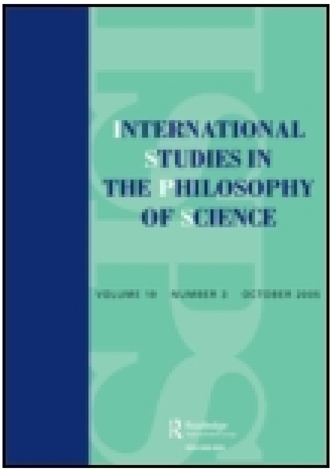

International Studies in the Philosophy of Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cisp20

Rationality, sociology and the symmetry thesis

John Worrall [a]

[a] Department of Philosophy, Logic and Scientific Method, The London School of Economics, London, WC2A 2AE, United Kingdom

Published online: 09 Jun 2008.

To cite this article: John Worrall (1990) Rationality, sociology and the symmetry thesis, International Studies in the Philosophy of Science, 4:3, 305-319, DOI: 10.1080/02698599008573370

To link to this article: http://dx.doi.org/10.1080/02698599008573370

Rationality, sociology and the symmetry thesis

JOHN WORRALL
Department of Philosophy, Logic and Scientific Method, The London School of Economics, London WC2A 2AE, United Kingdom

Abstract *This paper attempts to clarify the debate between those philosophers who hold that the development of science is governed by objective standards of rationality and those sociologists of science who deny this. In particular it focuses on the debate over the 'symmetry thesis'. Bloor and Barnes argue that a properly scientific approach to science itself demands that an investigator should seek the same general type of explanation for all decisions and actions by past scientists, quite independently of whether or not she or he happens to agree with those decisions or approve those actions as 'correct' or 'rational'. I try to improve on previous treatments of the 'rationalist' position (by Lakatos, Laudan, Newton-Smith and Brown) and clarify the exact asymmetries to which the 'rationalist' is, and is not, committed.*

1. Introduction

A major theoretical change occurred in the science of optics in the 1820s to early 1830s. In the 18th century, under Newton's influence, the dominant account of light had been a sort of positivistic version of the theory that light consists of material particles. (The theory was *officially* committed only to 'parts' of light, and the identification of these parts with real material particles was often explicitly disowned—again following Newton; but when it came to it, suppositions were made about these 'parts' which made them remarkably like material particles.) Due almost entirely to the work of Augustin Jean Fresnel, whose first paper was published in 1815, this particulate theory was challenged, and eventually displaced, by the theory that light consists of periodic disturbances (waves) transmitted through an all-pervading mechanical medium.

This change did not occur all at once: some scientists interested in optics adopted Fresnel's theory more or less immediately, while others continued for varying periods to 'hold out' for particles. In England, in particular, we find Sir George Biddel Airy, British Astronomer Royal, writing in an explanatory text published in 1831:

> 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)

On the other hand, in a 'Report' delivered that same year, Brewster—no less a knight of the realm than Airy—was writing (1831, p. 361):

> [the undulatory theory] is still burthened with difficulties and cannot claim our implicit assent...

Brewster continued throughout his career to express his experimental results in the language of the Newtonian particulate theory; while Airy dedicated himself to the development of the wave theory.[1]

Scientific theories themselves undergo modification in the light of empirical and conceptual difficulties and are judged relative to available, and therefore changing, empirical evidence. Hence, the question of what it was rational to believe about the two theories is highly time-dependent. But there is, I believe, no doubt that, *by 1831*, the wave theory had made sufficiently many impressive predictions and had shown itself sufficiently powerful in overcoming conceptual problems to rank objectively way ahead of the particulate theory (which had by then become little more than a ragbag of disconnected assumptions).[2] Suppose that some account of scientific rationality does deliver the judgement that the wave theory ought to have been adopted as the most likely way forward. Is that same account of rationality then also committed to the view that the beliefs and actions of Airy and his fellow 'revolutionaires' are to be explained in an entirely different way from the beliefs and actions of 'hold outs' like Brewster? In particular, does such an account of scientific rationality imply that Airy's beliefs and actions are to be explained simply on the grounds that they were the scientifically correct or 'rational' ones; while, on the other hand, Brewster's beliefs, being 'irrational', require explanation in terms of 'external' sociological factors—perhaps in terms of factors which made him fundamentally conservative in matters political as well as scientific? Or should both men's beliefs be explained in the same general way? After all, Airy had a personal and social history too and his intellectual outlook bore the marks of that history no less than did Brewster's.

The claim that there is in general an important explanatory asymmetry in cases such as these has been defended by many historically minded philosophers of science such as Laudan (1977) and, perhaps most vigorously, Lakatos (1970). Lakatos insisted that a historian should first try to explain developments in science as rational using whatever theory of rationality he holds; only if he fails to explain an event, in particular the acceptance of some theory by some group, as rational will he resort to 'external' (non-rational) factors. Newton-Smith (1981, p. 238) has nicely tagged this as the view that 'sociology is for deviants'.

On the other hand, and as is well known and also hardly suprising, the sociologists of science themselves have not been entirely happy with this restricted view of their intellectual province. Bloor and Barnes in particular have insisted on the *symmetry thesis*: a proper account of the development of science should give the same general type of explanation of the actions and beliefs of scientists, *independently* of whether those actions and beliefs were 'correct' or 'incorrect' according to any evaluative standards we happen to hold. Thus, applying their claims to my historical example, the fact, if it is one, that you or I think of Airy as 'right', the fact that we would endorse Airy's view against Brewster, can play no role in legitimate explanation. As Bloor put it:

> The investigator should not assess the beliefs he studies so as to use that assessment in deciding what kind of explanation to offer, e.g. offering a causal account of beliefs he rejects and treating beliefs he accepts as self-explanatory, self-evident or generally unproblematic. (Bloor, quoted in Newton-Smith, 1981, p. 250)

The engineer does not, Bloor remarks elsewhere, supply one type of explanation for bridges that fall down and another for those that stand up, both are explained scientifically, 'causally', in terms of stresses and strains, strength of materials, and so on; similarly the physiologist does not give one type of explanation of X's health and quite another of Y's ill health; the physiologist aims to explain both 'normal' and 'abnormal' functioning of organisms in terms of the same underlying mechanisms (subject to a range of internal differences and to different 'external' influences). Why then, Bloor asks, should the student of science look for entirely different kinds of explanation in the case of 'true' (he really means 'allegedly rationally justified') and of 'false' ('rationally unjustified') beliefs? Thus, unlike his misguided rationalist philosopher opponents, 'the sociologist seeks theories which explain the beliefs which are in fact found, regardless of how the investigator evaluates them' (Bloor, 1976, p. 3).

This question of symmetry in historical explanations of 'theory-choice' has close, more or less obvious, links with other fundamental issues: for example that of how (if at all) methodologies of science, that is, accounts of scientific rationality, are to be 'tested' against scientific practice. Moreover the symmetry claim is not restricted in its application merely to historical cases and nor is it restricted to choices between (allegedly) good and better scientific theories. Instead it applies equally well, for example, to the *present day* 'choice' between Darwinism and 'scientific' (or pseudoscientific?) creationism: according to the symmetry thesis, Ernst Mayr's or John Maynard Smith's preference for the former and Duane T. Gish's or Oral Roberts' preference for the latter are all to be given the 'same type' of explanation. Thus the whole idea of a special scientific rationality and, with it, the whole idea of an evaluative distinction between science and pseudoscience are at stake, as Bloor, Barnes and the rest make entirely clear. No wonder then that the question has received a good deal of discussion of late; the most direct philosophical discussions being those of Newton-Smith (1981) and Brown (1989).

I here try to sketch what seems to me the right view about scientific rationality and the explanation of events in the history of science such as the adoption of certain beliefs or certain 'theory choices'. I shall not attempt to defend the rationalist's position (except incidentally) but only try to clarify what exactly that position is. Although I have learned a good deal from thinking about the treatments of the issue developed by Newton-Smith and by Brown, I have several disagreements of detail and some of general approach. Enough I think to justify taking the liberty of presenting an independent treatment, rather than describing their views and developing mine by contrast (though I shall underline some of the disagreements and points of overlap as I go along).

2. The irrelevance of our present views for historical explanation

The problem

One of the arguments underlying Bloor's advocacy of the symmetry claim rests on the obvious point that what anyone thinks of as good science *now* cannot in itself play any role in explaining why some historical theory choice was made *then* (say in the 19th century). Any explanation of, for example, Airy's attitude toward the rival theories of light of his time must invoke facts about Airy and not about Carnap, Lakatos or indeed Bloor. And this is true quite independently of whether we construe the important explanatory factors as reasons or as causes. Obviously, if Airy was caused to hold the preference he did, then the causes operated on Airy, not on 20th-century philosophers; and even if we hold that Airy held his preference for (good) reasons, they clearly have to have been Airy's reasons—considerations which counted as reasons for Airy himself. This point would seem to be so obvious as to be scarcely worth making except that Bloor and others seem to think that their opponents deny it. Indeed Newton-Smith agrees with them: Lakatos at any rate did not take this obvious point into account.

Newton-Smith expresses the point raised by Bloor in the form of a question:

> What should our judgments now about who was right, who was wrong, who was being reasonable, who unreasonable during some past scientific controversy have to do with the explanation of why things turned out as they did? (1981, p. 240)

The fact that the obvious answer is a resounding 'nothing!' is taken by Newton-Smith as a criticism of Lakatos's account of normatively based historical explanation and hence of his account of 'testing' evaluative methodologies against history. Lakatos assumes that showing that a past scientist made what the methodology of scientific research programmes pronounces the scientifically correct choice in the evidential circumstances in which that scientist found himself is enough in itself to explain that choice. But clearly

> ... to show that past scientists made what we regard as the right choice in no way explains why they made the choice. For that we need to known what their goals and beliefs were. (*op. cit.*, p. 244)

Newton-Smith infers that Lakatos's account of how methodologies may be confirmed by history—at any rate on my (Worrall, 1976) elaboration of this account—is badly mistaken. Suppose that methodology *M* says that as the evidence stood at time *t* theory *X* was scientifically superior to theory *Y* and suppose that scientist *S* (or a large group of scientists *G*) did indeed reject *Y* in favour of *X*. Still it would be a mistake to conclude that this supports methodology *M*—support really requires that *the reasons* S *(or* G*) had for preferring* X *were* (roughly speaking) *the same reasons as those cited by* M.

Suppose, absurdly that Airy had had a little known colleague, G. B. Fairy. Fairy also accepted the wave theory, but unlike Airy had a rather lurid history. Fairy's parents liked to shower each other with small particles of dirt while having sexual intercourse and little George was frequently the unseen witness of these 'perverted' acts. Suppose furthermore that Fairy's hated father had later, while promenading on the seafront at Brighton, been swept away by a tidal wave, leaving George and his mother alone for a decade of prepubertal bliss. Suppose further (I said the example was absurd) that Freudian theory were true and that it is a prediction of that theory that anyone with this history will have a strong antipathy toward anything involving small particles and, on the contrary, be strongly predisposed toward anything involving waves. Suppose finally and to make the case stronger, that it turned out that Fairy paid no attention to any of the experimental results that our favoured methodology *M* tells us were the important results which swung the balance in favour of Fresnel's theory. In such a situation the correct explanation of Fairy's preference for the wave theory would clearly not be the scientific superiority of that theory (in view of the evidence then available). And, by the same token, the fact that he made what some general account of scientifically correct choices, *M*, tells us was the scientifically correct decision would scarcely count intuitively as a success for *M*.

On the other hand, Newton-Smith points out, the fact that *S* (or *G*) preferred some theory which *M* tells us was *not* the best then available hardly counts, on its own, as establishing that *S* (or *G*) acted 'irrationally' (even for someone who accepts *M* as correct). Suppose, to use Newton-Smith's own and much more sensible example, that the physics community at the turn of the century preferred Planck's quantum theory as an account of radiation over a 'combination' of Wien's law and the Rayleigh-Jeans law (the former to be applied at high temperatures, the latter at low) because the members of the community all endorsed the aim of obtaining theories which are maximally unified (subject to the constraint that they also be empirically adequate). Then, whatever methodology one endorses and *whether or not that methodology includes the injunction to maximize unity*, that scientific community surely acted rationally in preferring Planck's theory—they adopted that theory which best achieved the goals they set for scientific theories:

> Whether we endorse that principle [of unity] or not is irrelevant to the success of the explanation [of the choice of Planck's theory as resulting from the desire for unity]. What matters is that the community whose activities we are seeking to explain held that [unity] was an important *desideratum.* (Newton-Smith, 1981, p. 245)

The solution: the rationalist should claim that good scientists have always applied his principles

Whether or not Lakatos or any other 'rationalist' ever *really* denied it is an exegetical issue that I intend to avoid. Whatever the truth on that issue, it is *clearly* true that the factors cited to explain a historical agent's actions must indeed relate to the agent himself. On the other hand, as Newton-Smith allows, if we are to cite the agent's choice as an example of progress as opposed to mere change, then we must *endorse* his aims and norms. The issue for the 'rationalist' as I construe her is *not* merely—using Newton-Smith's own terminology—'minirationality' (that is, 'hypothetical', means-ends rationality) but 'maxirationality'. That is, the 'rationalist' envisages explaining, say, Airy's (1831) preference for the wave theory not (or not simply) as constituting the best means to the cognitive ends Airy happened to have, but rather as the scientifically correct preference, as constituting, if you like, the best means, not (or not simply) to Airy's ends, but to the ends of *science*, objectively construed.

There is a straightforward way of reconciling the claim to present a 'maxirat' explanation with the 'concession' that the explanatory reasons must be the agent's own. The 'rationalist' need only adopt the conjecture that *the general principles articulated in his favoured methodology were in fact the general principles applied—of course often implicitly—in a wide range of (perhaps nearly all) cases in the history of science.* So the proponent of methodology M who applies this rationalist model of historical explanation will claim that Airy, for example, preferred the wave theory because he 'at bottom' applied the principles of M (without having articulated those principles in exactly the terms used by the later methodologist). Indeed the conjecture is still stronger that (competent) scientists invariably do ('in effect') apply M, except when some 'external' factor intervenes.[3]

At least three interesting criticisms can be raised against this claim that scientists standardly have 'in effect' always judged theories in accord with a methodology M which was only recently fully articulated. The first (implicit in Newton-Smith's treatment) is that it is the wrong claim for a rationalist to make: the general rationalist line, complete with endorsement of some ('scientifically justified') choices, does not require that we share the cognitive aims of the historical agent whose choice we endorse as rational. I shall argue that this is wrong. The rationalist is in fact committed to the full claim here: if we take the methodology of scientific research programmes (MSRP) as an example, then the MSRP proponent *must* claim that Airy (and Newton(!), Einstein and the rest) were 'at heart' themselves MSRPists (when doing their science). However (and here come the other two criticisms), this is, you might think, surely absurd. Or, if not downright absurd, at least monumentally implausible. I shall argue that it is neither and in fact I claim (though shall not argue here) that it is true.

First, the criticism that the claim is the wrong one for a rationalist to make.

First criticism of the solution: 'correct' aims and norms have themselves (progressively) evolved

I claim then that a rationalist (henceforth for convenience 'she') should avoid the charge that she mysteriously makes historical decisions dependent on general theories of 'reason' articulated only later by asserting that historical agents did indeed apply what she sees as 'reason' all along (in their genuinely scientific work[4]); all she, the rationalist, did was articulate the principles that those historical agents implicitly already accepted. Newton-Smith agrees that the rationalist needs to endorse the aims and goals of some historical agents, precisely those who accepted those theories which constituted *progress* over their predecessors. It is exactly this endorsement which turns a 'minirat' into a 'maxirat' explanation. He insists, though, that endorsing a scientist's cognitive aims and goals does not need to mean that the current rationalist actually *shares* those aims and goals. This is because

> We obviously [!] have to allow that there can be reasonable disagreements about the proper goals and methods of science. (Newton-Smith, 1981, p. 246)

'[G]oals and/or methods alter through time' (*op cit.*, p. 246) in a way which, if reasonable, is 'itself progressive, representing an improvement in our ways of learning about the world' (*op cit.*, p. 245).

Newton-Smith's insistence on change (and progress) in methodology as well as in first-level scientific knowledge anticipates Laudan's (1984) much more elaborate development of this thesis. However, as I have argued at length against Laudan (Worrall, 1988, 1989), the position collapses into relativism *unless* some core of methodological principles are regarded as fixed—as constituting rationality—and therefore outside of (because governing) the historical evolution. The scientist whose actions a philosopher wants to explain as rational must be assumed to have 'effectively' held at least these core principles. Newton-Smith talks of 'progress' in methodology, but how is *this* to be judged? If it is *merely* a report of how things look from 'our' present vantage point, then rationality is, after all, relative to (changing) standpoint, an 'improvement' in methodology is whatever those 'scientists' at a particular time happen to have accepted as an improvement. This is pure historical relativism, a position which Newton-Smith energetically (and to my mind correctly) resists. The only alternative is that the standards for judging methodological improvements at least are fixed, 'outside the game'. Some principles have to be assumed fixed (and therefore to have been implicitly held by all 'rational' persons): I do not see any reason to think these principles are meta-methodological rather than simply methodological.

Newton-Smith may have been misled, in a way in which Laudan was, I believe, definitely misled, by the elasticity of the term 'methodology' (and correlated elasticities in terms like 'norms' and 'goals of science'). There are in 'ordinary' usage narrower and wider senses of these terms. If 'methodology' is construed in its wider sense, then the history of science undoubtedly exhibits methodological change. It is in this wider sense that there 'obviously' can be 'rational disagreements about methods'. But it is the *narrower* sense of 'methodology', as formal, 'logical' criteria of theory appraisal, that Newton-Smith has in mind most of the time (and indeed ought to have in mind since these are the important ones). There are not only no convincing historical cases of changes in 'narrow' methodology, but the rationalist cannot coherently allow the possibility of such change.

I argue these points at length in my recent interchange with Laudan (1989; Worrall, 1988, 1989). I shall here simply re-sketch the bare outlines of the argument.

It would, as Laudan points out, be strange if we had not in some sense learned how to do better science alongside actually doing better science. Thus the idea that the best theories were deterministic, and correlatively the goal of producing fully deterministic accounts of given ranges of phenomena, were in a sense parts of the methodology of 18th and 19th century science. Since the advent of quantum mechanics (and the apparent failure of 'hidden variable' reductions) this is no longer the case. Similarly (one of Laudan's own favourite examples) it was until comparatively recently no part of the 'methodology' of clinical trials that they be performed 'double blind'. Yet nowadays no such trial would be taken seriously if it were not performed in this way. Surely this is an example of methodological change and of methodological progress.

However, the claim that these changes did constitute progress presupposes that, underlying the changes, there is a core of unchanging evaluative principles—methodological principles in the narrower, more formal sense implicit in the approaches of the logical positivists, Popper, Reichenbach and others. The reason for the success of quantum mechanics, and the consequent abandonment of the 'requirement' of determinism, was that repeated attempts to satisfy these formal and fixed principles with a theory which was deterministic failed, while the quantum theory itself did satisfy those same formal principles, while being non-deterministic. Similarly the only way to account for the 'double blind' breakthrough as constituting genuine progress is as the result of a substantive discovery—roughly that the efficacy of some treatment for some disorder may be (partly) due to the 'placebo effect'—which was plugged into an *unchanging formal principle of good science:* that hypotheses should always be tested against plausible rivals. (The discovery that the placebo effect sometimes operates means that placebo is always a plausible alternative to the claim that the characteristic features of some drug treatment, say, are the major causal determinants in the observed effect.)

Thus the assumption that the basic formal (narrow) methodological principles are fixed is needed if change in 'substantive' (wide) methodology is to be explained as rational. I do not, then, believe that the rationalist can assume anything less than that those scientists who made the 'right' decisions did so because they were 'basically' applying her general principles of rationality.

Second criticism of the solution: it is absurd

Many critics have regarded this assumption as absurd. Suppose, purely for the sake of illustration, that I claim that some (revised) version of MSRP captures the general principles which scientists standardly use in appraising theories. Suppose, in particular, that I claim that Airy preferred the wave theory because he implicitly accepted MSRP. The idea that this is absurd seems to be based on the belief that if this claim does not actually require Airy to have read Lakatos, it does at least require him to have used the conceptual apparatus of 'research programmes', 'protective belt', 'degeneration' and the rest (or at least some close and obvious equivalents of these notions). But once articulated, this belief is clearly mistaken.

To see why, consider the case of deductive logic. Suppose (ahistorically of course, but the details do not matter to the general point) that Aristotle was the first to codify (a significant part) of the principles of deductive reasoning. Would someone who claimed—as Aristotle surely would, or at any rate as I want to on Aristotle's behalf—that people had been 'doing Aristotelian logic all along' be committed to the absurdity that they somehow had preknowledge of Aristotle's work or that they had talked in terms of 'syllogisms', 'major' and 'minor premises' and the rest? Obviously not: hence the famous phrase 'man didn't need Aristotle to make him rational (logical)'. The claim need only be that, often without being aware of it, and certainly without themselves being able to articulate the general principles they were 'in fact' relying on, people before Aristotle were reasoning in accord with Aristotelian syllogistic principles. The idea that the same holds not only for deductive logic, but also for 'inductive logic'[5] may be false but is surely not absurd. There is nothing absurd, for example, in assuming that for its whole history the human race has not only been standardly inferring *q* from if *p* then *q* and *p*, but has also been standardly impressed by theories that make successful predictions and standardly suspicious of non-independently testable *ad hoc* manoeuvres.

Third criticism of the solution: too much 'false consciousness'

A more sophisticated objection to the claim that scientists have always implicitly judged theories in accord with some recently proposed appraisal system is that it requires so much 'false consciousness' on the part of scientists of the past as to be, if not ridiculous, then at any rate highly implausible (e.g. Laudan, 1989).[6] Pursuing the Aristotle/deductive logic analogy: while people before Aristotle may not have described their reasonings in syllogistic terms at least there are not many documented cases of apparently successful reasoners expressing views on their reasoning directly at odds with Aristotle's system. However, the modern day proponent of, say MSRP is, it has been alleged, confronted with scientist after scientist who not only failed to use the explicit categories of that methodology but who did use the explicit categories of a methodology entirely at odds with it. The most famous case here of course is Newton, who often did explicitly describe the method he followed in arriving at his theories and who made claims, for example about 'deducing' his theories 'from the phenomena', which seem clearly at odds with MSRP.

There is always the let-out, exploited to the full by Lakatos himself, of doling out liberal helpings of 'false consciousness'. One should not necessarily expect great scientists to be very great at describing how they did their great science; after all, as Lakatos used to joke, no one expects fish to be experts in hydrodynamics.

However, this let-out surely becomes less and less plausible, the more it needs to be employed. No doubt some scientists have exhibited false (methodological) consciousness, but were Lakatos (or the proponent of any similar methodology making similar claims to 'explain' the history of science) to be forced to regard false consciousness as virtually universal in the scientific enterprise then his position would be highly unattractive. As Laudan has remarked:

> [The idea that] scientists' implicit judgments about theories and evidence are virtually never wrong, [while] their explicit accounts of their reasons for theory preferences are virtually never right ... is a monumental psychological implausibility (1989, p. 382)

I accept that this would be an implausible claim. But notice, first, that it is not at all implausible, let alone monumentally implausible, that Newton, for example, exhibited 'false consciousness' when considering the evidential claims of his alchemical hypotheses.[6] After all he seems *implicitly* to have given these hypotheses greater weight than they deserved from a scientific point of view. What *would*, I happily concede, be implausible is that Newton should have been consistently and seriously awry when discussing the methodology he followed in making what are generally agreed to be his major scientific breakthroughs in gravitation and in optics. (Although even here, *some* clash between explicit and implicit methodologies is not at all implausible.) In general it is only when considering 'correct' scientific moves that it seems implausible that scientists should be consistently wrong in describing their procedures.

Fortunately the amounts of 'false consciousness' that need be invoked so far as unambiguously scientific developments are concerned—and assuming the right account of theory appraisal—are, I believe, in fact very small. I cannot argue this claim in detail here. Newton is surely the most striking apparent counterexample and I try to show in a forthcoming paper that claims about his explicit methodology being out of line *both* with his actual practice, *and* with more recent methodologies have been greatly exaggerated. Certainly in the example we have been considering Airy's remarks about the importance of the successful empirical predictions which had been made by Fresnel's wave theory, about the heuristic promise of that theory and the fact that the alternative, particulate approach had no theoretical resources which had not already been exhausted without success are all perfectly in line with what I take to be the correct appraisal criterion.[7]

3. The symmetry thesis

Suppose, then, that I make the claim that scientists standardly do judge theories in accordance with some general methodology M. How exactly does this claim figure in historical explanations? Is the style of such explanations 'symmetric' with respect to 'reasonable' and 'unreasonable' beliefs (as of course judged by M)?

I think it instructive to approach this question *via* the analagous question concerning *deductive* logic, where the issues are, I think, altogether clearer. Suppose, then, I propound the, in this case standard, account of deductive validity of inference. How exactly does this account become involved in explanations of the way people actually do evaluate inferences? And are such explanations 'symmetric' with respect to correct and incorrect beliefs about validity?

One difficulty in this case is to make sure that the subjects whose evaluations are to be explained have distinguished validity and what is sometimes called soundness (sometimes 'correctness') of argument: that is, that they have separated the question of whether the premises seem true from the real question of validity—*if* the premises were true would the conclusion *have* to be true as well?

Assuming that this problem has been overcome, suppose that two arguers are both confronted by the same inference, I, which is, let us suppose, as a matter of (logical) fact invalid—B (for convenience 'he') says it is valid, A (for convenience 'she') that it is invalid.

We investigate A's 'reasoning' and find that although she has not been formally trained in deductive logic and does not have the formal categories at her fingertips, she has not 'sleepwalked' to the correct conclusion either; instead she has really got the right 'basic' idea (perhaps she talks about it being 'possible' for I's conclusion to be false even if its premises are true—in some intuitive sense of possibility that she is, at any rate for the moment, unable further to clarify).

On the other hand, investigation of B's (incorrect) belief that the inference is valid reveals, let us say, a desperate need to believe in the truth of the conclusion— perhaps B is expecting to die shortly and the conclusion of the argument is that there is 'life' after death.

It seems plausible to say then that part of the explanation for B's belief that the inference is valid is his desire to believe the conclusion. Of course, this explanation may be correct without being 'ultimate' or basic (failure to make this simple distinction is at the root of a good deal of confusion in this area and others). No doubt there is a deeper story to be told in terms of B's personality and a still deeper explanation in terms of his genetic and environmental history.

As for A, it would clearly be incorrect to say that the invalidity of the inference (as established, say, by representing it in first-order logic and producing an interpretation in which the premises are true and the conclusion false) in itself explains A's belief that it is invalid. As in the case of scientific reasoning, it is trivial that A's belief must be explained in terms of factors pertaining to A. But it does seem plausible to say that part of the explanation of her belief that the inference is invalid is that she has an intuitive grasp of the principles of correct deductive reasoning, an intuitive grasp which has been applied flawlessly in this case. As before, it is a further question whether this correct explanation is 'ultimate' and again the obvious answer is that it is not. There are deeper stories to be told: of how A came to grasp those correct principles and again, no doubt, of her capacity to grasp those principles, a capacity which itself may be explained in terms of her genetic and environmental history.

So, without denying that we could go deeper if we wished, it seems perfectly in order to say that an explanation (really an explanation sketch) of A's belief in the invalidity of I is that she recognized I's invalidity; while an explanation of B's belief in the validity of I is that he was 'blinded' to the real logical facts by his desire to believe the conclusion.

In what senses are these two explanations 'asymmetric'? Firstly, they are certainly not asymmetric in a sense several times objected to by Bloor—the above analysis does *not* amount to the claim that A's belief is 'self-explanatory' and only B's needs explanation. A's belief is explained (by her recognition of I's invalidity) just as much as B's belief is explained (by his desire to believe the conclusion). Bloor seems to have confused 'self-explanatory' and 'explained in a generalizable way'. It seems plausible that the 'same' explanation would apply to A's beliefs about a whole range of other inferences $I', I'', \ldots$ and also to a whole range of other agents' beliefs about inferences (including, no doubt, the beliefs of B about inferences whose conclusions are less emotionally charged for him). The explanation of B's attitude to the specific inference I is, on the other hand, more special to him and to this case. This specificity is clearly a matter of degree rather than kind, indeed the particular reason I envisaged for B's mistake has, of course, a certain generality: perhaps there is a general tendency to be 'softer' on inferences with appealling conclusions. However, suppose C also thinks incorrectly that I is valid and investigation reveals a more humdrum 'reason', perhaps that C was drunk at the time he was considering I. This explanation is not very likely to 'carry over' to D's belief in I's validity, which is instead likely to have a quite different source. In contrast, the explanation of A's belief in invalidity may well carry over to E's belief in invalidity, both simply having recognized that I is indeed invalid (though to repeat there is of course no necessity here: E *may* have the right belief for the wrong reason).

Secondly, to underline what was said in Section 2, there is no question of an asymmetry in this deductive case in the sense of the explanation of A's belief being dependent on my (or Russell's or Hilbert's) evaluation of I. The explanations are through A's (intuitive or implicit) grasp of the principles of deductive reasoning; as well as through factors operating on B to occlude his view of those principles.

Thirdly, and undoubtedly most important for Bloor and Barnes, is there an asymmetry here in the sense that only the explanation of B's belief is a real scientific explanation, while the explanation of A's belief appeals to abstract, non-physical facts—an appeal which is at odds with a truly scientific 'naturalistic' picture? Is the explanation of B's belief psychological or sociological or whatever *and therefore scientific*, while the explanation of A's belief is *rational* in a sense that precludes full scientific treatment?

Bloor does not explicitly address the deductive case I have been considering. However, he does insist that in the 'inductive' case the rationalist's explanations of true/reasonable beliefs and judgements, on the one hand, and of false/unreasonable beliefs, on the other, *are* asymmetric in this sense and are to be rejected on that ground. Brown, for one, holds that Bloor here sees a 'clash' between himself and the rationalist where none really exists:

> Suppose that by 'the same type of explanation', Bloor is merely advocating scientific (by which he means causal) explanations for all events: for bridges which are standing as well as fallen ones, for beliefs which are true/rational as well as false/irrational, and so on. . . . Well . . . both sides agree with this. The rationalist will cheerfully give 'the same type of explanation' for all beliefs if this just means giving causal explanations for them; though of course the rationalist will pointedly add: 'Some causes are reasons'. (1989, p. 39)

Some sociologists like to style themselves as doing 'anthropology of science', but, says Brown (1989, p. 32)

> Rationalists are anthropologists of science too; it is just that they have a different theory about what generally causes a scientist's behaviour, namely, they think that often good reasons and evidence rather than social forces are the determining influences.

However, it is Brown who misses what is indeed a real clash. Let me begin to explain why by first clarifying Bloor's position a little.

The main motivating claim behind Bloor's position, as I understand it, and the main (perhaps only) argument for the symmetry thesis is that *science must itself be studied scientifically*. Bloor actually says:

> If sociology could not be applied in a thorough-going way to scientific knowledge it would mean that science could not scientifically know itself. (1976, p. 40)

However, this contains an obvious mistake: science could presumably 'scientifically know itself' without invoking any irreducibly *social* factors at all and certainly without invoking *nothing but* social factors. All that follows if we accept that science must itself be studied scientifically is, as Bloor himself sometimes concedes, that the right approach is what used to be called 'science of science' or, more often nowadays, 'naturalized epistemology'. It is then a *further* (and, incidentally, entirely implausible) claim that the predominant 'natural' determinants of the development of science have been 'social' factors. Thus, Bloor the *sociologist* of science would no doubt be embarrassed by a successful explanation of our logical capacities and of the development of science in terms of, say, a hard-wired genetic predisposition to make appropriate deductive and inductive inferences, a predisposition that was notably insensitive to environmental influences right across the (known) board. However, Bloor's *argument* for symmetry, even if accepted, entitles him only to be a *scientist* of science and this envisaged genetic account would be perfectly in line with this naturalistic (but non-sociological) position. It follows that, even if Brown could show that scientists invariably accept conclusions arrived at by a single specified code of reason, this would be perfectly consistent with the core of Bloor's position here, *provided no claim was made about the correctness of this code.*

The rationalist (as here construed) *does*, however, make just such a claim about the correctness of the code of scientific reasoning: it is not just that most 'scientists' do, most of the time, draw conclusions in accord with the code the rationalist alleges he has identified, but the rationalist also asserts that scientists *ought* to reason in this way, that they are *correct* to do so. Thus, even if we eliminated from the code all the claims about 'social causation' of beliefs, a major disagreement would remain between Bloor's position and that of my rationalist. Bloor wants *evaluations*, as irreducible independent entities, out of the picture entirely. No doubt reasoners do evaluate arguments and no doubt scientists do evaluate theories in the light of evidence, but how they do so is itself determined by causal factors—the evaluations themselves, or rather their expressions, being then mere epiphenomena. The rationalist, on the other hand, insists that there are correct and incorrect evaluations and that an important reason why science has developed as it has is that scientists have often made the *correct* evaluations.

So with this point in mind let us turn back directly to the question of symmetry as seen by the (true) rationalist and her opponents. Again consider first the deductive case as likely to be clearer than the 'inductive' scientific one: is there an asymmetry in terms of scientific or causal character between the explanation of *A*'s (correct) belief in the invalidity of *I* and *B*'s (incorrect) belief in its validity?

As I see it, what the rationalist accepts and her naturalizing opponent denies is a world of logical facts over and above any psychological ones. Given these facts, the rationalist then sees a straightforward (and important) difference between *A* and *B*: *A*'s insight into this world of logical facts is, in the case of inference *I*, unclouded by inattention or contrary interests or the like, whereas *B*'s 'logical intuition' is thus clouded. I see no reason (though the point is of course controversial) why there should not be a good and full scientific (evolutionary) explanation of the development of logical intuition. Moreover there is undoubtedly a causal explanation (though we may have little interest in it) of how *A*'s logical intuition came to be unclouded in the particular circumstances at issue. Perhaps one would want to say that *A*'s insight into the realm of logical facts (alongside the fact that this insight was not occluded by other factors) *caused A* to believe that *I* was indeed invalid. (This would require adopting some fairly abstract account of causation.) So, on any construal, there is much of scientific interest and much that is causal to be said about *A* as well as about *B*. The overall explanation *in both* A*'s case and* B*'s case surely involves both 'causal' and logical factors*. It is not, as Bloor seems to hold, that the rationalist sees an asymmetry between a purely 'rational' or logical explanation in *A*'s case and a purely 'causal' explanation in the case of *B*. Instead the asymmetry (or difference) according to the rationalist is simply that the 'normal' logical–rational faculty operated 'normally' in the one case, while in the other, other factors interfered with that faculty's 'normal' operation.

Turning now from the case of deductive logic to the 'inductive', scientific case, the rationalist's claims are entirely analogous. The real issue between the rationalist and her 'naturalizing' opponents is that she holds there to be real (though non-physical) facts about which of a set of scientific theories is the best in the light of available evidence. The asymmetry in the explanation of, in our example, Airy's beliefs and those of Brewster is not that the former explanation is rational and the latter 'causal'. The distinctive rationalist claim is that both explanations may legitimately appeal to (methodological or epistemological) facts about the scientific standing of rival theories as well as to 'material' facts about Airy and Brewster. And the 'asymmetry' between the rationalist's explanation of the two cases is that Airy's insight concerning these facts was unclouded while Brewster's was, on the contrary, clouded by further factors. There may be much to say of a 'scientific' 'causal' kind about the development of this faculty; there is certainly a further explanation (though probably one without much interest) as to how Airy's methodological insight came to be unclouded in the particular circumstances in which his beliefs were formed. However, the difference between the two sets of beliefs remains.

If the rationalist's position is still unclear, it may help to consider the (close) analogy with the case of low-level perceptual judgements also analysed by Newton-Smith (1981, pp. 252–253). Consider Isabel and Icabod, who both (sincerely) believe they are seated on chairs; Isabel is indeed sitting on a chair, but Icabod is sitting on the floor. Would the explanations we give of these two beliefs be 'symmetric' or 'asymmetric'?

Newton-Smith argues for asymmetry, giving two accounts of why there is an asymmetry, accounts which he seems to think are equivalent but which are, I think, quite different. He *first* says that the asymmetry here rests on the fact that, in the case of Isabel's (correct) belief, the state of affairs which gives her belief its truth value (namely, her actually sitting on a chair) is part of the explanation of her belief, whereas in the case of Icabod's (incorrect) belief, the state of affairs which gives his belief its truth value (the absence of a chair, making his belief false) is no part of the explanation of his belief. But who can say without further details about Icabod? Suppose, for instance, that he is suffering from an extreme form of (of course, subconscious) perceptual countersuggestibility, so that he never believes anything at the perceptual level unless it is false. In that case the state of affairs which gives his belief its truth value is indeed part (and a crucial part) of the explanation of that belief. In general, it seems quite clear that there can be no *a priori* reason to deny the actual state of the world a causal role in the formation even of false perceptual beliefs.

However, Newton-Smith then goes on to characterize this case of perceptual beliefs as one in which

> if the belief is true, that it is believed is to be explained by causal mechanisms of normal perception [while] cases in which it is false are to be explained by some causal interference with normal perceptual mechanisms. (1981, p. 253)

This is surely different from the earlier characterization, since the result of the 'causal interference with normal perceptual mechanisms' may well (and in general will) also depend on the input from outside. This *second* characterization seems to me the right one and to provide the correct analogue for the rationalist's explanations of true/reasonable and false/unreasonable theoretical beliefs. The rationalist is assuming an intuitive grasp of logical and methodological facts on a par with sense perception of (low-level) physical facts. 'Normally' this intuitive grasp works well. This does not mean that there is no 'causal' story to be told in cases where it does work well, only that in such cases the specifics of the 'causal' story will be of a negative kind: a story of how other potentially obscuring factors in practice *failed* to interfere. Sometimes there is interference with this intuitive grasp, and then the interfering factors, likely to be different in different cases, need to identified and described.

Newton-Smith himself rejects the idea that the lessons learned from the case of low-level perceptual beliefs extend to the case of beliefs about scientific theories. He rejects the idea on the grounds that in the methodological case there are (or 'may be') no facts of the matter, as we assume there are in the case of the low-level perceptual beliefs. He goes on to give an account of those theory choices which are progressive in terms of our 'standing interest' in being rational—(presumably) in preferring those theories that are predictive, non-*ad hoc* and so on. He claims that it is a 'brute fact' that we have an interest in being rational:

> For the brute fact is the simple one that we have an interest in survival which brings with it an interest in following the dictates of reason. (1981, p. 257)

This then underlies the asymmetry seen by Newton-Smith: cases like that of Airy are explained by the 'standing interest' while cases like that of Brewster are explained by interests that interfere with this 'standing interest'.

Even if we accepted that we do have this standing interest, this would not be enough to supply a rationalist account as I understand it. It is not enough to allege that we have this interest as a matter of (descriptive) fact. After all, the creationists have the (for them) standing interest in defending the literal truth of the Bible. Is this just a clash of interests or is one side scientifically justified? A rationalist should give the latter answer, I do not see how Newton-Smith can.

If the answer were that the creationists' 'standing interest' is not one that is conducive to the survival of our species, whereas the preference for successfully predictive theories is thus conducive, then we need only ask how we 'know' this. Part of the answer is obviously that we 'know' that the theory of evolution is true or, more modestly (and accurately), that we know that this is the theory which at present is best supported by the evidence. But *how* do we know that? The answer clearly presupposes some prior account of evidential support.

Essentially this argument surely shows that *any* attempt to use the evolutionary version of naturalized epistemology to avoid relativism, while at the same time avoiding commitment to logico-epistemological truths, is doomed to failure. It is prefectly possible that the procedures for evaluating assertions about the universe adopted by many (though by no means all) humans have arisen as adaptive characteristics. However, anyone who believes that by asserting such an adaptive account a way is found of avoiding relativism while also avoiding commitment to anything irreducibly epistemological should ask themselves why they believe in such an account. The answer in part is clearly a belief in evolutionary theory. But is it simply a fact about them that they have this belief or is it justified by the evidence? If the former, then their position does not avoid relativism; if the latter then it does not avoid irreducible epistemology: for the canons by which evolutionary theory is justified by the evidence must be logically prior to the theory itself. It *may* be evolutionarily adaptive to believe in the theory of evolution and, in particular, to believe in some evolutionary account of our principles of theory appraisal (though I would need some convincing), but to think that any such 'self-explanation' could break the circle here is illusory. This method of explanation could succeed in justifying the correctness of our appraisal principles only if the correctness of the theory of evolution (and of this particular application of it) is assumed. But what underpins the judgement that the theory of evolution is correct?

The idea that there are, contrary to Newton-Smith's sceptical attitude, methodological facts of the matter—(irreducibly epistemological) truths about the evidential status of scientific theories at a given time—just as there are descriptive truths about the spatial positions of tables and chairs at a given time, may be difficult for hard-headed nominalistically inclined philosophers (and I plead 'guilty' here) to swallow. However, the alternative, no matter how it is dressed up, is relativism and that means facing up to digestive tasks that I for one find still more unpalatable: for example, swallowing the claim that orthodox and creationist biologists are just two different groups with different aims and norms and that there is no question of one of them being right ('right full stop': of course each is right and the other wrong according to its own standards).

Acknowledgements

I gratefully acknowledge a grant from STICERD (The Suntory-Toyota International Centre for Economics and Related Disciplines) which enabled me to take leave from my teaching duties from September 1989 to January 1990, and thus enabled me to do research, on part of which this paper is based.

Notes

1. Some of the details of Brewster's and Airy's positions are discussed in Worrall (1990) and rather more fully by Cantor (1983). My paper distinguishes three attitudes that a scientist might have to a theory, each of which might be called 'accepting' that theory: (a) believing the theory to be a completely true (or almost completely true) description of reality; (b) believing that working on trying to develop that theory is the likeliest way forward in the branch of science concerned; and (c) simply accepting that that theory is presently the one best supported by the facts without necessarily believing it to be true or even the best way forward. Each of these attitudes raises quite different issues of rationality. Here I try to avoid these subtleties and complications.
2. For the purposes of the present essay, I leave it vague what exactly is involved in one theory's ranking 'objectively higher' than another, given available evidence. For details of this complicated issue, the reader should refer to Worrall (1990).
3. Appeal to external factors need not, as Newton-Smith's second criticism shows, be inconsistent with a *mini*rational account. Indeed most of the standard examples cited, for example by Lakatos, of episodes requiring external explanation are easily construed as 'minirat': some Soviet scientists preferred Lysenkoism for reasons of political expediency, that is, because they aimed to avoid unpopularity with the state authorites, and no doubt the best means of achieving that aim in the circumstances was to prefer (or at any rate profess preference for) Lysenkoism over orthodox neo-Mendelian genetics. But Lakatos's account is all about maxi-rationality: this minirat account is precisely what Lakatos had in mind by 'handing over the problem to the sociologists', for here they need to explain how some scientists came to adopt this non-scientific aim rather than the 'normal' objective ones. (The explanation in this case is fairly obvious.)
4. The rationalist need not claim that even great scientists *always* applied her principles of appraisal, but only that they applied those principles in those parts of their work that are generally accepted to be successful. Thus, for example, the rationalist is under no obligation to claim that Newton applied her principles of rationality when working in alchemy or when developing his religious views. (I am grateful to Newton-Smith and Whyte for pointing to the need to clarify this point.)
5. I use this term in a very wide sense, roughly to cover any account of weighing empirical evidence for theoretical claims. It is not meant to endorse any particular approach to such weighing, say that of Carnap.
6. Whether or not Newton *did* exhibit false consciousness here is another matter and not one on which I am competent to pronounce.
7. See Worrall (1990) where the positions of Airy and his friend Baden Powell, on the one hand, and of Brewster, on the other, are treated in some detail with quotations and full references.

References

AIRY, G.B. (1831) *A Mathematical Tract on the Undulatory Theory of Light* (London, Macmillan).

BLOOR, D. (1976) *Knowledge and Social Imagery* (London, Routledge and Kegan Paul).

BREWSTER, D. (1831) 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.*

BROWN, J.R. (1989) *The Rational and the Social* (London, Routledge and Kegan Paul).

CANTOR, G. (1983) *Optics after Newton: theories of light in Britain and Ireland. 1704–1840* (Manchester, Manchester University Press).

LAKATOS, I. (1970) History of science and its rational reconstruction, reprinted in his *Methodology of Scientific Research Programmes, Philosophical Papers, Vol. 1* (Cambridge, Cambridge University Press, 1978).

LAUDAN, L. (1977) *Progress and its Problems* (Berkeley, University of California Press).

LAUDAN, L. (1984) *Science and Values* (Berkeley, University of California Press).

LAUDAN, L. (1989) If it ain't broke, don't fix it, *The British Journal for the Philosophy of Science*, 40, pp. 369–375.

NEWTON-SMITH, W.H. (1981) *The Rationality of Science* (London, Routledge and Kegan Paul).

WORRALL, J. (1976) Thomas Young and the "Refutation" of Newtonian optics, in: C. HOWSON (Ed.) *Method and Appraisal in the Physical Sciences* (Cambridge, Cambridge University Press).

WORRALL, J. (1988) The value of a fixed methodology, *The British Journal for the Philosophy of Science*, 39, pp. 263–275.

WORRALL, J. (1989) Fix it and be damned, *The British Journal for the Philosophy of Science*, 40, pp. 376–388.

WORRALL, J. (1990) Scientific revolutions and scientific rationality: the case of the "elderly hold-out", in: C. WADE SAVAGE (Ed.) *The Justification, Discovery and Evolution of Scientific Theories* (Minnesota, University of Minnesota Press).