REASON AND SCIENCE

Stathis PSILLOS (University of Athens, Greece)

Abstract. Among the many issues that relate to the role of Reason in science, I will focus my attention on two. The first concerns the problem of the justification of scientific method – and of induction in particular, which is the most basic and indispensable ampliative method of science. The second is related to the problem of theory-change in science: how can it be that theory-change is rational? In addressing these two issues (highlighting both their conceptual development and their present status), I will try to stress the need for a conception of method and rationality that leaves room for values.

1. Reason and Method

The rationality of science is typically associated with the scientific method and its justification. The very idea that the scientific method needs justification emerges from the fact that, whatever its detailed structure, it should satisfy two general and intuitively compelling desiderata: it should be ampliative *and* epistemically probative. Ampliation is necessary if the method is to deliver informative hypotheses and theories, viz., hypotheses and theories whose content exceeds the observations, and data that prompted them. Yet, this ampliation would be illusory, if the method was not epistemically probative; if, that is, the method did not convey epistemic warrant to the excess content thus produced (viz., hypotheses and theories). The philosophical problem of scientific method is that there are *prima facie* good reasons to think that these two plausible desiderata are not jointly satisfiable. The tension between them arises from the fact that ampliation does not carry its epistemically probative character on its sleeves. The following question then arises: what makes it the case that the method conveys epistemic warrant to its intended output? This is known as the problem of induction.

Hume's far-reaching point was that the alleged necessity and generality of causal claims cannot be proved *either* by means of demonstrative reasoning *or* by reason aided by experience. There can be no a priori demonstration of any general claim, since the cause can be conceived without its usual effect and conversely. Besides, any attempt to show, based on experience, that a regular pattern that has held in the past will or must continue to hold in the future, and hence any attempt to move from the particular to the general based on experience, will be circular and question-begging. It will presuppose a principle of uniformity of nature. But this principle is *not* a priori true. Nor can it be proved empirically without circularity. Any attempt to prove it empirically will have to assume what needs to be proved, viz., that since nature has been uniform in the past it *will* or *must* continue to be uniform in the future. This Humean challenge to any attempt to establish the necessity and generality of causal connections on either a priori or empirical grounds has become known as his scepticism about induction.

Faced with the dilemma “ampliation or justification,” Hume opted for jettisoning justification; better: he opted for a rejection of the traditional call for justification as a requirement for rationality. Hume should then be conceived as a proto-naturalist who took it that being natural and irresistible, inductive inferences are not under the legislative control of Reason and its quest for independent justification. They are governed by “custom,” which Hume took it to be “the great guide of human life.” The price, of course, is that generality (ampliation) can never be associated with necessity and certainty.

This kind of attitude was an anathema for rationalists – even an enlightened rationalist such as Kant. His distinctive attempt to ground Newtonian mechanics in a set of principles that were universal, necessary and certain was motivated by the thought that the very possibility of scientific knowledge required placing synthetic a priori restrictions on the set of models of the world that are consistent with experience. This distinctively epistemological a priori justification of (at least some) general principles came with a severe penalty: there cannot possibly be knowledge of the world *as it is in itself*. Hence, the special non-inductive method that can nonetheless yield unconditional generality and certainty can only apply to the phenomena, and this because its products (the general principles of pure science) are constitutive of them. More importantly, there is no guarantee that the world will co-operate with the synthetic a priori principles that make experience possible. Far from being unique, indispensable and unreviseable, the Kantian framework for science came to grief by developments in the formal and empirical sciences – notably the development of non-Euclidean geometries and the General Theory of Relativity.

The other major reaction to Hume was much friendlier, coming as it was from a fellow empiricist: John Stuart Mill. He was a thoroughgoing inductivist, who took all knowledge – even mathematical knowledge – to arise from experience through induction. Hence, Mill denied that there could be any certain and necessary knowledge. But, unlike Kant, Mill never thought there was a *problem* of induction. He took it that, being irresistible, induction did not need any justification: any attempt to rationally doubt it will end up in failure because we will keep on using it. What then is the scope of the scientific method? It should be such that it leads to secure – but not certain – knowledge of the world, where security is a function of the steps taken to eliminate error. The well-known Methods of Agreement and Difference where precisely meant to reduce the possibility of error, in light of certain substantive metaphysical assumptions, such that (a) events have a *limited* number of possible causes; and (b) same causes have same effects, and conversely.

Given the divergence between Kant and Mill, it might be an irony that the Kantian and the Millian approaches to method came together in the mature work of logical empiricists, especially Carnap. In a sense, Carnap borrows from Mill the claim that all substantive general knowledge of the world should be based on induction (while, of course, disagreeing with Mill's generalised inductivism – arithmetic and geometry are not inductive sciences). But he also borrows from Kant the thought that the justification of scientific method should be a priori (while, of course, disagreeing with Kant's view that this justification should be *synthetic* a priori).

Carnap, like many of his contemporaries, operated within a generally Fregean, anti-psychologist and anti-naturalist, intellectual milieu. What came to be known as *logic of science* was an attempt to capture the logical form of scientific method and to raise the issue of its justification within a logical context, where all that really matter are logical relations among propositions – those that express the evidence and those that express the theory. Given this, induction could be seen as a formal method with a definite logical form such that when certain evidence-statements are plugged in, a certain degree of probability is assigned to a general statement.

Despite the fact that probability was appealed to by Venn and Russell, it was John Maynard Keynes's (1921) employment of it in the characterisation of induction that set the stage for what was to follow. His key thought was that induction should be seen as a kind of logic: the logic of partial entailment of a hypothesis by relevant observational evidence. It should be developed into a formal system based on the probability calculus that aims to capture in a logical and quantitative way the notion of *inductive support* that evidence accrues to a hypothesis or theory. If this programme were successful, there would be an end to seeing induction as "a scandal of philosophy," as C.D. Broad put it. Its justification would be broadly logical, and hence unproblematic, because inductive logic was meant (a) to rely on logical principles, and (b) to mimic the content-insensitive structure of deductive logic. The traditional problem of induction was thereby supposed to give way to the well-defined task of devising confirmation functions, that is probability-functions which capture the logical relation between statements: those that express the evidence and those that express the hypothesis. The traditional dilemma "ampliation or justification" was meant to be dealt with swiftly: ampliation comes from induction, as traditionally understood, and justification comes from confirmation, seen as the logic of partial entailment.

This needs to be stressed: the traditional problem of induction – the problem of whether there can be reasonable acceptance of general truths either on the basis of reason or of experience – is split into two problems. The first – which, by and large, was taken to be beyond the ken of reason – has to do with the formation of general hypotheses; the second – which is amenable to reason-based treatment – has to do with the degree to which it is reasonable to accept a given general hypothesis in the light of given evidence. Thus seen, it is one thing to generate ampliative hypotheses, and quite another to confirm them, on the basis of experience. Traditionally induction was taken to be a method by means of which general propositions were generated and *accepted*. With probability coming into the picture, induction can no longer serve as a rule of acceptance. Instead, the rules of inductive logic were meant to specify the degree of credibility of an already given general hypothesis on the basis of already given evidence.

The thought, however, was that being the *logic of partial entailment*, inductive logic would capture the "rational degree of belief" that an agent should have in a hypothesis, given the evidence. Accordingly, given the evidence and given that different agents know it, they will attribute exactly the same probability to the relevant hypothesis. So inductive logic might not provide rules of acceptance, but – the thought was – it will provide *rational degrees of belief*. All this required what Keynes took it to be "logical intuition," which is such that anyone who possesses it can "see" the logical relation between the evidence and the hypothesis. This remnant of the Kantian intuition was as problematic as Kant's original and was criticised severely by Frank Ramsey. As he graphically put it: «I do not perceive them [the logical relations] and if I am to be persuaded that they exist it must be by argument» (Ramsey 1926: 63).

It was left to Carnap (1950) to resuscitate Keynes's programme while excising *intuition* from scientific method. In his own system of inductive logic, Carnap claimed, sentences expressing relations of partial entailment between the evidence and the hypothesis are analytic truths. And if inductive logic is analytic, it is also a priori – without the mystery of intuition. Carnap went as far as to claim that the contentious principle of uniformity of nature was also analytic *within* his system of inductive logic: it is a statement of logical probability asserting that «on the basis of the available evidence it is very *probable* that the degree of uniformity of the world is high» (Carnap 1950: 180-1). Carnap was certainly right in noting that, expressed as above, the principle of uniformity of nature could neither be proved nor refuted by experience. But of course, if anything, this principle is synthetic a priori and not analytic. It follows that if "inductive logic" had to rely on substantive *synthetic* principles, it was no longer *logic*.

Carnap hoped to devise certain quantitative functions that captured statements of the form: the degree of confirmation of H by e is r, where r is a real number between 0 and 1. He hoped he could thereby determine uniquely and quantitatively which of two competing hypotheses was more confirmed by some piece of evidence. To do this, he relied on a quasi-logical Principle of Indifference, which dictated that all equally possible outcomes should be given equal prior probabilities of happening. But different applications of this Principle lead to inconsistent results. For instance, one could start with the admission that all ways the world might be (what came to be called *state-descriptions*) are equally probable. But then it turns out that no evidence could raise the (posterior) probability of a state-description to more than what it was before the evidence rolled in. Alternatively, one could start with giving a bonus (higher) prior probability to some ways the world might be (in particular those ways in which certain universal regularities are present in the world). Then, it was shown that the evidence *does* raise the posterior probability of them, but now it was no longer the case that this relation of confirmation was independent of substantive assumptions as to how the world is likely to be.

Indeed, as was pointed out by Keynes long before Carnap, the very possibility of inductive inference requires that some hypotheses are given non-zero prior probability, for otherwise fresh evidence cannot raise their probability. And, if some hypotheses must be given finite non-zero prior probability, an infinite number of their rivals will have to be given zero prior probability – a priori! In the end, when Carnap (1952) devised the continuum of inductive method, he drew the conclusion that there can be a variety of actual inductive methods whose results and effectiveness vary in accordance to how one picks out the value of a certain parameter, where this parameter depends on formal features of the language used. But obviously, there is no *a priori* reason to select a particular value of the relevant parameter, and hence there is no explication of inductive inference in a unique way. Carnap suggested that it is left to the scientists to choose among different inductive methods, in view of their specific purposes. Where an a priori justification of induction was sought, the end product was based on a pragmatic decision.

The demise of the "logic" of induction left things where we started: the "scandal to philosophy" was there to stay. Already in his critique of Keynes, Ramsey took it that probabilities are not rational degrees of belief but *subjective* degrees of belief. Hence, there is no such thing as *the* rational degree of belief in the truth of a proposition; instead each individual is taken to (or allowed to) have her own subjective degree of belief in the truth of a certain proposition. Given that the probability calculus does not establish any non-trivial probability values, Ramsey argued that it was up to the agent to supply the initial probabilities. Then, the probability calculus, and Bayes's theorem in particular, can be used to compute values of other probabilities based on the prior probability distribution that the agent has chosen. The only requirement imposed on a set of degrees of beliefs is that they are probabilistically coherent, that is that they satisfy the axioms of probability.

The rationale for this claim is the so-called Dutch-book theorem. It is based on the significant mathematical result – proved by Ramsey (1926) and Bruno de Finnetti (1937) – that subjective degrees of beliefs (expressed as fair betting quotients) satisfy Kolmogorov's axioms for probability functions. The key idea is that unless the degrees of beliefs that an agent possesses, *at any given time*, satisfy the axioms of the probability calculus, she is subject to a Dutch-book, that is, to a set of synchronic bets such that they are all fair by her own lights, and yet, taken together, make her suffer a net loss come what may. The thrust of the Dutch-book theorem is that there is a *structural incoherence* in a system of degrees of belief that violate the axioms of the probability calculus. (For more on this see Skyrms 1984 and Howson 2000.)

The thought here is that the axioms of the probability calculus are an extension of deductive logic: the "logic of partial belief and inconclusive argument" as Ramsey put it. The demand for probabilistic coherence among one's degrees of belief is a *logical* demand: a demand for logical *consistency.* So it might be argued that the Dutch-Book theorem explains why we should strive for synchronic probabilistic coherence, if we are to be rational. This is certainly partly right, but as many Bayesians note, logic is about consistency and *not* about rational belief. In any case, the view that synchronic probabilistic coherence is a canon of rationality would require a non-question-begging *demonstration* that any violation of the axioms of the probability calculus is positively irrational. But no such proof is forthcoming.

Be that as it may, the demand for synchronic probabilistic coherence does not have anything to do (at least *prima facie*) with induction and confirmation. To accommodate the idea of learning from experience, Bayesians have tried to extend Bayesianism to belief-revision and belief-update by the technique of conditionalisation. It is supposed to be a canon of rationality (certainly a necessary condition for it) that agents should *update* their degrees of belief by conditionalising on the evidence:

$$\text{Prob}_{\text{new}}(\text{--})=\text{Prob}_{\text{old}}(\text{--}/e),$$

where *e* is the total evidence. (Conditionalisation can be either strict, where the probability of the learned evidence is unity, or Jeffrey – due to Richard Jeffrey – where the evidence one updates on can have probability less than 1.) The penalty for not conditionalising on the evidence is liability to a Dutch-book *strategy*: the agent can be offered a set of bets *over time* such that (a) each of them taken individually will seem fair to her at the time when it is offered; but (b) taken collectively, they lead her to suffer a net loss, come what may. (This is the Lewis-Teller argument, see Teller 1973.) But critics of Bayesianism point out that there is no general proof of the conditionalisation rule (see Earman 1992: 46-51). In fact, there are circumstances under which conditionalisation is an *inconsistent* strategy. When an agent is in a situation in which she contemplates about her Prob_{new}(--), she is in a *new* and different (betting) situation in which the previous constraints of Prob_{old} need not apply. A case like this is when the learning of the evidence *e* does upset the conditional probability Prob(--/e). Bayesianism has a point. Under *certain* circumstances, an agent should update her degrees of belief by conditionalising on the evidence. But it does *not* follow from this that Bayesian conditionalisation is a canon of rationality.

In so far as Bayesianism offers a theory of rationality (and as we have just noted this is by no means obvious), it offers a *structural conception of rationality*: rationality pertains to the structure of a belief system and not to its content. Hence, all that matters is how what you believe hangs together (at a certain time, or over time). According to the Bayesian structural conception of rationality, it is *not* irrational to maintain unjustified opinion. For subjective Bayesians, prior opinion can come from anywhere. And so can the prior probabilities. The standard (subjective) diachronic Bayesian picture is that people start with some prior opinion and then update it by conditionalising on the evidence. This is purely *logical* updating. It’s not ampliative. It does not introduce new content; nor does it modify the old one. It just assigns a new probability to the old opinion. In any case, it can be argued against the Bayesian view of rationality that the content of a belief matters to its rationality. For, to put it crudely, one could be a perfectly consistent Bayesian agent, even if one believed that the earth was flat. There seems to be nothing in Bayesianism that would render irrational an agent who neglected evidence which pointed to the roundness of earth in order to safeguard her belief that the earth is flat. In fact, a Bayesian agent could rationalise her attitude by giving zero prior probability to the hypothesis that the earth is round. More generally, it is entirely open to Bayesians to argue that some (perhaps all?) evidence against a certain belief can be neglected. (For more on this see Psillos 2007.)

The Bayesians' reliance on subjective prior probabilities has been a constant source of dissatisfaction among their critics. It is claimed that purely subjective prior probabilities fail to capture the all-important notion of rational or reasonable degrees of belief. In all fairness, it's been extremely difficult to articulate the notion of *rational* degree of belief. It may be argued that prior probabilities are informed by plausibility considerations, but this notion of "plausibility" resists further articulation. This has led Bayesians to insist on the indispensability of subjective priors in inductive reasoning. In fact, it can be proved that, in the long run, the prior probabilities wash out: even widely different prior probabilities will converge, in the limit, to the same posterior probability, if agents conditionalise on the same evidence. But this is little consolation because, apart from the fact that, as Keynes famously put it, in the long-run we are all dead, the convergence theorem holds only under limited and very well-defined circumstances that can hardly be met in ordinary scientific cases.

Despite all its merits (and its well-known successes), subjective Bayesianism cannot serve as a substitute for induction. There is still need for rules of acceptance – that is rules which entitle us to accept a general hypothesis on the basis of the evidence. After all, scientific theories tell us what statements to accept as expressing laws of nature and not what their degree of confirmation is. If this is recognised, the problem of the justification of these rules of acceptance seems to be alive and well.

We enter here a very controversial territory. But what I take it to be the right attitude to this problem is broadly pragmatic. Where it differs from standard pragmatism is that it is taken to amount to a conception of *justification* (of induction).

In "Truth and Probability," Ramsey summed up the attitude we already saw Mill advocating. Ramsey says:

> We are all convinced by inductive arguments, and our conviction is reasonable because the world is so constituted that inductive arguments lead on the whole to true opinions. We are not, therefore, able to help trusting induction, nor if we could help it do we see any reason why we should, because we believe it to be a reliable process (1926: 197).

This, Ramsey adds, is a pragmatic attitude, since induction is judged, ultimately, on the basis of its own success: that it leads to true beliefs more often than not and that it leads to true conclusions more often than other non-inductive methods do.

What's important, and what Ramsey suggests but does not develop, is that induction can be employed in its own justification without this circularity being vicious. Indeed, a kind of straightforward way to vindicate induction is by the following inductive argument:

> (I) Induction has worked in the past; therefore induction is likely to work in the future – and hence to be reliable.

A more exact formulation of this argument would use as premises lots of successful individual instances of induction and would conclude (by a meta-induction or a second-order induction) to the reliability of induction *simpliciter*. Arguments such as this have been employed by many philosophers, such as Braithwaite (1953), van Cleve (1984), Papineau (1992), Psillos (1999) and others.

To see their force, we must distinguish between two types of circularity. Let's call "premise-circular" an argument such that its conclusion is among its premises. This is a viciously circular argument. The charge of *vicious* circularity is an epistemic charge – a viciously circular argument has no epistemic force: it cannot offer reasons to believe its conclusion, since it presupposes it; hence, it cannot be persuasive. If the charge of circularity were logical and not epistemic, (if that is, a circular argument lacked validity altogether and not just epistemic force), all deductive arguments would be viciously circular. There is an obvious sense in which all deductive arguments are such that the conclusion is "contained" in the premises – this grounds/explains their logical validity.

Hence, deductive arguments can be circular without being *viciously* circular. And similarly, *some* deductive arguments are *viciously* circular, (without thereby being invalid). Premise-circularity is vicious! But (I) above (even in the rough formulation offered) is *not* premise-circular.

There is, however, another kind of circularity. This, as Braithwaite put it «is the circularity involved in the use of a principle of inference being justified by the truth of a proposition which can only be established by the use of the same principle of inference» (1953: 276). It can be called rule-circularity. In general, an argument has a number of premises $P_1,...,P_n$. *Qua* argument, it rests on (employs/uses) a rule of inference R, by virtue of which a certain conclusion Q follows. It may be that Q has a certain content: it asserts or implies something about the rule of inference R *used* in the argument; in particular that R is reliable. So: rule-circular arguments are such that the argument itself is an instance, or involves essentially an application, of the rule of inference whose reliability is asserted in the conclusion.

If anything, (I) is rule-circular. Is rule-circularity vicious? Obviously, rule circularity is *not* premise-circularity. But, one may wonder, is it still vicious in not having any epistemic force? To address this question, let us note that if one is not an inductive sceptic (that is, if one does not think that induction is not – or cannot be – justified), there are two options available when it comes to the issue of its justification: non-inferential justification and inferential justification. A non-inferential justification of induction, if possible at all, would have to rely on some a priori rational insight. An inferential justification of induction would have to rely on some rule of inference. But the very idea of a non-inferential justification presupposes something whose existence is dubious (rational insight). What about an inferential justification? If the rule is distinct, there is the issue of how the two rules are inferentially connected. If the rule is the self-same, we end up in rule-circularity.

The good news is that this is not a conceptual tangle that arises only in the case of induction. It arises already when it comes to the justification of deductive logic. In the case of the justification of *modus ponens* (or any other genuinely fundamental rule of logic), if logical scepticism is to be forfeited, there are two options available: either non-inferential justification or inferential (rule-circular) justification. There is no non-inferential justification of modus ponens. Therefore, there is only rule-circular justification. Indeed, any attempt to justify modus ponens by means of an argument has to employ modus ponens itself (see Dummett 1974). Why is there no non-inferential justification of modus ponens? There are many routes to this conclusion but a prominent one has to do with Quine's (1936) argument against basing logic on conventions.

But, one may wonder, couldn't *any* mode of reasoning (no matter how crazy or invalid) be justified by rule-circular arguments? A standard worry is that a rule-circular argument could be offered in defence of "counter-induction." This is supposed to move from the premise that "Most observed As are B" to the conclusion "The next A will be not-B." A "counter-inductivist" might support this rule by the following rule-circular argument: since most counter-inductions so far have failed, conclude, by counter-induction, that the next counter-induction will succeed.

The right reply here is that the employment of rule-circular arguments rests on or requires the absence of specific reasons to doubt the reliability of a rule of inference. We can call this, the *Fair-Treatment Principle*: a doxastic/inferential practice is innocent until proven guilty. This puts the onus on those who want to show guilt. The rationale for this principle is that justification has to start from somewhere and there is no other point to start apart from where we currently are, that is from our current beliefs and inferential practices. Accordingly, unless there are specific reasons to doubt the reliability of induction, there is no reason to forego its uses in justificatory arguments. Nor is there reason to search for an active justification of it. Things are obviously different with counter-induction, since there are plenty of reasons to doubt its reliability, the chief being that typically counter-induction have led to false conclusions.

It may be objected that we have no reasons to *rely* on certain inferential rules. But this is not quite so. Our basic inferential rules (including induction, of course) are rules we value. And we value them because they are *our* rules, that is rules we employ and reply upon to form beliefs. Part of the reason why we value these rules is that they have tended to generate true beliefs – hence we have some reason to think they are reliable, or at least more reliable than competing rules (say counter-induction).

Rule-circularity is endemic in any kind of attempt to justify basic method of inference and basic cognitive processes, such as perception and memory. In fact, as Ramsey noted, it is only via memory that we can examine the reliability of memory. Even if we were to carry out experiments to examine it, we would still have to rely on memory: we would have to *remember* their outcomes. But there is nothing vicious in using memory to determine and enhance the degree of accuracy of memory. For there is no reason to doubt its general reliability and have some reasons to trust it.

If epistemology is not to be paralysed, if inferential scepticism is not to be taken as the default reasonable position, we have to rely on rule-circular arguments for the justification of basic methods and cognitive processes.

There cannot be an absolute justification of induction – and of scientific method more generally. But unless we thought that the very idea of offering a justification of induction was meant to persuade the inductive sceptic, we can certainly live with this lack.

2. Reason and Judgement

A major challenge to the rationality of science has come from the process of theory-change and more particularly from scientific revolutions. In its most severe form, this challenge has been associated with Kuhn's claim that competing paradigms are incommesurable and that the ordinary canons of theory-appraisal break down during revolutionary transitions. In what follows, I will not discuss this matter in any detail, since I think the challenge can be raised even if we leave aside – as we should – the typically extravagant claims associated with incommensurability. The fact is that conceptual and theoretical change cannot be fully captured by standard approaches to scientific method. At stake here is the very idea of rational ways by means of which rival theories can be assessed and on the basis of which theory-choice can acquire rational force.

Note that Bayesian conceptions of scientific method face a rather acute problem, here: they cannot accommodate radical theory change, unless they allow violations of the mechanism of belief-updating by conditionalisation. This is what Marc Lange has aptly called «the problem of incorrigibility» (1999: 300). Bayesian agents are enslaved to their prior probability distribution: degree-of-belief updating merely alters the probability agents assign to the propositions that express their initial beliefs. This feature of Bayesianism is in conflict with the fact that during radical theory-change scientists revise or abandon some of their prior beliefs.

My own favourite, *Inference to the Best Explanation* (IBE), does not fair a lot better. The crux of IBE, no matter how exactly it is formulated, is that explanatory considerations should inform (perhaps, determine) what is reasonable to believe. An explanation should be such that it incorporates the explanandum into the rest of the reasoner's background belief-corpus by providing some link between the explanandum and other hypotheses that are part of the reasoner's background beliefs. Accordingly, explanation has a coherence-enhancing role. Explanatory coherence is a cognitive virtue because it is a prime way to confer justification on a belief or a corpus of beliefs. The warrant IBE confers on the chosen explanation is related to the fact that it fares better than its competitors in an explanatory-quality test and, as a result of this, it enhances the explanatory coherence of the belief corpus. All this might of course be contested (for its fuller defence see Psillos 2009). But the relevant point here is that, even if we were to grant all this, IBE cannot straightforwardly be applied to cases of revolutionary changes precisely because explanatory considerations are related to the relevant background beliefs, and it is these that are in doubt in a revolutionary episode.

All is not lost, however, and this because the very idea of a good explanation is subject to some structural standards, which, at least partly, fix explanatory merit and mark the explanatory power of a hypothesis. These standards operate crucially when the substantive information contained in the relevant background knowledge cannot forcefully discriminate between competing potential explanations of the evidence. They also operate when background beliefs are themselves at stake. Kuhn himself argued that there are some important traits that characterise a "good scientific theory": accuracy, consistency, broad scope, simplicity and fruitfulness (cf. Kuhn 1977: 321-322). But he also thought that though these *values* are trans-paradigm, they cannot be the basis of a rational adjudication between competing paradigms, because how they are applied and interpreted differs in different paradigms. No doubt, Kuhn was right in stressing that there could not be an algorithmic and value-free account of scientific method. Insofar as the previous rational reconstructions of science had aimed to equate rational judgement in science with the application of an algorithmic method, Kuhn was right in his criticism of them. But why should one accept this equation in the first place? And certainly, it does not follow from Kuhn's own criticism that the only alternative is to accept that scientific judgement is an irrational and unconstrained-by-evidence-and-reason activity. The Kuhnian values, after all, should be seen as part and parcel of sound and rational scientific judgement.

In support of this claim, let us take a leaf from Duhem's *Aim and Structure of Physical Theory* (1906). Perhaps more than anyone else, Duhem felt the fundamental tension between the strict conception of scientific method that he himself had advocated and the need for a broader conception of rational judgment in science. Despite that at first blush, he equated scientific method with experience+logic, he went on to argue that there is space for rational judgements in science which is *not* captured by the slogan: scientific method=experience+logic. What's important here is that experience plus logic are not enough even to decide when a theory should be abandoned (or modified).

This problem relates to the well-known Duhem(-Quine) thesis and doubly so. On the one hand, Duhem himself recognised what Poincaré made famous by saying that though experience does not, strictly speaking, contradict a theory, yet it can *condemn* it. On the other hand, experience and logic cannot dictate how to revise theories in the face of recalcitrant evidence. Duhem is well-known for his view that crucial experiments are «impossible in physics» (Duhem 1906: 188). Take a crucial experiment to be an experiment that would prove one theory wrong – one that would strictly contradict the theory. If a situation such as this is not possible, how do theories get abandoned? Obviously, any answer would have to go beyond the strict limits of experience and logic. And Duhem's own answer does. He employed other criteria of assessment. Here are some that he suggests: the scope of the theory, the number of hypotheses, the nature of hypotheses, novel predictions (1906: 28, 195), compatibility with other theories (1906: 221, 255), unification into a single system of hypotheses (1906: 293). These are, of course, the usual suspects. They are *values* or *virtues* of a theory that transcend logic (or, at least, they defy a rigorous logical formulation). What Duhem saw clearly was that the employment of such criteria is (a) indispensable, and (b) not algorithmic. Their employment requires the exercise of *judgement*.

Though Duhem was pessimistic (and rightly so) about the prospects of a fully articulated theory of rational judgement in science, he was clear that (a) there is space for rational judgements; and (b) this judgement exceeds the (perhaps artificial) limits posed on scientific method by a pure and strict adherence to experience plus logic (including probability theory, I would add). He classified this judgemental character of scientific method under the rubric "good sense." This is something that scientists acquire in their training and practice. It is something that can be cultivated and sharpened. It is also something that is akin to common sense. But no matter what the details are, good sense permits scientists to "decide" among competing hypotheses. And though «the reasons of good sense do not impose themselves with the same implacable rigour as the prescriptions of logic do» (Duhem 1906: 217), they are still *reasons*. Duhem was a great admirer of Pascal. So, for him these reasons are «reasons which reason does not know» (1906: 217).

But who said that judgement should *not* be part of the scientific method? An extreme positivistic understanding of scientific method, encapsulated in the fiction of Carnap's robot, as a fully-determined-by-exact-rules algorithmic process which delivers yes-no answers (or exact degrees of confirmation) for each hypothesis formulated in a formal language, is not just a chimera. It is, in addition, a model that does not bear any resemblance to whatever happens in science. Does it follow from the judgemental character of the method that the criteria that guide these judgements are unreasonable? Unworkable? Subjective? Opaque? None of the above follows.

I think we should take to heart Duhem's lesson. Scientific method is not algorithmic. It requires the exercise of judgement. This judgement is constrained by several virtues that theories should possess. It can be rational even if it is not dictated by experience plus logic. Its rationality depends, ultimately, on taking account of the *reasons* that favour a certain option and condemn another. But we should also go beyond Duhem. A *theory* of rational judgement is not doomed from the start. It cannot be reduced just to an ineffable “good sense.” A good starting point for such a theory is to focus on the role that explanatory considerations and the virtues of theories play in rational judgement – a point that Duhem himself made available.

This need for an account of rational judgement which goes beyond logic plus experience has been articulated recently by Ernan McMullin. As he aptly notes: «Values do not function in assessment as rules do» (1996: 19). It’s not just that different scientists may weigh different values (or virtues) differently. This, as Kuhn has already noted, is true enough. But it is also true that even if they are weighed similarly, they may be in conflict with each other (say, simplicity vs. informativeness). Hence, judgement is required in balancing them out. Values are unlike rules in that they cannot be followed algorithmically. No recipe is there for choosing among competing theories. It would be too quick, however, to conclude from this that these values have no rational force. This would amount to intellectual paralysis. For take the prime empiricist virtue (and don’t forget that it is a *virtue* too): *empirical fit*. Of course, theories should be consistent with the evidence (or entail it). But judging empirical fit is no (much) less value-laden than judging, say, explanatory fit. It’s not just that many competing theories can be consistent with the same observations. It’s also that the very empirical fit of a theory to facts requires judgement: Which are the relevant data? Which measurements are reliable? What error-margins are allowed? Etc.

3. Concluding Thoughts

The prospects of offering a *unified* theory of the role of reason in science are tied to the prospects of telling a story that brings under a single chapter the role of method in theory-appraisal and the role of method in theory-change. As we have seen, the very idea of a rational way to deal with theory-change in science – especially with revolutionary theory change – requires that the traditional conception of method be broadened with an account of values and virtues. These can create a space for rational judgement where algorithmic approaches break down. But these have to be *shared* values and virtues; traits of theories that scientists – *qua* cognitive agents – value and strive for their theories to have them. We also saw, however, that values enter into the very idea of offering a justification to induction – the very basic scientific method. In the absence of a way to persuade the inductive sceptic, looking for a justification of induction amounts to looking for warrants for a rule (or a method) that we already *value*; a method that we already use successfully and have no specific reasons to doubt. There is need to take seriously the task of developing an account of method and rational judgement in science which places values in centre-stage.

References

Braithwaite R.B. 1953. *Scientific Explanation*. Cambridge: Cambridge University Press.

Carnap R. 1950. *Logical Foundations of Probability*. Chicago: The University of Chicago Press.

Carnap R. 1952. *The Continuum of Inductive Methods*. Chicago: The University of Chicago Press.

De Finetti B. 1937. Foresight: Its Logical Laws, Its Subjective Sources. In H.E. Kyburg Jr. and H.E. Smokler (eds.), *Studies in Subjective Probability*, 53-118. New York: John Wiley and Sons, 1964.

Duhem P. 1906. *The Aim and Structure of Physical Theory*, 2nd ed. 1914, P. Wiener transl. Princeton: Princeton University Press, 1954.

Dummett M. 1974. *The Justification of Deduction*. Oxford: Oxford University Press.

Earman J. 1992, *Bayes or Bust?* Cambridge, MA: MIT Press.

Howson C. 2000. *Hume's Problem*. New York: Oxford University Press.

Keynes J.M. 1921. *A Treatise on Probability*. Cambridge: The Royal Economic Society.

Kuhn T.S. 1977. Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension.* Chicago: The University of Chicago Press.

Lange M. 1999. Calibration and the Epistemological Role of Bayesian Conditionalisation. *Journal of Philosophy*, 96: 294-324.

McMullin E. 1996. Epistemic Virtue and Theory Appraisal. In I. Douven and L. Horsten (eds.), *Realism in the Sciences*, 13-34. Louven: Louven University Press.

Papineau D. 1992. Reliabilism, Induction and Scepticism. *Philosophical Quarterly*, 42: 1-19.

Psillos S. 1999. *Scientific Realism: How Science Tracks Truth*. London and New York: Routledge.

Psillos S. 2007. Putting a Bridle on Irrationality: An Appraisal of Van Fraassen's New Epistemology. In B. Monton (ed.), *Images of Empiricism*, 134-164. Oxford: Oxford University Press.

Psillos S. 2009. *Knowing the Structure of Nature*. London: MacMillan-Palgrave.

Quine W.V.O. 1936. Truth by Convention. In O.H. Lee (ed.) *Philosophical Essays for A.N. Whitehead*, 90-124. New York: Longmans.

Ramsey F. 1926. Truth and Probability. In R.B. Braithwaite (ed.), *The Foundations of Mathematics and Other Essays*. London: Routledge, 1931.

Skyrms B. 1984. *Pragmatics and Empiricism*. New Haven and London: Yale University Press.

Teller P. 1973. Conditionalisation and Observation. *Synthese*, 26: 218-58.

Van Cleve J. 1984. Reliability, Justification and the Problem of Induction. *Midwest Studies in Philosophy,* 9: 555-567.