extending from our own eyes to the remotest verge of the starry heavens' (*ibid.*). This was always too much – or perhaps *too little* – for Brewster to swallow. So Brewster's view was that a fully acceptable theory would share many of the structural assumptions implicit in the current wave theory but would reject – at least – that theory's invocation of the luminiferous ether. And he had, moreover, not abandoned the hope that some (highly modified) version of the Newtonian corpuscular account might prove to be such a theory.

Alongside this general argument, Brewster produced a second argument relating to the details of the particular version of the wave theory then current. Brewster in effect pointed out that while the wave theory's predictive success might be impressive, it was by no means complete. The wave theory failed and failed badly, in the case of at least two empirical phenomena: those of dispersion and of selective absorption.

I concentrate here on the second, a phenomenon which Brewster actually discovered, though identical morals could be drawn from the first failure. Brewster found that if a beam of sunlight is passed through certain gases and then dispersed in a prism, the spectrum that emerges from the prism is marked by a series of dark lines – indicating that, speaking in wave-theoretical terms, the components of the sunlight of certain sharply defined wavelengths have been absorbed during passage through the gas. Brewster – entirely reasonably – pointed out that, rather than simply refuting some particular version of the wave theory, this phenomenon provided a general difficulty for the whole wave approach. Whatever the details, the general story the wave theory seemed forced to tell looked extremely far-fetched. Referring to one particular absorption line (in 'oxalate of chromium and potash'), the wave theory needed to claim that the ether within that gas 'freely undulates to a red ray whose index of refraction in flint glass is 1.6272, and also to another red ray whose index is 1.6274 while ... its ether will not undulate at all to a red ray of intermediate refrangibility whose index is 1.6273!'

In other words, an infinitesimal change in the length of a wave must be supposed to produce a discrete change from free passage through the ether within the gas to no passage at all. Brewster pointed out that:

> There is no fact analogous to this in the phenomenon of sound, and I can form no conception of a simple elastic medium so modified by the particles of the body which contains it, as to make such an extraordinary selection of the undulations which it stops or transmits .... (1833a, p. 321)[2]

(*iii*) *Brewster disagreed with the wave theorists over the heuristic issue of the likeliest way forward.*

The defenders of the wave theory in Britain, notably Airy and Baden Powell, had no problems in accepting Brewster's claims – so long as they were understood as simply about the *present state* of the wave theory. Each acknowledged (they could scarcely do otherwise) that, as it stood in the 1830s, the wave theory had, for example, no explanation for selective absorption. However they each went on to point out that the wave theory had earlier had no explanation for polarisation either and that in particular it had seemed to be refuted by Fresnel's and Arago's experiments on the interference of polarised beams. (If the famous two slit experiment is modified so that the light coming through the two slits is polarised in mutually orthogonal planes – by the interposition of suitably oriented quartz plates, for example – then the interference fringes visible in the original experiment disappear.) But, rather than give up the theory, Fresnel had taken the bold step of instead developing it – in fact by switching from the assumption that light waves are longitudinal to the assumption that they are transverse (and hence switching from the idea that the ether is an elastic fluid to the idea that it is an elastic *solid*). This step, they pointed out, had led to exactly the sort of predictive success that Brewster himself applauded. In particular, Fresnel's move had led to the prediction, mentioned earlier, of the hitherto unsuspected phenomena of internal and external conical refraction. Moreover, as Airy and Powell justly asserted, there was nothing to match this success in the whole track-record of the emission theory.

While Brewster seemed to have some relatively vague belief in the revivability of the emissionist/corpuscularian approach, the wave theory had recently, and more than once, shown its ability to change major problems into major predictive successes. In that situation, Baden Powell argued:

> No sound philosopher would for a moment think of abandoning so hopeful a track, and none but the most ignorant or perverse would find in the obstacles which beset the wave theory anything but the most powerful stimulus to pursue it. (1841, p. iii)

3. KUHN'S LATER ACCOUNT OF 'THEORY-CHOICE': 'OBJECTIVE' AND 'SUBJECTIVE' FACTORS

The idea of a scientist 'holding out' against a new theory seems hopelessly vague. We now have an altogether more detailed account of a hold-out-scientist's position. This will eventually enable us to ask more penetrating questions about Kuhn's general views on reason and theory-change. But first we need to have a clearer picture of those general views themselves.

Kuhn developed – initially in the 1970 *Postscript* to his *The Structure of Scientific Revolutions* and then, in rather more detail in chapter 13 of his (1977) book *The Essential Tension* – a fuller account of the factors underlying what he there called 'theory-choice', than anything found in the original book. The 'mob psychology' gibe, he argued in 1977, 'manifests total misunderstanding' because he had always allowed a crucial role to the 'objective factors' from the philosopher's 'traditional list' (and he mentions five such objective factors: empirical accuracy and scope, consistency, simplicity and 'fruitfulness'). Kuhn says:

> I agree entirely with the traditional view that [these objective factors] play a vital role when scientists must choose between an established theory and an upstart competitor . . . [T]hey provide the *shared* basis for theory choice. (1977, p. 322)

His claim had simply been all along that, while important, these objective factors fail to supply an 'algorithm for theory-choice'. At any rate when the choice between rival theories is a live issue in science, the objective factors never *dictate* a choice. Amid a good deal of rather flabby talk about methodological rules operating as *values* that 'influence' choice rather than as *rules* that dictate it, Kuhn supplies two sharp reasons for this failure of objective factors to provide an 'algorithm' for 'theory-choice'. (The reason for the 'scare quotes' round 'theory-choice', which from henceforth will be taken as implicit, will be explained below.)

The *first* reason is that single objective factors often turn out to deliver no unambiguous preference when applied to the theories *as they stood at the time when the choice was being made* – 'Individually the criteria are imprecise: individuals may legitimately differ about their application to concrete cases' (1977, p. 322).

For example, it is often assumed that the Copernican heliostatic theory was empirically more accurate than the Ptolemaic theory. This *eventually* became true but only as a result of the work of Copernicus, Kepler, Galileo and others – who had clearly then 'chosen' the Copernican theory for other reasons (if, indeed, for any *reasons* at all).

The *second* source of the failure of the objective factors generally to deliver a definite choice of theory is that 'when deployed together, they repeatedly prove to conflict with one another'. That is, even where single objective factors do point clearly in the direction of one of the rival theories, different factors may – again at the time when the choice was actually being made – point in *opposite* directions: so, for example, while simplicity (in a certain sense) favoured Copernican theory, consistency (with other, then accepted, theories) undoubtedly favoured the Ptolemaic theory. Kuhn concluded that the objective factors, while supplying the shared criteria of choice,

> are not themselves sufficient to determine the decisions of individual scientists. For that purpose, the shared canons must be fleshed out in ways that differ from one individual to another. (1977, p. 325)

In other words,

> every individual choice between competing theories depends on a mixture of objective and subjective factors, or of shared and individual criteria. (1977, p. 325)

The intent of Kuhn's further explanation of his views was to show that, on the topic of theory-choice, they differed less from philosophical orthodoxy than had generally been believed: his critics' remarks about irrationality 'manifest total misunderstanding', and indeed he had chosen not to write on this topic earlier precisely because his real views on it diverge rather little from 'those currently received', as compared to other topics (1977, p. 321). Notice, however, that this more elaborate account is presented as explaining, and *endorsing* the earlier account in *Structure* and in particular as endorsing the claims about hold-outs. He also explicitly re-emphasised the specific entailment that, whenever a new theory is developed to challenge an older one, 'there are always at least some good reasons for each possible choice' (1977, p. 328) – that is, good reasons for sticking to the older theory as well as good reasons for switching to the new.

Much could be said about Kuhn's treatment of each of the 'objective' (or shared) factors, but the main point at issue in the present paper will be his general account of the distinction between, and necessity for, both 'objective' and 'subjective' factors.

4. KUHN'S ACCOUNT AND PERSONALIST BAYESIANISM

Kuhn himself seems, then, to have given here a direct answer to the question raised in my title: his account of theory-change in science, far from being revolutionary, is in close agreement with that given by 'the' philosophers of science. Several important issues can be clarified by pursuing this claim.

The first point to be made is of course that Kuhn's view of 'the' philosophers of science seems unjustifiably monolithic – it is difficult to think of a single issue on which philosophers of science speak with a single voice and certainly the issue of theory-change is not one of them. Much of Kuhn's account can indeed be interpreted as cohering quite well with one well-supported tradition within current philosophy of science – that of personalist Bayesianism. But that tradition, of course, as well as invoking fervent support, also invokes fierce resistance.

Bayesianism, as is well-known, makes the rationality of an 'agent' depend on two requirements, and – at least in the *pure* version (which, I would argue, is the only clear version so far articulated) – *only* two requirements. The first is that, at any given stage in the development of science, the agent distribute degrees of belief over the various statements available to her in such a way as to satisfy the probability calculus; and the second requirement is that, whenever new evidence e comes in and nothing else of epistemic significance occurs (that is, the agent's 'background knowledge', relative to which all probabilities are implicitly relativised, remains otherwise constant), then the agent's new degrees of belief be related to the old by the 'principle of conditionalisation'. This principle requires that the agent's new 'prior' degree of belief in any assertion A be, in those circumstances, her old degree of belief in A, *conditional on* the evidence e.[3]

The 'posterior probability' $p(A/e)$ is, *via* Bayes' theorem, dependent on, amongst other things, the prior probability $p(A)$. Personalist Bayesians think of the prior as measuring a purely subjective degree of belief. It is true that the agent may, and generally will, have arrived at those priors themselves by conditionalisation on *earlier* evidence – but that conditionalisation will itself have been dependent on an earlier prior and so on: the subjective element is ineliminable.

Given the subjectivity of the priors, then, as I suggested in my (1990), there seems to be no real problem in reconciling at least some aspects of Kuhn's account of 'theory-choice' with this particular Bayesian philosophy. (A similar point is made by Wesley Salmon in his (1990) article 'Tom Kuhn meets Tom Bayes'.) The 'subjective factors' are taken care of by the priors – the fact that two equally 'reasonable' investigators might disagree about the merits of two rival theories just means that they assign different priors to the same theory. Whether or not all the objective factors can be delivered as consequences of Bayes' theorem is another matter – but some have argued, with varying degrees of plausibility, that they can. If so, then, as Salmon pointed out, although Kuhn took himself to be denying the existence of an algorithm for ranking theories, he could be seen, at least in part, as *endorsing* an algorithm – namely Bayesian conditionalisation, while at the same time acknowledging, as personalist Bayesians anyway do, that subjective preferences need to be taken as inputs into this algorithm in order to produce a definite theory-choice. The fact (if it is one) that Airy and Brewster made different choices between the wave and corpuscular theories of light, even given all the evidence equally available to both, may be explained by the fact that they in effect assigned different *prior* probabilities to those theories ahead of the evidence.

But accepting that Kuhn's account is broadly coherent with that given by personalist Bayesians at best amounts, in the eyes of many philosophers (myself included), simply to a restatement of 'the problem with Kuhn': that he makes theory-change in a science a much more subjective affair than many of us believe it to be. (Remember that the outcome of Kuhn's analysis was still that there are always 'at least some good reasons for every choice'.) Conversely, to the extent that Kuhn's account can be reconciled with personalist Bayesianism, this simply underlines what critics of *Bayesianism* have always insisted on: that it allows much too large a role to subjective, personal factors to provide – without further augmentation – an adequate account of reasoning about theories in science.

Kuhn's claim that his views diverge relatively little from 'those currently received' is only in fact true of a proper subset of them; and is true of that proper subset because of a feature of their philosophy that many other philosophers think makes it ultimately indefensible.

The objections of the 'objectivists' stem first and foremost, of course, from the outright subjectivism of the priors. There are some much touted results concerning the 'swamping' or 'washing out' of priors – results which have led some commentators to hold that this source of subjectivism is, in the end, much less worrying than might initially be imagined. The relevant theorems prove that, under certain conditions, the posterior probabilities that two agents assign to some pair of rival theories will, given evidence of a certain sort, converge *whatever prior probabilities* (short of zero or one) they may have assigned to those theories. But, aside from detailed issues about whether or not the necessary conditions can plausibly be taken to hold in particular cases, the fact is that these results guarantee agreement – even agreement on the ranking of the rival theories – *only* in the limit, which of course is never achieved in practice. Given any actual theory-ordering in the light of the available evidence that a sensible person would regard as frankly ridiculous – such as, for example, a preference now for special creationism over Darwinism as an account of the present biological furniture of the earth – there must, quite trivially, be priors that the 'ridiculous' agent could have had, such that they conditionalised away fully and accurately in accord with Bayesianism on *all* the accumulating evidence, about the fossil record, homologies and the like and *still* arrived at their clearly unreasonable ranking.[4]

But, aside from the much-advertised problem of the priors, there are two further sources of subjectivism in the Bayesian approach that have received relatively little publicity yet which surely ought to be equally damning in the eyes of anyone who thinks of science as governed by strong objective principles of sound reasoning.

First the notion of evidence itself is subjective in this approach. When reconstructing episodes from the history of science in their terms, Bayesians invariably assume that all sensible agents come to take as evidence what we would all take as evidence. But this is an extra assumption for which there is no sanction in the 'pure' Bayesian account. According to that account, any synthetic assertion e is evidence for an agent (relative to epistemic situation S) if and only if she happens to assign, when in that situation, a subjective probability of one to e. (I here ignore wrinkles about Jeffrey conditionalisation which allows an agent to conditionalise on evidence about which she is less than certain – but again the, in that case non-extreme, probability the agent ascribes to e will be a subjective affair.) Anything that anyone comes to be subjectively entirely convinced of the truth of 'that they were abducted and raped by aliens', 'that god exists', 'that Jesus walked on water', 'that the needle in the meter points to "5" when all the rest of us see it as pointing to "10"', counts as new evidence for that 'agent'. At the point at which such an agent *becomes* convinced of the particular piece of 'evidence' at issue, she is required by Bayesian rationality to modify all her erstwhile degrees of belief by conditionalisation on it – however strange that 'evidence', and therefore those shifts in degrees of belief, look to you or me.

There is still more subjectivism. Let's assume that some Bayesian 'agent' in fact takes as evidence what any sensible scientist would take as evidence. There is still the issue of what counts for that agent as implicit 'background knowledge'. Bayesians are quite clear that all probabilities, at any given time, are to be thought of as relativised to the agent's background knowledge at that time. The distinction between evidence and background knowledge is blurred – perhaps inevitably so. Once an item of background knowledge is articulated, then, I suppose, it automatically counts as evidence for that agent (since she is bound to regard it as having probability one). Again in applications, Bayesians quietly assume that what counts as background knowledge at a given time in science is a more or less universal, intersubjective affair. But, *first*, this will be plausible even as an idealisation of the actual state of affairs in a particular science at a particular time only courtesy of a very selective attitude toward who counts as a competent scientist (as a *bona fide* member of the relevant 'scientific community'). And, *secondly*, there is, so far as I can tell, no official sanction within personalist Bayesianism for this assumption: on the contrary, according to the official 'pure' position, background knowledge is another entirely agent-specific factor – whatever the agent regards as 'given' or as delimiting the space of conceptual possibilities[5] is background knowledge for her, *whatever anyone else might think*.

This also means that there is another almighty slice of subjectivism lurking in the Bayesian's account of belief-*dynamics*. There is a crucial, but under-emphasised, clause in the principle of conditionalisation: your new prior probability on A must be your old probability on A, conditional on e, if in the meanwhile (that is, between the 'old' and 'new' times) e has turned from possible evidence to actual evidence, *and nothing else of epistemic significance has occurred* – that is, no other change has occurred between the two historical stages in your 'background knowledge'. There is again here a significant difference between the way Bayesian theory is standardly applied and the pure general theory. In applications, it is quietly assumed that the general epistemic situation, supplied by 'background knowledge', will be the same for all agents; and that, where the 'reasonable' assumption is that the only change to agreed background knowledge is the addition of some new evidence, this assumption too will be generally shared. However there are no such constraints in Bayesian theory itself – which, so far as I can tell, must leave this ingredient too as an agent-relative affair. A Bayesian agent is, apparently, officially allowed to assert at any point that her personal epistemic situation, her personal background knowledge, has suddenly changed, and hence call for an entirely new round of bets. Indeed she is *required* to do so in order to be Bayesian-rational, if her subjectively perceived epistemic framework somehow changes. Bayesian rationality imposes no need for any *argument* as to why the agent's 'background knowledge' has suddenly shifted and hence imposes no need for any particular argued connection between the new and old priors. Indeed this is not an area where normative considerations play any role – it will just be a descriptive matter whether or not the agent's background 'knowledge' (really, on this approach, set of background *beliefs*) remains unchanged.

Suppose, for example, a special creationist Bayesian sees her initially massive prior for creationism being steadily eroded by conditionalisation on evidence. Viewing her diminishing posterior with alarm, she (of course quite unrelatedly!) suddenly feels that her whole epistemic situation has undergone an abrupt change. Pointing out that all Bayesians agree that all degrees of belief are implicitly relative to general epistemic framework or background knowledge and that this general epistemic framework too is agent-dependent, she can simply 'call for a new round of bets', that is, insist on redistributing her priors against the, for her, new background. Suppose again that, against the new background, the special creationist theory has a massive prior.

This seems the antithesis of how to rank theories scientifically in the light of the evidence, yet the personalist Bayesian cannot but sanction it. Of course, the Bayesian does *not* sanction an agent's simply *pretending* to have undergone a shift in her general epistemic framework in order to 'defend' a theory favoured on non-evidential grounds (and we would all be suspicious of the veracity of any real scientific creationist who made such assertions about their subjective degrees of belief). Nonetheless, if there *were* to be such an agent who genuinely felt that the whole epistemic earth had moved and who ended up with a suddenly increased 'prior' for her scientific creationist view then the Bayesian could not regard her as in any sense irrational.

5. EARMAN AND SHIFTS IN THE 'SPACE OF CONCEPTUAL POSSIBILITIES'

Although such alleged shifts in general epistemic background may appear particularly suspect to the objectively inclined, John Earman has in effect argued (in his 1993) that there are cases where no such intuitions are elicited and where, on the contrary, the intuitively correct analysis does involve such a shift. Indeed Earman argues that the occurrence of such a shift is exactly what characterises a scientific *revolution*.

Earman, like Salmon and myself, investigates ways in which Kuhn's views on theory-change and those of the Bayesians can be reconciled. (His direct aim is to investigate relationships between Kuhn and Carnap, but the latter of course was or became a tempered personalist.) He arrives however at interestingly different conclusions: Earman in fact sees two ways in which Kuhn and the Bayesian must remain at odds and then a third issue on which fruitful cooperation is possible (indeed where the Bayesian must, in Earman's view, accept Kuhn's insights and hence radically augment her position).

On the first allegedly irreconcilable difference, Earman seems to me mistaken. He complains that:

> Kuhn's list of criteria for theory choice is conspicuous for its omission of any reference to the degrees of confirmation or probabilities of theories .... (1993, p. 21)

But surely the 'omitted' criteria should in fact be thought of as implicitly present on Kuhn's list – as either definable in terms of criteria that *are* explicitly mentioned (those of simplicity, perhaps consistency and especially empirical scope and adequacy), or, more strongly, as providing the means of defining those criteria. Simplicity, consistency (both internal consistency and consistency with other well-supported theories) and empirical scope and adequacy are criteria that, intuitively speaking, feed into judgements about degree of confirmation and the probabilities of theories in the light of evidence. Indeed some Bayesians would make the stronger claim that all these other criteria can themselves be defined in terms of probabilities. It seems hard to believe that Kuhn could consider that his views on theory-choice diverged comparatively little from 'those currently received' unless those views could accommodate, in one way or another, something very like degrees of confirmation.[6] This is how Salmon and I have reconstructed Kuhn.

Earman's second reason for seeing Kuhn and the Bayesian as at odds – essentially that Kuhn's notion of 'theory-choice' is fundamentally irreconcilable with a probabilistic approach – is altogether weightier and will be the subject of the next section. Here I want to concentrate on the issue on which Earman sees the Bayesian as needing to benefit from Kuhnian insights – basically in dealing with Kuhn's favourite topic: *scientific revolutions*.

Almost every postwar philosopher of science has implicitly recognised the importance of 'background knowledge' both in the generation and in the appraisal of particular scientific theories. But until recently, few have done much to turn this implicit recognition into a fruitful tool of analysis by looking more precisely at how science depends on 'background knowledge'. Even within the Bayesian approach, and despite the fact that that approach explicitly recognises the dependence of all probability assignments at a particular time on the background knowledge of the time, nothing is said about the exact nature of this dependence. On the contrary, and as we have seen, the way in which probabilities are 'relativised' to background knowledge is generally left as just another area in which subjective judgement inevitably intrudes. Earman – surely correctly – takes it that the Bayesian *needs* to say something more and (wearing his Monday-Wednesday-and-Friday clothes) starts to try to say it.

The chief role played by background knowledge for Earman is in specifying the background 'space of conceptual, theoretical possibilities'. If, in the light of background knowledge, this space is finite or can be finitely partitioned, then eliminative induction *via* observational evidence becomes possible, for example. The important feature for present purposes, however, concerns his view of *changes* in this background space of possibilities. According to Earman such changes are precisely what characterise scientific revolutions – of which he distinguishes two flavours: mild and strong.

> A mild form of scientific revolution occurs with the introduction of a new theory that articulates possibilities that lie within the boundaries of the space of theories to be taken seriously but that, because of the failure of actual scientists to be logically omniscient, had previously been unrecognized as explicit possibilities. The more radical form of revolution occurs when the space of possibilities itself needs to be significantly altered to encompass the new theory. (1993, pp. 24, 25)

Earman claims that even the 'mild form' shows the inadequacy of the standard form of Bayesianism – even mild revolutions cannot be explained by

Bayesian conditionalisation:

> For conditionalizing (in any recognizable sense of the term) on the information that just now a heretofore unarticulated theory T has been introduced is literally nonsensical, because such a conditionalization presupposes that prior to this time there was a well-defined probability for this information and thus for T, which is exactly what the failure of logical omniscience rules out. (1993, p. 25)

And, of course, matters are still worse in the case of 'strong' scientific revolutions when some genuinely new, as opposed to simply hitherto unrecognised, conceptual possibility is introduced. Even if an agent were logically omniscient she could not assign at time t a well-defined probability to a theoretical possibility that did not yet exist at t.

Given this account, Earman sees Kuhn's subjective factors playing a role *not* (as Salmon and I had) in the assignment of prior probabilities within a *given* epistemic framework, but rather in *reassigning* probabilities after shifts in the conceptual space:

> In typical cases [of either mild or strong revolutionary shifts] the scientific community will possess a vast store of relevant experimental and theoretical information. Using that information to inform the redistribution of probabilities over the competing theories on the occasion of the introduction of the new theory or theories is a process that, in the strict sense of the term, is *a*rational: it cannot be accomplished by some neat formal rules, or, to use Kuhn's term, by an algorithm. On the other hand, the process is far from being *ir*rational, since it is informed by reasons. But the reasons, as Kuhn has emphasised, come in the form of persuasions rather than proof. In Bayesian terms, the reasons are marshalled in the guise of plausibility arguments. The deployment of plausibility arguments is an art form for which there currently exists no taxonomy. (1993, p. 26)

The first question that arises about Earman's account is whether or not the idea of extensions of the conceptual space provides a satisfactory analysis of what goes on in a scientific revolution.[7] This is clearly a big issue but there seem to me several reasons for doubt.

Notice two significant features of the account. *First*, Earman admits that the distinction between the two flavours of revolution 'mild' (in which some hitherto unrecognised but actually 'available' theoretical possibility begins to be taken seriously) and 'strong' (in which the new theoretical possibility is genuinely new) is 'blurred, perhaps hopelessly so' (1993, p. 25). *Secondly*, and relatedly, he admits (indeed he insists) that it is generally possible *post hoc* to reconstruct even the 'strongest' revolutions (such as the transition from classical to relativistic physics) as having taken place within a common linguistic and conceptual framework (1993, p. 24).

But, in view of these concessions, the Bayesian might seem well advised to idealise and take it that the warring parties in a 'revolution' are working against a common background of conceptual possibilities, while disagreeing only over which possibility to prefer. After all, any account of the rationality of science is bound to idealise in some ways. This is certainly true of personalist Bayesianism in its standard form and even in the modified version towards which Earman is working. Bayesian agents are assumed, for example, to be perfect deductive logicians at least in the sense of assigning probability one to all logical truths.[8] But suppose s, a complicated sentence of the propositional calculus involving

38 atomic sentences, is in fact a tautology. Even the most logically acute real Bayesian would have problems in immediately recognising – indeed she may live her life without *ever* recognising – that s is a tautology and hence that, whatever she may initially think (if indeed she thinks anything at all), her *real* degree of belief in s is one. Nonetheless the Bayesian supposes that this is what every agent's real degree of belief in s is – and that seems (to me at least) exactly right. So the complaint against taking warring parties in a 'scientific revolution' as working against a commonly agreed background cannot be merely that it idealises.

Earman himself, as we shall consider in more detail in the next section, reacts to the fact that, outright, some scientists *believe* certain theories (that is, in Bayesian terms, assign them probability one) by taking the line that they *ought not* to:

> One can cite any number of cases from the history of science where scientists seem to be saying for their pet theories that they set $p = 1$. Here I would urge the need to distinguish carefully between scientists as advocates of theories versus scientists as judges of theories. The latter role [alone] concerns us ... and in that role scientists know, *or should know*, that only in very exceptional cases does the evidence rationally support a full belief in a theory. (1993, p. 23 emphasis supplied)

But, if sensible idealisation is permitted, what is wrong with assuming, at least in the case of 'mild' revolutions, that, however it may appear psychologically to a given agent/scientist, she does in fact have some initial degree of belief in the 'revolutionary' possibility? The 'revolution' would then, for the Bayesian, simply consist in the (perhaps sudden) decrease, through conditionalisation, in the probability of the 'old' theory, and a corresponding increase in the probability of the 'new' theory.

And if this is the right way to treat 'mild' revolutions, and if the distinction between 'mild' and 'strong' is 'blurred, perhaps hopelessly so', and if it is always possible, as Earman claims it is, to reconstruct the theoretical dispute after the event as taking place against the background of a conceptual space held in common by the disputants, why not treat *all* scientific revolutions, admittedly somewhat idealistically, as taking place against the background of a fixed space of conceptual possibilities? So far as I can tell, Earman's main counter argument is, indeed, that it is psychologically unrealistic – the scientists involved in revolutions did not as a matter of fact themselves explicitly internalise the conceptual possibilities that make it possible to see the dispute as occurring against an agreed conceptual background. But, as I indicated, this seems to me, even if true, not necessarily either here or there.

But is it true? I find it difficult to see Earman's model as instantiated in a range of scientific revolutions. For example, neither the Copernican nor the Darwinian revolutions involved essentially new conceptual possibilities. The heliocentric model had, after all, been articulated as long ago as Aristarchus and was certainly not in any sense unthinkable for his Ptolemaic opponents – they conceded the possibility that the earth moved around the sun, and simply believed that this was (very likely to be?) false for a range of evidential and non-evidential reasons. As for Darwin, pretty well all of his contemporaries seem to have agreed that species have evolved, the only dispute was over the mechanisms, and their relative weights; and here Darwin spent much time stressing the analogy with artificial selection precisely in the attempt to make natural selection seem a natural, non-novel idea. Moreover, Wallace independently had arrived at essentially indistinguishable ideas – surely showing that they were 'in the air' at the time. (Indeed, the strikingly frequent phenomenon of simultaneous discovery in science, even of 'revolutionary' ideas, seems to indicate how much 'in the conceptual air' they *generally* are. Hooke, Wallis and Wren, for example, really did have the idea of universal gravitation and its inverse square relationship to distance – Newton really did hold that his genius was to have 'proved' the theory from Kepler's phenomena while the others 'merely' conjectured it.)

The Einsteinian revolution seems to provide the main stimulus for John Earman's general account. Although I would not dare cross swords with him concerning Einstein, it is a fact that there are other analyses that show the axioms of relativity as derived from new experimental results plus already generally accepted background principles.[9] (Also, remember that so far as *special* relativity goes at any rate, Poincaré is, with apparent justice, regarded as a simultaneous – or even pre- – discoverer with Einstein.)

If one were thinking – not too rigorously – about the history of science with Earman's intuitive distinction in mind, then probably it is the quantum revolution that would seem *prima facie* the 'strongest' of them all. But here too the quantisation of energy has been persuasively argued – in this case, by John Norton – to have been arrived at as a deduction from the phenomena (where this means, of course, deduction from the phenomena *plus* already existing background knowledge).[10] It took the genius of Bohr to show that energy-quantisation could be derived deductively from new experimental results plus already existing, and arguably, generally accepted background principles; but if there is such an argument, then it seems hard to deny that Bohr was showing that everyone, genius or not, *implicitly* recognised energy quantisation as in the space of conceptual possibilities ahead of his 'innovation'. New experimental results may be *surprising*, but it is hard to think of them ahead of their discovery as actually inconceivable; but if 'all' that it takes to arrive at a 'revolutionary' new theory is to plug some new experimental results into a general framework, that is not only conceivable but arguably part of generally accepted background knowledge, then it is hard to see how that new theory can itself have been outside the space of conceptual possibilities beforehand.

The situation so far as the historical episode considered in this paper goes is even clearer. Far from being an hitherto unrecognised conceptual possibility in the early nineteenth century, the wave theory of light had, of course, already been around (in altogether less impressive but still recognisably similar forms) for at least a century and a half. (Hooke, Huygens and, in the eighteenth century, Euler had all developed versions of it.) And again, the chief 'revolutionary' in this case, Fresnel, claimed (with perhaps surprising plausibility) that, given the premise that the corpuscular theory had proved so problematic in view of the experimental phenomena as to be out of the game,[11] his version of the wave theory (complete with the 'luminiferous aether') could be straightforwardly deductively inferred from experimental results plus uncontentious principles of background knowledge.[12]

In sum, it seems to me that Kuhn greatly exaggerated the revolutionary nature of 'revolutionary' change in science; and that John Earman is following suit. But now suppose that, for the sake of argument, we go along with Earman's analysis of scientific revolutions and therefore accept his claim that the Bayesian position needs to be augmented; and suppose further that we agree that that position needs to be augmented using Kuhn's account of the factors involved in 'theory-choice' as something like the right account of how prior probabilities get re-assigned after a revolutionary 'shake-up' of the space of conceptual possibilities. Such an account, in line with Kuhn, would admittedly be 'arational' since it does not conform to some 'neat set of formal rules' (there is no 'algorithm') but this does not mean, suggests Earman, that the process is actually *ir*rational. The process is 'informed by reasons' – though Kuhn is right that these reasons take the form of 'persuasions rather than proof', or of 'plausibility arguments' – an 'art form for which there currently exists no taxonomy' (1993, p. 26).

This would then seem to amount to just another version of the earlier story of 'the problem with Kuhn' finding itself underlined by the partial agreement of his view with that of the personalist Bayesian. 'Objectivists' like myself want to insist that there is, at every stage of the development of science, such a thing as 'the intellectual argument' between two or more competing theories, and at each stage, there is an objectively correct view – no matter how complex it might be – about the state of that argument. No one expects such arguments to be purely deductive ones beginning from uncontentious premises and entailing one of the rival theories (if that is what Earman means by 'proof'). But if all that we have is 'persuasion' relying on 'plausibility' and if, as Kuhn's insistence on the idiosyncratic nature of the subjective factors seems to suggest, and as Earman's endorsement of his Bayesian version of Kuhn suggests he supports, one man's plausibility is the next woman's far-fetched implausibility, then all talk of 'reasons', 'persuasions' and 'art forms' is surely a smokescreen to cover the admission of a sizeable chunk of relativism into the account of scientific theory-change. If, on the contrary, what counts as plausible is meant to be an objective matter, then the whole problem would seem to be to articulate and defend the objective principles that govern plausibility. If the 'neat formal rules' that John Earman recognises on behalf of the Bayesian are not up to this task, then we need to find other, stronger rules. To talk in Earman's way seems simply to surrender the game to the Kuhnian relativist.

And yet Earman must surely be right that we cannot plausibly expect to capture the whole of the complex and rich process of scientific theory-change in anything likely to count as a neat set of formal rules.

Not the *whole* of the complex and rich process of theory-change, certainly; but then, I shall argue, we should never have expected to. Philosophers of science, following Kuhn, have got themselves into a mess and have proved an easy target for *some* of the barbs of Kuhn-inspired, social constructivist-inclined critics by expecting too much. We need a proper (and at the same time more nuanced and yet more modest) identification of what features of this process of theory-change are, and what features *are not*, governed by considerations of 'rationality'. This identification can be made, as I shall indicate in the next section, by following through in some detail the second of the features of Kuhn's account that Earman sees as inconsistent with personalist Bayesianism: namely Kuhn's overly simple notion of 'theory-choice'.

6. WHY 'CHOOSE' A THEORY? RATIONALITY REGAINED

Both in the *Postscript* and in chapter 11 of *The Essential Tension*, Kuhn analyses theory-change in terms of scientists making theory 'choices'. As so often, Kuhn is less clear than one might like, and certainly he attempts no explicit definition of this notion. Implicitly however he *seems* to take choosing a theory to involve, not just the view that that theory is the best available so far in light of the evidence, not just the decision to work on it to see where it leads and how it can be developed, but also taking the theory fully to one's breast, 'accepting' it, believing it to be true.

Earman complains about this from a Bayesian perspective – arguing, that the only sense in which a scientist might 'choose' or 'accept' a theory consistently with the Bayesian approach is exactly 'the innocuous sense [of] choos[ing] to devote [her] time and energy to' that theory (1993, p. 22). To show how 'baffling' for the Bayesian is the idea of choosing or accepting a theory T in a sense that reflects a judgement about T's epistemic status, Earman considers a researcher who performs some introspection and decides that her subjective probability for T in the light of all evidence available to her is p. One Bayesian-kosher sense in which the researcher would surely be said to accept T is if $p = 1$ or is 'so near to 1 as makes no odds'. But

> Such cases ... are so rare as to constitute anomalies. Of course, one can cite any number of cases from the history of science where scientists seem to be saying for their pet theories that they set $p = 1$. Here I would urge the need to distinguish carefully between scientists as advocates of theories versus scientists as judges of theories. The latter role concerns us here, and in that role scientists know, or should know, that only in very exceptional circumstances does the evidence rationally support a full belief in a theory. (1993, p. 23)

While applauding, as indicated earlier, this willingness to override psychological facts (even about eminent scientists) in the name of good general sense, it is not at all clear to me that cases of full belief are either as rare or as unjustified as is here suggested. It surely depends how 'far down' the hierarchy of theories we go: the assertions that perpetual motion machines are impossible, that the heart pumps the blood round the body, that cells contain energy-providing mitochondria, that water consists of molecules consisting in turn of two atoms of hydrogen and one of oxygen, etc., all seem to me to be, given the present evidence, perfectly proper objects of outright or total belief (whatever that might precisely mean).

'Fundamental', 'explanatory' theories – precisely the sort of theory that has triumphed over others inconsistent with it as a result of a 'scientific revolution' – are, though, a different kettle of fish. And concerning them, John Earman is surely correct – although there are cases of scientists who seemed to assign them probability one (indeed I have been told by some scientists that they *need* to believe in the truth of their theories in order to work successfully on them), the sensible view, precisely because of the historical record (a record that underwrites the so-called pessimistic meta-induction), is that they *ought not* to.

Suppose, then, that T is such a fundamental theory (the general theory of relativity, quantum theory, or whatever) and that some sensible research scientist has a high but less than total degree of belief in T. What might the further fact that she *accepts* (or 'chooses') T mean? One possibility, Earman points out, is that, having decided initially that her degree of belief in T is, say, 0.75, she then, by 'accepting' T, converts that probability to one. Earman is again surely right that

> This is nothing short of folly, since she has already made a considered judgment about evidential support and no new relevant evidence occasioning a rejudgment has come in. (1993, p. 23)

But then the only other possibility is that she retains her initial degree of belief in T ($p = 0.75$) but 'acts *as if* all doubt were swept away in that she devotes every waking hour to showing that [all relevant] observations can be explained by the theory, she assigns her graduate students research projects that presuppose the correctness of the theory' and so on. But this simply amounts to an alternative expression of the view of 'acceptance' of a theory as a purely pragmatic decision not reflecting any judgement about the epistemic status of the theory.

I agree, then, that the Bayesian has good reason to be unhappy with Kuhn's idea of theory-choice. But justified unhappiness on this score is not restricted to Bayesians. Our historical case shows precisely why.

Did Brewster continue to 'choose' the Newtonian, corpusular theory of light, despite the availability of Fresnel's wave theory? This question, I suggest, with its implicit commitment to measuring attitudes to theories along one dimension, is inherently unsatisfactory. In so far as one can give an answer at all, it is 'yes-and-no' (or perhaps, for reasons to be explained, 'yes-no-and-no').

Brewster had, remember, rather than a single view, three main, related but independent views about the corpuscle/wave rivalry as it stood around 1830. *First* he made various concessions about the empirical power and predictive success of the wave theory, that can, I think, plausibly be interpreted as allowing that *as things stood* the wave theory had much the stronger empirical support. *Secondly*, he clearly held that, despite its empirical success, the wave theory was not true (or at any rate, not at all likely to be true), and in particular that the elastic solid ether it centrally postulated was not real (or at any rate, not at all likely to be real). *Thirdly*, he seems to have disagreed with Airy, Baden Powell and others about the way forward in optics – seeing grounds for optimism that developing the corpuscular theory further might turn the evidential scales at present favouring its rival.

Kuhn's notion of theory-*choice*, as we saw, seems to involve not just preferring that theory as the best empirically supported theory, not just deciding to work on it to see where it leads and how it can be developed, but also involves taking the theory fully to one's breast, believing it to be true. Brewster 'chose' no theory on this characterisation – I translate his view as entailing that he (a) regarded the wave theory as presently best supported by the evidence, (b) believed in the truth of no *available version* of *any* theory of light and (c) (roughly speaking) chose to *work* on the corpuscular theory.

Earman, as we saw, complains from a Bayesian perspective about Kuhn's idea that choosing a theory involves 'accepting' it (that is, presumably, believing it to be true). But taking choice to involve commitment to truth was not Kuhn's only implicit mistake; his treatment also seems clearly to presuppose that choice is a single, all or nothing affair – you either choose a theory or you choose some rival. And here Bayesianism in a sense follows suit: it allows of course for *degrees* of belief, and suggests that the general case will be that several rival theories have non-zero probabilities, but it is still committed to the idea that brownie points for theories are, so to speak, scalars – an agent ranks theories simply according to the degree to which she believes the theory is likely to be true, given the evidence she has. In fact nearly all philosophers of science have been trapped into thinking about scientific rationality in terms of a single dimension: this theory is more probable than that, this research programme is progressive, that one is degenerating ... and therefore the reasonable guys 'prefer' the first.

But, as the case of Brewster illustrates, the truth is surely that what it is and is not reasonable to believe about a theory is a somewhat more complex matter – one with quite different aspects involving perhaps quite different considerations. So we should ask separately about the rationality of each of Brewster's three different views about the two rival theories he considered. Was any of the three views 'irrational' – or, perhaps better, in order to avoid the unnecessarily aggressive overtones of that word, was any of them contrary to sound scientific reasoning?

Well, clearly not the first view – that, as things stood, the evidence favoured the wave theory – since this was uncontroversial (and correct). According to Kuhn's much-discussed analysis, the claim invariably underlying the positions of the hold-outs is that the phenomena cited by the revolutionaries as telling evidence in favour of their new view can in fact be 'shoved into the box' provided by the older paradigm.[13] And one of the main reasons (perhaps the main reason) that hold-outs cannot justifiably be regarded as 'illogical or unscientific' is, he suggests, that this claim is not demonstrably incorrect. In fact something stronger can be said – there invariably are ideas around at the time of the revolution about which direction the proponents of the older paradigm should 'shove' in: that is, positive ideas about how the evidence that seems to tell in favour of the new theoretical framework might be accommodated by the old. In the case of the wave/corpuscle rivalry, for example, there were ideas around in the late eighteenth and early nineteenth centuries about how to give corpuscular explanations of the phenomena of interference and diffraction *either* as the result of very complicated diffracting forces emanating from 'gross matter' and, at different distances, either repelling or attracting the light-corpuscles, *or* as some sort of physiological effect.

Again Kuhn is not entirely clear, but *if* he is suggesting here that all it takes for the hold-outs to balance the evidential scales is for such 'shoving' to succeed, then he makes a major mistake about the nature of evidential support. As I and others have argued,[14] whatever one's precise account of evidential support, a general adequacy requirement is that such an account entail a big difference between the support lent by phenomena that are 'shoved' *ad hoc* into a theoretical framework and phenomena that are genuinely predicted by such a framework. If Brewster, for example, stood ready to elaborate on his acceptance that the wave theory was, as it then stood, better supported by the evidence by adding that all it would take to bring the evidential scales back into balance would be *any sort* of *post hoc* accommodation within the corpuscular theory of the phenomena predicted by the wave theory, then he too would be making a significant mistake about the nature of empirical support in science. There is, however, no historical evidence that this is the case: in particular Brewster is very modest about his *suggestion* that interference may be a physiological phenomenon (of course this suggestion left him with a great deal to be modest about).[15]

What of the other two elements of Brewster's position?

Brewster could not bring himself to believe in the wave theory and in particular in the ether, 'invisible, intangible, imponderable, inseparable from all bodies and extending from our own eye to the remotest verge of the starry heavens.' He predicted that the wave theory would eventually give way to a quite different one 'after it has hung around for another hundred years or so.' Was he being 'irrational' on this score? Well this would be a strange judgement to make in view of the fact that Brewster was *right*! Indeed if anything he was overgenerous to the wave theory and its elastic solid ether which was to last at best another seventy or so years before being unambiguously rejected.

The history of theory-change in science in general surely requires a separation of judgements about which of the available theories is currently picked out by the evidence from judgements about which theory if any is true (or even likely to be true). The fear, felt by many philosophers, is perhaps that the former sort of judgement is weak to the point of vacuity if separated from the latter – what does it mean for a theory to be 'favoured by the evidence', if not that the evidence makes it more likely to be true than available rivals? This is a legitimate worry, but it is nonetheless just true that Brewster's position – that the evidence favoured the wave theory but that the wave theory was very likely to be false – was consistent, and indeed more than that: clearly reasonable. It seems to follow that we had better make this separation. More on this after considering the third element in Brewster's view of the then current state of play between the wave and corpuscular theories of light.

Brewster seems to have believed that the near-monopoly on talented advocacy and development then enjoyed by the wave theory was bad for science. Let's assume that this means that he believed that there was 'heuristic steam' left in the corpuscular theory, so that development of *it* might eventually lead to a version which was still better favoured by the evidence than the current version of the wave theory. Was *this* view 'irrational'?

Well of course the dominant view in philosophy of science until two or three decades ago was that the contexts of justification and of discovery are quite separate and that rationality considerations come into play only in the former context. Hence Brewster could think what he liked about the way forward in optics without fear of contravening any rule of scientific logic. Nowadays we are more sophisticated. But, however interconnected these two contexts might in fact be, the connecting principle quite plainly cannot be the simple one that the only reasonable course of action is to try to develop the theory that is presently best empirically supported. The obvious point has often been made that such a principle would, apart from anything else, automatically condemn the great innovators of theoretical science – who, almost by definition, are those who start to work on a theory *before* it is the best empirically supported in its field and who, through their work, *turn it into* the best supported theory. There must therefore again be room for a separation between the theory one judges best on the available evidence and the theory one chooses to devote most effort to.

When Lakatos advocated the view that the *primary* domain of rationality is simply the area of empirical support – that is, judgements about which direction the evidence at present tends in, the almost universal reaction was that this was to weaken the notion of rationality to the extent of making it uncontroversial. If all that is needed for, say, a defender of classical physics in 1920 to count as a 'rational' is that she admit that relativity theory is ahead in terms of empirical support as things stand, but is then free to pursue any classical physics project she likes, then, Paul Feyerabend famously remarked, Lakatos' position is simply 'anarchism in disguise'.[16] In fact, though, such judgements of the present 'evidential score' and the fact that scientific rationality demands unanimity concerning them is surely not as trivial a matter as Feyerabend suggested. It is no easy matter, for example, to get a 'scientific' creationist to admit that her theory is presently massively behind evolutionary theory in terms of empirical support – even if you were to provide her with the comforting (though surely false) thought that there have been cases of theories that have started massively behind a rival in terms of the evidence and have eventually managed to turn the tables. But Feyerabend and of course others were right that there ought to be more to good reasoning in science than mere recognition of the present empirical score; and there is.

There is no straightforward connection between (i) the present evidential support enjoyed by some set of rival theories, (ii) the likely truth (or 'approximate truth') of those rivals and (iii) the reasonableness of various research strategies – in particular the strategy of concentrating all one's research efforts on the presently best supported theory. But no straightforward connection does not of course entail no connection at all; and the fact – if it is one – that the first thing to be straight about when it comes to good reasoning in science is the relative degrees of support enjoyed by the available rival theories does not entail that this is *all* one should be concerned about. *Of course* the fact that Darwinism is streets ahead of creationism on the evidence we have does not on its own entail that it is logically impossible for a creationist to produce a theory within her own approach that reverses the evidential tables. But if someone were to tell us that she intended to exploit this possibility it would be sensible to ask her exactly how she intended to proceed. It is difficult to see any sort of heuristic idea within the creationist programme the pursuit of which might turn the trick: indeed the whole *modus operandi* of that programme seems to be to come along after the (empirical) event and absorb evidence as it independently arises. God created the universe in 4004 BC roughly as it now is. How is that? Well, experiment and observe and whatever you find is how God made it! The programme's leading idea supplies an indefinite set of 'free parameters' that the creation scientist fills in as she goes along and this, I hold, is a recipe for creating specific theories that enjoy no real empirical support. The creationist who felt that there are unexploited heuristic possibilities within her programme would, I think, simply be making a mistake.

Returning to the more serious case of Brewster, just as in the case of the question of which theory he 'chose', the question as to whether or not Brewster was irrational or 'illogical or unscientific' (in Kuhn's terms) in holding out against the wave theory has no straightforward single answer. This does not mean, however, that it has no answer (as Kuhn suggests), but rather that it has a slightly more complicated answer. Brewster was right to concede that the wave theory was presently ahead in terms of empirical support. He was right that this does not entail that the wave theory is true (and of course right in particular that it could not be true unless it eventually gave an explanation of the phenomena of dispersion and of selective absorption).[17] As for his views about 'the way forward', we need to ask for more information.

How exactly, except by wishful thinking, did he think that developing the corpuscular theory in 1830 was going to lead to specific theories that might conceivably enjoy predictive and explanatory successes on par with, or perhaps surpassing, those enjoyed by Fresnel's wave theory? The corpuscular programme was by then as bereft of (unused) heuristic ideas as the scientific creationist approach always has been – the difference of course is that there had been significant heuristic ideas behind the corpuscular approach initially, it was just that by 1830 they had all been tried and failed.

In barest outline, the idea of the corpuscular programme was to reduce optics to the Newtonian mechanics of moving objects. Initially the idea had been to effect the reduction to *particle* mechanics – the particles of light being simple entities (though perhaps with different masses or different velocities according to the colour they produced) subject to forces emanating from 'gross matter' (at reflection, refraction and, in passing by the edge of ordinary matter, diffraction). Naturally, the theories that were thought of first gave these forces the forms of other already known forces, but it was clear right from the start that all such theories fail to yield the phenomena. There were special difficulties in the case of diffraction, where it became obvious that, if anything worked to accommodate the known phenomena of the diffraction fringes, it would have to be a highly complicated force law, one making the force switch from attractive to repulsive and back again as the distance from the 'diffracting object' (such as the slit-screen or straightedge) changed minutely. Polarisation phenomena (first discovered via double refraction through crystals such as calcite) clearly showed that light rays could be made to be 'sided' – that is, to exhibit different properties in different planes through the direction of propagation of the ray. This meant presumably that the light-'particles' themselves must be treated, not as Newtonian particles, but as extended bodies with different properties in different 'sides' – a suggestion made by Newton himself and investigated in gory detail by J.B. Biot in the early years of the nineteenth century. Biot succeeded, *partially*, in 'shoving' some of the phenomena predicted by the wave theory into the 'box' of the corpuscular approach, but without any hint of *independent* testability, without any hint of any testable prediction. Brewster faced a 'particle' theory that had already invoked the most complex of forces, had already endowed the light particles with 'poles' and complicated axial movements with respect to those poles and had still not produced anything resembling an empirical success.

If Brewster had some other view and believed that there was some unexhausted general idea behind the corpuscular approach that might yet yield a version of the theory that turned the evidential tables on its wave-based rival, then, so far as I can tell, he was just plain wrong.

Suppose he felt instead that by pursuing some already heavily pursued idea – perhaps if the expression for the diffracting force went to the 25th power of the distance, rather than the 24th – then everything would change: instead of lagging constantly behind the facts the corpuscular theory would suddenly become predictive. The right response then seems to me to be meta-inductive: of course it is logically possible that this might happen, but the evidence from the history of physics seems to be that no amount of flogging has ever revived a horse as dead as corpuscular optics was in 1830.

If, finally, he was simply relying on wishful thinking, serendipity, the idea that maybe by pondering the corpuscular approach some new idea would crop up that turned out to revolutionise the situation, then aside from making obvious remarks about flying pigs, one would need to ask whether the corpuscular approach with some essentially new idea would really be the corpuscular approach rather than some entirely new research programme (and one would need to ask whether even new research programmes arise 'out of thin air' rather than in some methodical way from old background knowledge and new phenomena).

Certainly Brewster's complaint that, in effect, the wave theory was ahead in terms of predictive success because it had more, smarter advocates is at 180° to the truth. Unlike the corpuscular approach, the wave approach had clear unexhausted heuristic resources in 1830. For example, dispersion – the fact that what the wave theory identifies as beams of light of different wavelengths travel at different velocities in the same transparent medium – was, as Brewster emphasised, an anomaly for the then current wave theory. But the wave theoretic prediction of the independence of velocity from wavelength followed only from the assumption that the ether within transparent bodies was the very simplest form of elastic medium: one that obeys Hooke's law exactly. This assumption was always too simple to be good – more complicated elastic media were known, there seemed every reason to think that by complicating the force law somewhat, dispersion would be dealt with.

This is precisely what Cauchy and others attempted. Moreover, and as Airy and Baden Powell pointed out, there were successful precedents to be cited in the wave approach – cases, such as Fresnel's shift from longitudinal to transverse waves, which had proved strikingly empirically (that is, predictively) successful. The wave theory (or rather wave programme) did not have more empirical success because it had more, smarter advocates; rather it had more, smarter advocates because they could see within the approach unexhausted theoretical opportunities for empirical success.

7. CONCLUSION

In this paper I have addressed, occasionally somewhat tangentially, the question of how revolutionary Kuhn's views – more especially, his views on theory or paradigm *change* – really are. I have argued in effect that, like the question of which theory Brewster 'chose' and the question of whether or not Brewster's hold-out views were 'rational', the answer is not straightforward.

Kuhn's general comments about hold-outs and their fundamental rationale are not revolutionary at all. His claim that these hold-outs are right (or, rather, not necessarily wrong) that the allegedly crucial phenomena can be 'shoved' into the older paradigm's 'box' amounts to no more than the Duhem problem with examples. And in so far as it implies (as it seems to) that shoving a phenomenon, predicted by a 'revolutionary' theory, into the box of the older theory means that that phenomenon can supply no reason to prefer the newer theory, it is plain wrong.

In so far as Kuhn's account can be reconciled with that of personalist Bayesianism it is not revolutionary enough – since this agreement simply underlines the insufficiency, the over-subjectivism of both accounts.

Finally I have argued that many of the problems, both with Kuhn and with Kuhn-influenced later studies, stem from another failure to be revolutionary enough: his talk of theory-choice repeats the mistake of taking scientists' attitudes toward the rival theories available to them as measured for rationality or reasonableness along only one dimension.

What is needed, then is a more elaborate and more revolutionary account of scientific 'rationality' – one that recognises the different elements of Brewster's view, explains more clearly what is involved in regarding a theory as the best supported by the evidence if this need *not* entail regarding that theory as the most likely to be true, and explains, more clearly than others have managed, the relationship between what are sometimes called 'acceptance' and 'pursuit'. I do not, of course, claim to have done any more than sketch some aspects of this more elaborate account here.

The right way to proceed, I think, is by concentrating in the first instance, not on individual scientists' choices in any sense of the term, but rather on reconstructing the *intellectual argument* between rival theories at different stages of science. The main objectivist claim is, or ought to be, that there is such a thing as the intellectual argument between competing theoretical views at any stage of science, and that there is such a thing as the objective state of that argument at each such stage. Once put in this way then it seems obvious that the 'state of the argument' may be a more complicated entity than can be reflected by a single set of numbers, in the way of the probabilists, or a single set of judgements – wave theory progressive, corpuscular theory degenerating, in the way of Lakatos, say.

Of course, nothing in the above account should be seen as denying that Kuhn was a major figure. Some aspects of his views will undoubtedly be recaptured in the promised, more sophisticated account. The account I see emerging from my current work will – by delimiting more carefully those attitudes towards rival theories where consensus amongst rational people really ought to be expected from those where different opinions are 'equally valid' – explain at least some of the motivation behind Kuhn's invocation of 'subjective factors', while preserving certain aspects of theory appraisal in the light of evidence as entirely objective (intersubjective). The progress of philosophy of science, like that of science itself (or so I have suggested), is really evolutionary rather than revolutionary.

London School of Economics

NOTES

* This paper is a modified and extended version of the Dyason Memorial Lecture given to the Australasian Association for the History, Philosophy and Social Studies of Science at Auckland in June 1997. I thank the Committee of the AAHPSSS, and especially Robert Nola, for honouring me with this invitation. And I thank Alan Chalmers both for comments on the lecture and for his kindness during my trip to Australasia. I am indebted for helpful research assistance to Antigone Nounou and James Ward.

1 See Worrall 1990.

2 The point about sound is, of course, that it was 'known' to consist of waves, transmitted in its case, through the air. Although by then Fresnel had shifted to the theory that light consists of *transverse* waves (sound waves are longitudinal, pressure waves) and hence it was known that the analogy was by no means complete, nonetheless given that they were, if the wave theory of light was correct, both wave phenomena provided a *prima facie* case for expecting any result found in light to have a counterpart in the case of sound.

3 The case of new evidence provides the most straightforward application of the principle of conditionalisation; but according to some versions of Bayesianism, at any rate (e.g., that advocated by Howson and Urbach, 1994), the idea of 'old' and 'new' probability assignments linked by conditionalisation may be applied whenever one is assessing the confirmatory weight of *any* piece of evidence, new *or old*. This requires some slick footwork concerning how to 'delete' known *e* from the operative 'background knowledge' relative to which all probabilities are assessed.

4 For elaboration of these arguments see my (1993).

5 The idea that the principal function of background knowledge is to delimit the 'space of conceptual possibilities' is one that John Earman has recently been developing in a number of ways, as we shall see below.

6 Earman in fact sees Kuhn's implicit rejection of probability and degrees of confirmation as intimately connected with his explicit rejection of a theory-neutral observation language and 'the largely tacit but pervasive anti-inductivism of *Structure*' (1993, p. 21). I see both of these views (especially the former) however, as confused and having no real influence on Kuhn's (1977) account of theory-choice. (I do though heartily endorse John Earman's remark that 'in the physical sciences there is in principle always available a neutral observation base in spatial coincidences, such as dots on photographic plates, pointer positions on dials and the like' (1993, p. 16). See for example my (1980) and (1985a).)

7 The second question, raised later, is whether, if we concede the accuracy of Earman's account of revolutions, his talk of plausibility arguments and art forms is anything more than a concession that relativism is correct.

8 Systems have been developed – for example in Garber (1983) and Niiniluoto (1983) – in an attempt to solve the old evidence problem, in which Bayesian agents *may* make purely logical discoveries, which may in turn affect their degrees of belief in substantive theories. (So the idea is that, although the facts about Mercury's perihelion may have been known ahead of Einstein's general theory of relativity hence those facts had probability one and so no confirmatory power, what was *not* known in 1914 was the *logical fact* that general relativity entails the precession of Mercury's perihelion.) However (i) this approach clearly requires a modification of the classical theory of probability; (ii) since this brings into the statement of the axioms themselves considerations of what the agent does or does not 'know', it involves replacing crisp mathematically precise axioms with vague ones; and (iii) the idea that it solves the old evidence problem is a non-starter (it is the substantive *evidence* of Mercury's perihelion, facts about Mercury's orbit that ought to confirm general relativity, not some logical truth).

9 See for example Zahar (1989).

[10] See especially Norton (1993).

[11] Needless to say, this 'premise' was itself not universally accepted.

[12] For details, see my (2000).

[13] See Kuhn (1970), pp. 151, 152.

[14] See, for example, my (1985b) and (1989a).

[15] The idea seems to have been that the light-particles in each of the two streams might arrive at the eye at distinctive intervals and that the two different intervals for the two streams might be such that the vibrations they each set up *within the eyeball* produce particular interference patterns at the retina. Of course there is nothing automatically unscientific in invoking physiology within optics (the wave theory, for example, *correctly* uses the limitations of our visual apparatus to explain the absence of observable interference patterns when two closely adjoining *but independent* point light-sources are trained on a screen). The problems with Brewster's suggestion on behalf of the corpuscular theory were that (a) no one ever succeeded in turning this explanation-sketch into anything like a full and adequate explanation and (b) there were never any independent tests of the idea. At best it showed how *one might* explain interference patterns on the corpuscular theory, but there was never any independent reason to take this possibility seriously.

[16] See, for example, Lakatos (1978, p. 110) and Feyerabend's (1975), dedicated, of course, to 'Imre Lakatos, fellow anarchist'.

[17] He even conceded, remember, that the wave theory's empirical success meant that 'it must contain among its assumptions some principle which is inherent in ... the real producing cause of the phenomena of light' (1838, p. 306). It might be argued from a structural realist perspective (see my 1989b) that, in the light of the history of scientific 'revolutions', this sounds like exactly the view it is reasonable to have concerning the truth claims of current theories.

REFERENCES

Airy, G.B., 1831: *A Mathematical Tract on the Undulatory Theory of Light*, London.

Brewster, D., 1833a: 'A Report on the Recent Progress of Optics', *British Association for the Advancement of Science, Report of the First and Second Meetings 1831 and 1832*, London.

Brewster, D., 1833b: 'Observations on the Absorption of Specific Rays, in reference to the Undulatory Theory of Light', *Philosophical Magazine*, 3rd Series, 2, 360–363.

Brewster, D., 1838: Review of Comte's *Cours de Philosophie Positive*, *Edinburgh Review* 67, 279–308.

Earman, J., 1993: 'Carnap, Kuhn and the Philosophy of Scientific Methodology' in P. Horwich (ed.), *World Changes: Thomas Kuhn and the Nature of Science*, MIT, Cambridge MA, pp. 9–36.

Feyerabend, P., 1975: *Against Method*, New Left Books, London.

Garber, D., 1983: 'Old Evidence and Logical Omniscience in Bayesian Confirmation Theory', in J. Earman (ed.), *Testing Scientific Theories*, University of Minnesota Press, Minneapolis, pp. 99–131.

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

Kuhn, T.S., 1970: *The Structure of Scientific Revolutions*, second enlarged edition, University of Chicago Press, Chicago.

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

Lakatos, I., 1978: *The Methodology of Scientific Research Programmes: Philosophical Papers*, Vol. 1, Cambridge University Press, Cambridge.

Niiniluoto, I., 1983: 'Novel Facts and Bayesianism', *British Journal for the Philosophy of Science* 34, 375–379.

Norton, J., 1993: 'The Determination of Theory by Evidence: the Case for Quantum Discontinuity, 1900–1915', *Synthese* 97, 1–31.

Powell, B., 1841: *A General and Elementary View of the Undulatory Theory of Light as Applied to the Dispersion of Light and Some Other Subjects*, London.

Salmon, W.C., 1990: 'Rationality and Objectivity in Science or Tom Kuhn meets Tom Bayes', in C. Wade Savage (ed.), *Scientific Theories*, University of Minnesota Press, Minneapolis, pp. 175–204.

Worrall, J., 1980: 'Feyerabend and the Facts' in H.-P. Duerr (ed.) *Versuchungen: Aufsatze zur Philosophie Paul Feyerabend*, Frankfurt (republished in G. Munévar (ed.), *Beyond Reason*, Kluwer Academic Press, Dordrecht, 1991, pp. 329–353).

Worrall, J., 1985a: 'The Background to the Forefront', in P.D. Asquith and P. Kitcher (eds.), *PSA 1984*, Vol. 2, Philosophy of Science Association, East Lansing, Michigan, pp. 672–682.

Worrall, J., 1985b: 'Scientific Discovery and Theory-Confirmation' in J. Pitt (ed.), *Change and Progress in Modern Science*, Reidel, Dordrecht, pp. 301–332.

Worrall, J., 1989a: '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 University Press, Cambridge, pp. 135–157.

Worrall, J., 1989b: 'Structural Realism: the Best of Both Worlds?' *Dialectica*, 43, 99–124 (reprinted in D. Papineau (ed.), *Philosophy of Science*, Oxford University Press, Oxford, 1996, pp. 139–165).

Worrall, J., 1990: 'Scientific Revolutions and Scientific Rationality: The Case of the "Elderly Hold-Out"', in C. Wade Savage (ed.), *Scientific Theories*, University of Minnesota Press, Minneapolis, pp. 319–354.

Worrall, J., 1993: 'Falsification, Rationality and the Duhem Problem: Grünbaum vs Bayes' in J. Earman, A.I. Janis, G.J. Massey and N. Rescher (eds): *Philosophical Problems of the Internal and External World*, University of Pittsburgh Press, Pittsburgh and Konstanz, pp. 329–372.

Worrall, J., 2000: '"Heuristic Power" and the "Logic of Scientific Discovery": Why the Methodology of Scientific Research Programmes is less than half the story', in G. Kampis, L. Kvasz and M. Stoeltzner (eds.), *Appraising Lakatos: Mathematics, Methodology and the Man*, forthcoming, Kluwer Academic Press, Dordrecht.

Zahar, E., 1989: *Einstein's Revolution: A Study in Heuristic*, Open Court, Chicago and La Salle.