5 Resisting the pessimistic induction

The explanationist defence of realism (EDR) has suffered a rather serious blow from Laudan's contention that the history of science itself destroys the credibility of realist explanation of the success of science. For it is full of theories which were once empirically successful and yet turned out to be false. Laudan's argument[1] against scientific realism is simple but powerful. It can be summarised as follows:

> The history of science is full of theories which at different times and for long periods had been empirically successful, and yet were shown to be false in the deep-structure claims they made about the world. It is similarly full of theoretical terms featuring in successful theories which do not refer. Therefore, by a simple (meta-)induction on scientific theories, our current successful theories are likely to be false (or, at any rate, are more likely to be false than true), and many or most of the theoretical terms featuring in them will turn out to be non-referential.
>
> Therefore, the empirical success of a theory provides no warrant for the claim that the theory is approximately true. There is no substantive retention at the theoretical, or deep-structural, level and no referential stability in theory-change.

Laudan has substantiated his argument by means of what he has called 'the historical gambit': the list that follows – which, Laudan says, 'could be extended *ad nauseam*' – gives theories which were once empirically successful and fruitful, yet were neither referential nor true. These theories were *just* false:

- the crystalline spheres of ancient and medieval astronomy
- the humoral theory of medicine
- the effluvial theory of static electricity
- catastrophist geology, with its commitment to a universal (Noachian) deluge

- the phlogiston theory of chemistry
- the caloric theory of heat
- the vibratory theory of heat
- the vital-force theory of physiology
- the theory of circular inertia
- theories of spontaneous generation
- the contact-action gravitational ether of Fatio and LeSage
- the optical ether
- the electromagnetic ether.

If Laudan is right, then the realist's explanation of the success of science flies in the face of the history of science: the history of science cannot possibly warrant the realist belief that currently successful theories are approximately true, at least insofar as the warrant for this belief is the 'no miracle' argument. In what follows, I analyse the structure of Laudan's argument and show how scientific realism can be defended.

Laudan's *reductio*

The 'pessimistic induction' is a kind of *reductio*. The target is the realist thesis that:

(A) Currently successful theories are approximately true.

Laudan does not directly deny that currently successful theories may *happen* to be truth-like. His argument aims to discredit the claim that there is an *explanatory connection* between empirical success and truth-likeness which warrants the realist's assertion (A). In order to achieve this, the argument compares a number of past theories to current ones and claims:

(B) If currently successful theories are truth-like, then past theories *cannot* have been.

Past theories are deemed not to have been truth-like because the entities they posited are no longer believed to exist and/or because the laws and mechanisms they postulated are not part of our current theoretical description of the world. Then, comes the 'historical gambit':

(C) These characteristically false theories were, nonetheless, empirically successful.

So, empirical success is not connected with truth-likeness and truth-likeness cannot explain success: the realist's potential warrant for (A) is defeated. As Laudan put it:

> Because they [most past theories] have been based on what we now believe to be fundamentally mistaken theoretical models and structures, the realist cannot possibly hope to explain the empirical success such theories enjoyed in terms of the truth-likeness of their constituent theoretical claims.
>
> (1984a: 91–92)

Hence, the pessimistic induction 'calls into question the realist's warrant for assuming that *today's* theories, including even those which have passed an impressive array of tests, can thereby warrantedly be taken to be (in Sellars' apt image) 'cutting the world at its joints' (Laudan 1984b: 157).

No realist can deny that Laudan's argument has *some* force. It shows that, on inductive grounds, the whole truth and nothing but the truth is unlikely to be had in science. That is, all scientific theories are likely to turn out to be, strictly speaking, false. This is something that realists seem to have to concede. However, a false theory can still be *approximately true*. The notion of approximate truth is discussed in detail in Chapter 11. For the time being, let me note that a theory is approximately true if it describes a world which is similar to the actual world in its most central or relevant features. So, what realists need to show is that past successful theories, although strictly speaking false, have been approximately true. This is the defensive line in which realists regroup and start their counter-attack.

Laudan's immediate challenge is that a theory cannot be said to be approximately true unless it is shown that its central terms refer (1981: 33). This requirement seems plausible. But one should be careful here. The intended realist claim is that from the genuine empirical success of a theory one can legitimately infer that the entities posited by the theory are real – they inhabit the world we live in. Without this assumption we cannot adequately explain the empirical success of a theory. There is, however, no way in which any proponents can 'step outside' of their theories and check whether these entities exist. We should simply have to rely on our theories as our best guide to what the furniture of the world is. What Laudan observes is that, given the past track-record of science, we simply cannot do that: the radical changes in the central ontological claims made by theories over the centuries suggest that any such claim is as likely to go as any other. None of them, in other words, enjoys any privilege over any other. Mary Hesse has put the same thought in the form of the 'principle of no privilege', which, she says, follows from an 'induction from the history of science'. According to this principle, 'our own scientific theories are held to be as much subject to radical conceptual change as past theories are seen to be' (1976: 264). In order to rebut the 'principle of no privilege', realists should show that:

1 the theoretical discontinuities in theory-change were neither as widespread nor as radical as Laudan has suggested;

2 instead, there has emerged a rather stable and well-supported network of theoretical assertions and posits which is our best account of what the world is like; and
3 theoretical terms that can be legitimately taken to have been central in past theories can still be referential, i.e. they can still be taken to refer to entities which feature in science's current theoretical ontology.

In sum, realists should try to reconcile the historical record with the realist claim that successful theories are typically approximately true. How can this be done?

Realist gambits

Before discussing this, let me make two preliminary points. First, one should note that scientists are not prone to acquire only false beliefs. As science progresses, they accumulate more evidence, further and fresh empirical data, which they can then use to update and modify their beliefs and theoretical commitments. Besides, scientists can come to know how to better test their theories and, in particular, how to identify those methods of theory-construction which are likely to generate false and unwarranted beliefs. Hence, they can form better-supported theoretical beliefs. They can learn how to gauge the requisite evidence for their beliefs, how to improve their methods, and how to avoid unreliable methods. There is no guarantee, of course, that this process of learning from past experience will lead from false to truer theories. However, if scientists can positively learn from past experience, they are in a better position to abandon false theoretical claims in favour of new ones that are better supported by the evidence. Hence, these claims have a better chance of being truth-like than did those now abandoned. Second, even a quick glance at current science suggests that there is a host of entities, laws, processes and mechanisms posited by past theories – such as the gene, the atom, kinetic energy, the chemical bond, the electromagnetic field etc. – which have survived a number of revolutions to be retained in current theories. That is, one can quickly see that Laudan has overstated his case against scientific realism. In its crudest form, the pessimistic induction boils down to the claim that, as science grows, we can certify only the accumulated theoretical falsehoods, while we invariably have no good reasons to believe that we have hit upon some theoretical truths. But this is far-fetched and implausible.

Success too-easy-to-get

It is now time to attempt a conclusive refutation of Laudan's *reductio*. In light of the structure of his argument outlined earlier, one way to block Laudan's *reductio* is to target the 'historical gambit' or premiss (C). One can substantially weaken premiss (C) simply by reducing the size of Laudan's

list. If we manage to restrict the meta-inductive basis, it no longer warrants the conclusion that genuine success and approximate truth are unconnected. Therefore, the 'historical gambit' is neutralised.

The form of Laudan's 'historical gambit' is this. It claims that all past theoretical conceptualisations of the several domains of inquiry $T_1, \ldots, T_n$ Laudan has sampled have been empirically successful yet false, and it concludes, inductively, that *any* arbitrarily successful scientific theory T_{n+1} is likely to be false (or, at any rate, more likely to be false than true).

This kind of argument can be challenged by observing that the inductive basis is not big and representative enough to warrant the pessimistic conclusion (cf. Devitt 1984: 161–162; McMullin 1984: 17). The basis for Laudan's induction can be eroded by querying whether all of the listed theories were, as a matter of fact, successful and whether they were representative of their disciplines at stages of development sufficiently advanced as to be reckoned theoretically mature.

One can dispute the claim that all theories in Laudan's list were successful. Laudan suggests that a theory is successful 'so long as it has worked reasonably well, that is, so long as it has functioned in a variety of explanatory contexts, has led to several confirmed predictions, and has been of broad explanatory scope' (1984a: 110). To be sure, he thinks that this is precisely the sense in which realists claim scientific theories to be successful when they propose the 'no miracle' argument (ibid.). However, the notion of empirical success should be *more* rigorous than simply getting the facts right, or telling a story that fits the facts. For any theory (and for that matter, any wild speculation) can be made to fit the facts – and hence to be successful – by simply 'writing' the right kind of empirical consequences into it. The notion of empirical success that realists are happy with is such that it includes the generation of novel predictions which are in principle testable.[2] Consequently, it is not at all clear that all theories in Laudan's list were genuinely successful. It is doubtful, for instance, that the contact-action gravitational ether theories of LeSage and Hartley, the crystalline spheres theory and the theory of circular inertia enjoyed any genuine success (cf. McMullin 1987: 70; Worrall 1994: 335). A realist simply would not endorse their inclusion in Laudan's list. On the contrary, the real question for a realist is this: are theories which were *genuinely* successful characteristically false?

Given the centrality of novel predictions in my defence of realism, it is prudent to analyse this notion a bit further so that it becomes clearer and certain misunderstandings are avoided. A 'novel' prediction is typically taken to be the prediction of a phenomenon whose existence is ascertained only *after* a theory suggests its existence. On this view a prediction counts as novel only if the predicted phenomenon is *temporally* novel, that is, only if the predicted phenomenon was hitherto unknown. This, however, cannot be the whole story. For one, theories also get support from their ability to explain already known phenomena. For another, why should the provenance

of the predicted phenomenon have any bearing on whether or not the prediction supports the theory? One can easily imagine a case in which, unbeknown to the theoretician whose theory made the prediction of a temporally novel phenomenon, the phenomenon had already been discovered by some experimenter. Would or should this information affect the support which the predicted fact confers on the theory? If we thought that *only* genuine temporally novel predictions can confer support on theories, then we would have to admit that once we were aware that the fact was known, the predicted fact would become impotent to support the theory. In order to avoid these counter-intuitive pitfalls, the notion of novelty should be broader than what is meant by 'temporal novelty'. Following Earman (1992: Chapter 4, section 8) we should speak of 'use novelty', where, simply put, the prediction P of a known fact is use-novel *relative to a theory* T, if no information about this phenomenon was used in the construction of the theory which predicted it.[3]

But how exactly are we to understand the claim that a theory T makes a use-novel prediction of a known phenomenon? I think that in order to appreciate the issue at stake, one must follow Worrall (1985; 1989c) and provide some analysis of the ways in which a known fact E can be accommodated in a scientific theory T. Generally, there are two such ways:

- Information about a known fact E is used in the construction of a theory T, and T predicts E.
- A phenomenon E is known the time that a theory T is proposed, T predicts E, but no information about E is used in the construction of T.

Tidal phenomena, for instance, were predicted by Newton's theory, but they were not used in its construction. Let me, then, call *novel accommodation* any case in which a known fact is accommodated within the scope of a scientific theory, but no information about it is used in its construction. Let me, moreover, contrast novel accommodation with *ad hoc accommodation*. Although the Lakatosian school has produced a fine-grained distinction between levels of ad hocness, (cf. Lakatos, 1968: 399; 1970: 175; Zahar, 1973: 101), I shall take the most general case, namely:

> *Conditions of ad hocness*: A theory T is ad hoc with respect to phenomenon E if and only if either of the following two conditions is satisfied:
>
> 1 A body of background knowledge B entails the existence of phenomenon E. Information about E is used in the construction of a theory T, and T accommodates E.
> 2 A body of background knowledge B entails the existence of phenomenon E. A certain already available theory T does not predict/explain E. T is modified into theory T' so that T' predicts E, but the *only* reason for this modification is the prediction/explanation of E. In particular T' has no other excess theoretical and empirical content over T.[4]

Given this analysis, novel accommodation (or use novelty) of known facts can be explicated as follows:

> *Use novelty*: A prediction P of a phenomenon E is use-novel with respect to a theory T if E is known before T is proposed, T does not satisfy either of the ad hocness conditions and T predicts E.

The real issue then is whether use novelty and temporal novelty have different bearings on the empirical support of a theory. I do not want to enter here the subtleties of this debate, for my purpose is to contrast novel accommodation with ad hoc accommodation. But, briefly, my view is that both use novelty and temporal novelty, so long as they are sharply distinguished from any ad hoc accommodation, are *complementary* aspects of theory confirmation. For, one can demand that a theory should accommodate known phenomena in a *non* ad hoc way, and *in addition* to this that it must yield temporally novel predictions. When, however, it comes to the *support* that use-novel and temporally novel predictions confer on a theory, that is, when it comes to the degree to which they confirm a theory, we may well assign different weights to these two sorts of prediction. It is natural to suggest that any temporally novel predictions which obtain carry an *additional* weight, because a theory that suggests new phenomena takes an extra risk of refutation. For there is always the possibility that a known fact can be 'forced' into a theory, whereas a theory cannot be forced to yield an hitherto unknown fact. Hence, predicting a new effect – whose existence falls naturally out of a theory – makes the theory more risky and susceptible to extra experimental scrutiny which may refute it.[5]

In sum, I want to stress that it is important *not* to contrast use novelty and temporal novelty, but both are to be contrasted with ad hoc accommodation. For, if anything, there is at most a difference in *degree* between use novelty and temporal novelty, whereas, there is a difference in *kind* between novel accommodation and ad hoc accommodation.[6]

Besides making the notion of empirical success more rigorous, another way to reduce the size Laudan's list is to suggest that *not all* past theoretical conceptualisations of domains of inquiry should be taken seriously. Realists require that Laudan's list should include only *mature* theories; that is, theories which have passed the 'take-off point' (Boyd) of a specific discipline. This 'take-off point' can be characterised by the presence of a body of well-entrenched background beliefs about the domain of inquiry which, in effect, delineate the boundaries of that domain, inform theoretical research and constrain the proposal of theories and hypotheses. This corpus of beliefs gives a broad identity to the discipline by being, normally, the common

ground that rival theories of the phenomena under investigation share. It is an empirical matter to find out when a discipline reaches the 'take-off point', but for most disciplines there is such a point (or, rather a period). For instance, in the case of heat phenomena, the period of theoretical maturity was reached when such background beliefs as the principle of impossibility of perpetual motion, the principle that heat flows only from a warm to a cold body and the laws of Newtonian mechanics had become well entrenched. If this requirement of maturity is taken into account, then theories such as the 'humoral theory of medicine' or the 'effluvial theory of static electricity' drop out of Laudan's list. Once Laudan's list is restricted to those past theories which were *mature and genuinely successful*, then it is no longer strong enough to warrant the pessimistic conclusion.

Although it is correct that realists should not worry about all of the past theories that Laudan suggests, the present move is not enough to defeat the 'pessimistic induction': for it does not account for the fact that at least *some* past theories which pass both realist tests of maturity and success are nevertheless considered false. Relevant examples are the caloric theory of heat and the nineteenth-century optical ether theories. If these theories are false, despite their being both distinctly successful and mature, then the intended explanatory connection between empirical success and truth-likeness is still undermined. How then can we defend this explanatory connection?

The *divide et impera* move

The crucial premiss in Laudan's *reductio* is (B) (see p. 102): if we hold current theories to be truth-like, then past theories are bound not to be truth-like since they posited entities that are no longer believed to exist, and posited laws and theoretical mechanisms that have now been abandoned. Without this premiss the pessimistic conclusion does not follow.

Can we defeat (B)? Here is a suggestion: it is enough to show that the success of past theories did not depend on what we now believe to be fundamentally flawed theoretical claims. Put positively, it is enough to show that the theoretical laws and mechanisms which generated the successes of past theories have been retained in our current scientific image. I shall call this the *divide et impera* move. It is based on the claim that when a theory is abandoned, its theoretical constituents, i.e. the theoretical mechanisms and laws it posited, should not be rejected *en bloc*. Some of those theoretical constituents are inconsistent with what we now accept, and therefore they have to be rejected. But not all are. Some of them have been retained as essential constituents of subsequent theories. The *divide et impera* move suggests that if it turns out that the theoretical constituents that were responsible for the empirical success of otherwise abandoned theories are those that have been retained in our current scientific image, then a substantive version of scientific realism can still be defended.

This move dissociates genuine empirical success from characteristic falsity. Moreover, it paves the way for the 'right kind' of explanatory connection between success and truth-likeness. Laudan, realists should say, has taught us something important: on pain of being at odds with the historical record, the empirical success of a theory cannot issue an unqualified warrant for the truth-likeness of everything that the theory says. Insofar as older realists have taken this view, they have been shown to be, to say the least, unrealistic. Yet, it would be equally implausible to claim that, despite its genuine success, everything that the theory says is wrong. The right assertion seems to be that the genuine empirical success of a theory does make it reasonable to believe that the theory has *truth-like constituent theoretical claims*.

Moreover, if the theoretical constituents that were responsible for the empirical successes of past theories have been retained in subsequent theories, then this gives us reason to be more optimistic about their truth-likeness: that all these theoretical constituents have been shown to be invariant and stable elements of our modern scientific image; they have survived several 'revolutions' and have contributed to the empirical success of science. I think realists should follow Philip Kitcher's lead (1993) and suggest that the best way to defend realism is to use the generation of stable and invariant elements in our evolving scientific image to support the view that these elements represent our best bet for what theoretical mechanisms and laws there are.

This preamble for the *divide et impera* move may resonate with two recent reactions to the 'pessimistic induction', those of Kitcher (1993) and of Worrall (1989; 1994). Both have defended the analogous view that realists should characterise which kinds of statement are abandoned as false and which are retained. Kitcher suggests a distinction between 'presuppositional posits' and 'working posits', while Worrall draws the line between the 'content' of a theoretical statement, which gets superseded, and its 'structure', which is retained. The position I defend is akin to Kitcher's, although some differences will be discussed shortly. However, the *divide et impera* move is not meant to reflect or capture Worrall's distinction between structure and content. The latter distinction and Worrall's position deserve a more detailed discussion and criticism, to which Chapter 7 is devoted.

How should realists circumscribe the truth-like constituents of past genuinely successful theories? I must first emphasise that we should really focus on the specific successes of certain theories, like the prediction by Fresnel's theory of diffraction that if an opaque disk intercepts the rays emitted by a light source, a bright spot will appear at the centre of its shadow; or Laplace's prediction of the law of propagation of sound in air by means of the hypothesis that sound's propagation is an adiabatic process. Then we should ask the question: how were these successes brought about? In particular, which theoretical constituents made essential contributions to them? It is not, generally, the case that *no* theoretical constituents contribute

to a theory's successes. Similarly, it is not, generally, the case that *all* theoretical constituents contribute (or contribute equally) to the empirical success of a theory. (What, for instance, was the relevant contribution of Newton's claim that the centre of mass of the universe is at absolute rest?) Theoretical constituents which make essential contributions to successes are those that have an indispensable role in their generation. They are those which 'really fuel the derivation' – to use one of Laudan and Leplin's recent expressions (1991: 462).

When does a theoretical constituent H indispensably contribute to the generation of, say, a successful prediction? Suppose that H together with another set of hypotheses H' (and some auxiliaries A) entail a prediction P. H indispensably contributes to the generation of P if H' and A alone cannot yield P and no other available hypothesis H^* which is consistent with H' and A can replace H without loss in the relevant derivation of P. Clearly, there are senses in which all theoretical assertions are eliminable, if, for instance, we take the Craig-transform of a theory, or if we 'cook up' a hypothesis H^* by writing P into it. But if we impose some natural epistemic constraints on the potential replacement – if, for instance, we require that the replacement be independently motivated, non ad hoc, potentially explanatory, etc. – then it is not certain at all that a suitable replacement can always be found. Worrall has recently noted that whenever a theory is replaced by another, 'the replacing theory alone offers a constructive proof of the "eliminability" of the earlier one' (1994: 339). There should be no doubt that the old theory as a whole gets eliminated. Yet, Worrall's observation does not establish the eliminability of the specific theoretical constituents that contributed to the empirical successes of the superseded theory. If the *divide et impera* move is correct, then these constituents are typically those that 'carry over' to the successor theory (admittedly, sometimes, only as limiting cases of the relevant constituents of the replacing theory).

So, when it comes to explaining the specific successes of a theory by means of the claim that the theory has truth-like constituent theoretical claims, realists should argue that the truth-like constituents are (more likely to be) those that contribute essentially to, or 'fuel', these successes. Realists need care only about those constituents which contribute to successes and which can, therefore, be used to account for these successes, or their lack thereof. Analogously, the theoretical constituents to which realists need not commit themselves are precisely those that are 'idle' components, impotent to make any difference to the theory's stake for empirical success.

What is required to successfully perform the *divide et impera* move? The key to this question lies in the careful study of the structure and content of past genuinely successful theories. What is needed are careful case-studies that will

- identify the theoretical constituents of past genuine successful theories that made essential contributions to their successes; and

- show that these constituents, far from being characteristically false, have been retained in subsequent theories of the same domain.

If all kinds of claims that are inconsistent with what we now accept were essential to the derivation of novel predictions and in the well-founded explanations of phenomena, then one cannot possibly appeal to their truth-likeness in order to explain empirical success. Then, Laudan wins. However, if it turns out that the theoretical constituents which were essential are those that have 'carried over' to subsequent theories, then the 'pessimistic induction' gets blocked. Settling this issue requires detailed study of some past theories that qualify as genuinely successful.

The good news for realism, as we shall see in detail in the next chapter, is that relevant studies of the several stages of the caloric theory of heat and the nineteenth-century optical ether theories suggest that both of the foregoing requirements can be met. However, as regards the *general* argument thus far, the details of these studies – illuminating though they may be – are not necessary. This argument has aimed to show that if realists successfully perform the two tasks outlined above, then a case can be made for scientific realism; it has also indicated how these tasks can be performed, in particular, what role the suggested case-studies are to play, what issues they should focus on and how they are relevant to settling the argument between scientific realism and the 'pessimistic induction'.

Is the *divide et impera* move perhaps too close to Kitcher's approach? Could one not simply identify the idle constituents of a theory with Kitcher's 'presuppositional posits' and the essentially contributing constituents with his 'working posits'? These identifications may be pertinent. However, there are differences. My distinction between idle and essentially contributing constituents is meant to capture how the successes of a theory can differently support its several theoretical constituents. Kitcher's distinction between presuppositional and working posits, however, is meant to capture the difference between referring and non-referring terms. Working posits are said to be 'the putative referents of terms that occur in problem-solving schemata', while presuppositional posits are 'those entities that apparently have to exist if the instances of the schemata are to be true' (Kitcher 1993: 149). But, so put, the distinction is problematic. For, in effect, we are told that the success of a problem-solving schema does support the existence of the referents of some of the terms featuring in it, but it does not support the existence of a putative entity the presence of which is required for the truth of the whole schema. But unless one shows how it is possible that the empirical success of the theory can lend support only to some, but not all, existence claims issued by the theory, then Kitcher's contention seems to be just grist to Laudan's mill. Kitcher suggests that the putative referents of presuppositional posits, such as the ether, were apparently only presupposed for the truth of the relevant schemata; in fact, they turned out to be eliminable without derivational loss (1993: 145). This suggestion is

retroactive and open to the charge that it is ad hoc: the eliminable posits are those that get abandoned. Yet, as we are about to see, the *divide et impera* move can improve on Kitcher's views by avoiding this charge.[7]

A central objection to my line thus far is the following: with the benefit of hindsight, one can rather easily work it out so that the theoretical constituents that supposedly contributed to the success of past theories turn out to be those which were, as it happens, retained in subsequent theories. So, the realists face the charge that they are bound to first identify the past constituents which have been retained and then proclaim that it was those (and only those) which contributed to the empirical success and which enjoyed evidential support. Can realists do better than that? Retention aside, can we independently identify the theoretical constituents that contribute to the successes of a given theory and show that it is only those that we deem truth-like?

In response to this objection, it should be pointed out that eminent scientists do the required identification all the time. It is not that realists come, as it were, from the future to identify the theoretical constituents of past theories that were responsible for their success. Scientists themselves tend to identify the constituents which they think were responsible for the success of their theories, and this is reflected in their attitude towards their own theories. This attitude is not an all-or-nothing affair. As we are about to see in some detail, scientists do not, normally, believe either that everything a successful theory says is truth-like or conversely that, despite its success, nothing it says is truth-like. Rather, the likes of Lavoisier, Laplace and Carnot – to mention just a few – had a differentiated attitude towards their theories (in this case the caloric theory), in that they believed in the truth-likeness of some theoretical claims while considering some others to have been too speculative, or too little supported by the evidence, to be accepted as truth-like. This differentiated attitude was guided by the manner in which the several constituents of the theory were employed in the derivation of predictions (e.g. Laplace's prediction of the correct law of the propagation of sound in air) and in well-founded explanations of phenomena (e.g. Carnot's explanation of the fact that maximum work is produced in a Carnot-cycle). So, theoretical claims which were not essential for the success of the theory were treated with suspicion, as for instance was the case with the assumption that heat is a material fluid; and those claims which 'fuelled' the successes of the theory were taken to enjoy evidential support and were believed to be truth-like, as for instance was the case with the claims that heat can remain in latent form, or that the propagation of sound in air is an adiabatic – rather than an isothermal – process.

My claim is that it is precisely those theoretical constituents which scientists themselves believed to contribute to the successes of their theories (and hence to be supported by the evidence) that tend to get retained in theory change. Whereas, the constituents that do not 'carry-over' tend to be those that scientists themselves considered too speculative and unsupported to be

taken seriously. If this view is right, then not only is the *divide et impera* move not ad hoc, but it actually gains independent plausibility from the way scientists treat their theories, and from the way they differentiate their commitments to their several constituent theoretical claims. If, therefore, there is a lesson which scientists should teach realists it is that an all-or-nothing realism is not worth fighting for.

In the next chapter, I try to substantiate these general philosophical points by means of two detailed case-studies. They concern the two controversial items on Laudan's list: the caloric theory of heat and the optical ether theories of the nineteenth century. Let me here just summarise the main points that these studies will raise and defend.

The study of the *caloric theory of heat* shows that the caloric representation of the cause of heat as a material fluid was not as central, unquestioned and supported as, for instance, Laudan (1984a: 113) has claimed. Caloric was not a putative entity to which the most eminent scientists had committed themselves as the real causal agent of heat phenomena. More importantly, the empirical success of the caloric theory was not essentially dependent on claims concerning the existence of an imponderable fluid which caused the rise (fall) of temperature by being absorbed (given away) by a body. The laws which scientists considered well supported by the available evidence and the background assumptions they used in their theoretical derivation were *independent* of the hypothesis that the cause of heat was a material substance: no relevant assumption was essentially used in the derivation–prediction of these laws. So, the laws which scientists considered to be well supported by the evidence and to generate the empirical success of the caloric theory did not support, nor did they require, the hypothesis that the cause of heat was a material substance. What this study suggests is that the parts of caloric theory which scientists believed in were well supported by the evidence and were retained in subsequent theories of heat, whereas the hypotheses that were abandoned were those which were ill-supported by the evidence. Hence, the point which the first case-study will highlight is this: when the laws established by a theory turn out to be independent of assumptions associated with allegedly central theoretical entities, it makes perfect sense to talk of the approximate truth of this theory, despite the recognition that not all of its theoretical terms refer.

The second case-study – which discusses the *dynamical optical ether theories* of the nineteenth century – aims to offer a different service to realism. It suggests that the most general theory – in terms of Lagrangian dynamics and the satisfaction of the principle of the conservation of energy – which was the backbone of the research programme around the dynamical behaviour of the carrier of light-waves has been retained in the subsequent framework of electromagnetism. This general theory was employed in the study of the *luminiferous ether* which was taken to be the dynamical structure which underlies light-propagation and which was such that it sustained the light-waves, and stored their energy (*vis viva*), during the time between

their leaving the source and until just before reaching the receiver. Given that the carrier of light-waves was a dynamical structure of unknown constitution, the application of Lagrangian dynamics to study its behaviour enabled the scientific community to investigate its most general properties (e.g. its general laws of motion) leaving out the details of its constitution. The investigation of the possible constitution of the carrier of light-waves was aided by the construction of models (e.g. Green's elastic-solid model of the ether), where this model construction was based on perceived analogies between the carrier of light-waves (e.g. its ability to sustain transversal waves) and other physical systems (e.g. elastic solids). It was mostly these models that were abandoned later on. This case-study will show that a reading of the nineteenth-century theories of optics which suggests that the content of these theories was exhausted by the elastic solid-like models confuses the model and the actual, yet concealed, dynamical system the behaviour of which scientists were trying to understand. The advocates of the pessimistic induction would simply make an illegitimate move, if they appealed to those past failed models which scientists took to be heuristic devices, in order to infer that any current or future physical theory is likely to be false.

One of the points that the second study raises relates to the status of the abandoned theoretical term 'luminiferous ether'. It is hard to deny that the postulation of a medium for the propagation of light – denoted by the term 'ether' – underwrote the development of optical theories during the nineteenth century. Yet, the term 'ether' has been seen as an exemplar of a non-referring scientific term. Does it, then, follow that the whole range of dynamical theories of optics in which ether had a central function cannot possibly be approximately true? Discussion of that issue is postponed until Chapter 12, where attention turns to theories of the reference of theoretical terms. There I motivate a causal–descriptive theory of reference and defend the view that it is plausible to think of 'luminiferous ether' as referring to the electromagnetic field.

6 Historical illustrations

THE CALORIC THEORY OF HEAT

Heat as an imponderable fluid or heat as motion?

The core problems of the theories of heat in the late eighteenth and the early nineteenth century were the following: the cause of the rise and fall in the temperature of bodies; the cause of the expansion of gases when heated; the change of state; and the cause of the release of heat in several chemical interactions, and especially in combustion. It was in this problem-nexus that scientists such as Joseph Black, Antoine Lavoisier and Pierre-Simon Laplace introduced the causal–explanatory model of caloric.

Caloric was taken to be a theoretical entity and 'caloric' was the theoretical term purporting to refer to a material substance, an indestructible fluid of fine particles, which causes the rise in temperature of a body which absorbs it (cf. Lavoisier 1790: 1–2). Heat was taken to be the observable effect of the transportation of caloric from a hot body to a cold one (ibid.: 5). Being a material substance, caloric was taken to be conserved in all thermal processes. In 1780s, Lavoisier used caloric as an important element in his anti-phlogiston system of chemistry (ibid.: Part I; also Lilley 1948). Moreover, the assumption that heat was conserved played an important role in the development and theoretical exploitation of experimental calorimetry (see Laplace and Lavoisier 1780: 156). In dealing with the change in the state of a substance (e.g. the vaporisation of water), where, although a large quantity of heat is needed, this change takes place at constant temperature, Black (1803) assumed that heat can exist in a latent form, too. Lavoisier had already suggested that caloric can exist in two forms: either free (*calorique sensible*) or combined. Combined caloric was thought to be 'fixed in bodies by affinity or electric attraction, so as to form part of the substance of the body, even part of its solidity' (1790: 19). So, the existence of latent heat was explained by means of caloric in combined form.

However, a dynamical conception of heat had been the rival of the caloric theory ever since the latter was put forward. According to the dynamical theory, the cause of heat was not a material fluid. Instead, it was the motion

of the particles which constitute a substance. So, heat was taken to be nothing over and above the result of the motion of the molecules of a body. Laplace and Lavoisier give the following account of the dynamical theory: '[H]eat is nothing but the result of the insensible motions of the molecules of matter. . . . According to the hypothesis we examine [i.e. the dynamical theory] the heat is the *vis viva* (*force vive*) which is the result of the insensible motions of the molecules of bodies' (1780: 151–152).

The dynamical representation of the cause of heat was less developed than the caloric theory. But, it could also explain the transmission of heat and the restoration of equilibrium between unequally heated bodies put in contact (ibid.: 152 and 154). Most proponents of the caloric theory considered the dynamical theory as a serious but, given the available evidence, less probable competitor (see Black 1803: 44). The main reason why the dynamical account attracted the attention of scientists was that it could explain the production of heat by friction. Davy (1799: 9–23) listed a series of experiments which constituted, as he said, a *reductio ad absurdum* of the thesis that heat was a material substance, since matter could not be produced or created by motion, that is, for instance, by rubbing two things together. So, this empirical fact was taken to undermine the claim that the cause of heat was a material substance which was never created or destroyed. Count Rumford (Benjamin Thomson) (1798: 70) took up Davy's misgivings against the caloric theory and performed several experiments in which heat was produced by friction. He also suggested that the cause of heat could not be a material substance since heat could be produced by friction in an *inexhaustible* manner, and no material substance can be inexhaustible. On the contrary, he said, if heat was motion, as the advocates of the dynamical theory suggested, then its generation by friction would be easily explained.

Most caloricists, however, were unmoved by Count Rumford's challenge because, after all, only a finite quantity of heat could ever be obtained before the bodies used for the production of heat by friction were rubbed away. Hence, their claim was that the production of heat by friction could not be inexhaustible. Besides, the dynamical representation of heat was physically and mathematically undeveloped and did not attract any significant attention until Clausius and William Thomson showed that this representation is compatible with the Carnot–Clapeyron theory of work and the basic laws of the caloric theory.

Yet, the caloric representation of heat was not without problems. Probably its most important difficulty was related to the problem of the weight of caloric. According to both the critics and the advocates of the theory, if caloric were a material substance, then it should have mass and weight. Up to 1785, all experiments performed had shown that a heated substance did not weigh more than when it was unheated. The absence of weight from caloric was an important problem for the caloric theory. Reviewing several experiments, Black (1803: 45) stated:

> It has not, therefore, been proved by any experiment that the weight of bodies is increased by their being heated, or by the presence of heat in them. This may be thought very inconsistent with the idea of the nature or cause of heat that I . . . mentioned [i.e. that the cause of heat is a material fluid]. It must be confessed that the afore-mentioned fact may be stated as a strong objection against this supposition [i.e. that the cause of heat is a material fluid].

Starting from 1787 and lasting until late 1790s, Count Rumford performed a series of experiments in order to calculate 'the weight ascribed to heat'. Rumford examined whether liquids change in weight when they lose heat by just cooling down. The results obtained were negative. So he concluded that the caloric theory could not explain away the absence of weight from caloric, unless it assumed that caloric 'is so infinitely rare, even in its most condensed state, as to baffle all our attempts to discover its gravity'. On the contrary, he argued, if one adopted the theory that 'heat is nothing more than the intestine vibratory motion of the constituent parts of heated bodies', then it would be clear that 'the weight of bodies can in no wise be affected by such a motion' (1799: 100). So, whereas the caloric theory had to perform an artificial manoeuvre in order to accommodate the absence of weight from caloric, the competing dynamical theory could accommodate this fact more naturally.

Does the superiority of the caloric representation of heat at this early stage suggest that scientists believed that the caloric theory was true? What I will show is that most of the eminent supporters of the theory were very cautious in expressing their attitude to the *epistemic value* of the theory. Let us consider the following points:

1. Most of the eminent proponents of the caloric theory were aware of the difficulties that this theory faced.
2. They knew the advantages of the alternative representation of heat, especially in explaining the production of heat by friction.
3. They were aware also of the shaky experimental evidence, and of the inaccuracy of most of the experimental results available.

Such factors made most of the eminent scientists working within the caloric theory of heat to be very careful in their statements and very cautious in their epistemic claims. Probably the example most illustrative of this behaviour concerns Black. In his lectures, Black presented *both* contemporary theories of heat. He emphasised moreover that '(O)ur knowledge of heat is not brought to the state of perfection that might enable us to propose with confidence a theory of heat or to assign an immediate cause of it' (1803: 42). He noted that 'the supposition' that heat was a material fluid appeared the 'most probable', but he added that 'neither of these suppositions [i.e. the material and the dynamical] has been fully and accurately

considered by their authors, or applied to explain *the whole facts and phenomena* related to heat. They have not, therefore, supplied us with a proper *theory* or *explication* of the nature of heat'.

Black was cautious in his attitude towards the caloric theory, in fact towards both theories of heat available at his time, because neither could adequately explain *all* the then-known phenomena of heat. He went on to say that most of the ways that caloricists followed in order to develop their theories in the light of recalcitrant experience were ad hoc. Black gives the following excellent account of ad hoc modifications:

> Many have been the speculations and views of ingenious men about this union of bodies with heat. But, as they are all hypothetical, and as the hypothesis is of the most complicated nature, being in fact a hypothetical application of another hypothesis, I cannot hope for much useful information by attending to it. *A nice adaptation of conditions will make almost any hypothesis agree with the phenomena.* This will please the imagination, but does not advance our knowledge.
>
> (1803: 46)

This attitude towards the hypothesis that the cause of heat is a material substance, which amounted to a suspension of judgement until better evidence came in, was not just Black's idiosyncratic behaviour. After presenting both theories, Laplace and Lavoisier also suggested that the theory of experimental calorimetry was independent of the considerations concerning the cause of heat. Here is their own account:

> We will not decide at all between the two foregoing hypotheses [material v. dynamical theory of heat]. Several phenomena seem favourable to the second [the dynamical theory of heat], such as the heat produced by the friction of two solid bodies, for example; but there are others which are explained more simply by the other [material theory of heat] – perhaps they both hold at the same time. So . . . one must admit their common principles: that is to say, in either of those, *the quantity of free heat remains always the same in simple mixtures of bodies. . . .* The conservation of the free heat, in simple mixtures of bodies, is, then, independent of those hypotheses about the nature of heat; this is generally admitted by the physicists, and we shall adopt it in the following researches.
>
> (1780: 152–153)

Their account suggests two things: on the one hand, the principle of conservation of heat was not adhered to because it was a consequence of the claim that the cause of heat is a material substance, but rather because it was taken to be a *theoretical generalisation* stemming from the experiments in calorimetry. On the other hand, since calorimetric laws were independent

of considerations about the cause of heat, they could not be used to test either of the theories of the cause of heat.

Lavoisier repeated his reservations about the caloric representation of heat in his monumental *Traite Élémentaire de Chimie* (1789). Although in this work he put forward the material theory of heat as a candidate for the cause of heat phenomena, he was careful to qualify his commitments: 'Strictly speaking, we are not obliged to suppose this to be a real substance; it being sufficient, as will more clearly appear in the sequel of this work, that it is considered as the repulsive cause, whatever that may be, which separates the particles of matter from each other' (1790: 5).

What follows from all this is that the scientists of this period were not committed to the truth of the hypothesis that the cause of heat was a material substance. Therefore, caloric was not as central a posit as, for instance, Laudan has suggested (1984a: 113). Equivalently, the theoretical attempt to discover the cause of heat did not revolve around the unquestioned belief that caloric was the wanted cause. Most scientists' cautious attitude was the product of some important methodological considerations:

1. The caloric theory faced anomalies which could not be explained easily.
2. An alternative theory was available, which could account for some of the anomalies that the caloric theory faced.
3. The hypothesis that the cause of heat was a material substance was not essentially and ineliminably involved in the derivation and explanation of the laws of calorimetry.
4. The modifications to which the caloric theory was subjected in order to overcome some anomalies were rather artificial and ad hoc.
5. Most of the work in experimental calorimetry was conducted independently of any theory of heat.

However, it would be wrong to infer that the scientists' attitude towards the caloric theory was instrumentalist. Rather, using current philosophical terminology, I would claim that: *semantically*, the scientific community's attitude towards the theory was realist. 'Caloric' was a putative referring term which stood for a material fluid whose transportation from one body to another caused changes in temperature. *Epistemically*, the scientists' attitude was one of cautious and differentiated belief. Their epistemic attitude was not an all-or-nothing matter, but rather was determined by the evidence which supported the several theoretical constituents of the theory.

Laplace's prediction of the speed of sound in air

One of the most notably successful predictions attributed to the caloric theory is Laplace's prediction of the speed of sound in air. In 1816 Laplace published a memoir in which he suggested that the transmission of sound takes place in an adiabatic way, thereby correctly predicting the speed of

sound. This was an amazing success, for Laplace corrected Newton's calculation of the speed of sound in air. Unlike Newton, who had assumed that the expansions and contractions of a gas, as sound passes through it, take place isothermally, Laplace suggested that the propagation of sound was an adiabatic process. He assumed that there was some quantity of latent heat which was released from the compression of the air. This quantity of heat is normally diffused in the gas. But, for Laplace, 'since this diffusion takes place very slowly relative to the velocity of the vibrations, we may suppose without sensible error that during the period of a single vibration the quantity of heat remains same between two neighbouring molecules' (1816: 181). He then approximated sound-propagation by an isothermal compression of the gas and followed by heating the gas at constant volume.

Laplace suggested that Newton had failed to appreciate the effect of the second process on the pressure (or elasticity) of the gas. For Laplace 'it is clear that the second cause [heating the gas at constant volume] should increase the velocity of sound since it increases the elasticity of the air' (ibid.). He was then able to show that the speed of sound is represented by the formula

$$v^2 = (c_p/c_v)\, dP/d\rho,$$

where c_p is the specific heat of air under constant pressure, c_v is the specific heat under constant volume, P is the pressure and ρ the density of air.[1] The result obtained was 345.18 m/sec. Laplace attributed the difference from the experimental value to 'the uncertainty in experimental measurements' (cf. 1816: 181). In fact, he was right, since he took γ ($= c_p/c_v$) = 1.5 based on the quite off-the-mark calculations by Delaroche and Berard.[2]

Was this successful and novel prediction in any way dependent on the hypothesis that heat is a material substance? Laplace's account does not explicitly rest on any particular representation of heat, although he happened to be an advocate of the caloric theory. It is also noteworthy that Laplace's explanation of the propagation of sound in terms of an adiabatic process is essentially correct and has been retained in the subsequent theoretical accounts of heat.

In 1823, Poisson established by theoretical means the general law which governs adiabatic processes, that is, PV^γ = constant, where γ is the ratio of the two specific heats of a gas under a certain temperature (cf. 1823: 328–329). Here again, however, this law was shown to be *independent* of any specific hypothesis about the cause of heat. To be sure, Poisson did rest his derivation on the hypothesis that the quantity of heat absorbed or released by a body is a *state function* of three macroscopic properties of the body – pressure P, temperature T, and volume V. And, it is worth observing, the assumption that the quantity of heat involved in a process is a state function of the macroscopic parameters (pressure, temperature and volume) should be taken as the fundamental hypothesis of the mature caloric theory. For if such a function of heat did exist, it would follow that, in a complete cycle from (V_1, T_1) back to (V_1, T_1), the quantity of heat absorbed was equal to the quantity of heat released, irrespective of the way that the changes took place; that is, it would follow that heat was a conservative quantity. After Clausius's work in thermodynamics, it was recognised that heat is not a state function of the macroscopic properties of a gas. On the contrary, the quantity of heat released or absorbed by a body depends on how the process happens. More specifically, when work is produced in a thermal cycle, the quantity of heat involved in this cycle does not uniquely depend on the initial and final states in which the substance undergoing the changes is found. As a result, heat is not conserved in all thermal processes.

However, Poisson's derivation of the theoretical law of adiabatic change, is approximately correct. For although heat is not a function of the state of a gas, one can approximate infinitesimal changes in the quantity of heat of a gas, such as those occurring in an adiabatic process, by the method employed by Poisson, that is by analysing an infinitesimal change in heat in terms of the partial derivatives of two macroscopic parameters (cf. Fermi 1936: 20, 21–26). So, although in the advanced caloric theory the hypothesis that the cause of heat is a material substance was made concrete by the assumption that heat can be mathematically represented as a state function, Laplace's account of the propagation of sound did not depend on this hypothesis. Moreover, Poisson's theoretical derivation of the law of adiabatic change was approximately correct despite the use made by the derivation of the mathematical representation of heat as a state function.

Carnot and caloric[3]

Let me now move on to discuss the role of the caloric theory in Carnot's work. Carnot devotes his 'Reflections on the Motive Power of Fire' to the theoretical study of the work which can be produced by a gas undergoing specific changes so that it returns to its initial state (i.e. it traverses a complete – and reversible – thermal cycle).

In his theoretical account of the motive power of heat, it seems as though Carnot had accepted the principle of the conservation of heat and the existence of a state-function. For instance, he wrote (although in a footnote of his text) that '(t)his fact [i.e. the conservation of heat] has never been called in question. It was first admitted without reflection, and verified afterwards in many cases by experiment with the calorimeter. To deny it would be to overthrow the whole theory of heat to which it serves as a basis' (1824: 19/76).[4]

However, Carnot was also aware of the difficulties faced by the hypothesis that heat is conserved in any process whatsoever. Even in his published paper, he questioned the soundness of the supposed central axiom of the caloric theory. He remarked:

> The fundamental law [i.e. that heat was a state function] which we proposed to confirm seems to us however to require new verifications in order to be placed beyond doubt. It is based on the theory of heat as it is understood today, and it should be said that this foundation does not appear to be of unquestionable solidity. New experiments alone can decide the question. Meanwhile, we can apply the theoretical ideas expressed above, *regarding them as exact,* to the examination of different methods proposed up to now for the realisation of the motive power of heat.
>
> (1824: 46/100–101; emphasis added)[5]

Concerning the motive power of heat, Carnot stated that the work produced in a steam engine was due to the *redistribution of caloric* among the parts of the engine. So, he took it to be the case that the steam produced in the boiler of an engine was used to transport caloric to the condenser, thereby producing mechanical work, without any quantity of heat being consumed in this process. The hypothesis that heat is a material substance entailed this thesis: if caloric was a substance, then it had to be indestructible; then it could produce work in a heat engine without being consumed, but by its mere redistribution.

However, Carnot was very careful not to employ the hypothesis of conservation of heat. In order to support this claim let us look at the demonstration of the theorems relating to the well-known Carnot's cycle. Carnot considers two bodies A and B kept at different, but constant, temperatures, T_1 and T_2 respectively, where $T_1 > T_2$ (see Figure 6.1). The working substance is a gas contained in a tank *abcd*, the top side, *cd*, of which is movable with a piston. Carnot studied a process which consisted of four steps (Carnot 1824: 17–19/74–76):

1. The gas is brought in contact with body A, at the constant temperature T_1, and is slowly left to expand, at a constant temperature T_1, to the position *ef* (i.e. isothermal expansion from V_1 to V_2).
2. Body A, then, is removed from the gas, and the latter is left to expand from the position *ef* to the position *gh*, where its temperature becomes equal to that of the body B, i.e. T_2 (i.e. adiabatic expansion from T_1 to T_2).
3. Then, the gas is brought in contact with body B, at a constant temperature T_2, and is compressed from *gh* to *cd*, at a constant temperature T_2 (i.e. isothermal compression from V_2 to V_1).
4. Body B is removed, and the gas is compressed from *cd* to *ik*, its final temperature being again T_1. Finally, the gas is brought to its initial state *ab* by contact with the body A (i.e. adiabatic compression from T_2 to T_1).

The process can be repeated indefinitely, by repeating the four steps in the same order.

Using his cycle, Carnot demonstrates the following propositions:

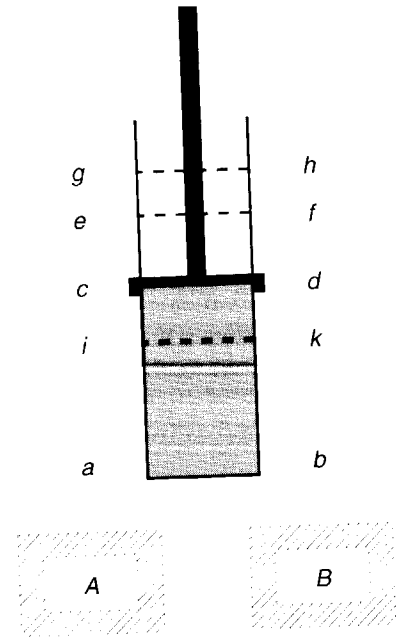

Figure 6.1 Carnot's cycle

Source: Adapted from Carnot 1824

(1) *The maximum quantity of work can be produced when and only when a substance undergoes transformations in a Carnot cycle* (see 1824: 19/76).

The demonstration of this theorem is most interesting since Carnot appeals to well-established independent background knowledge. Suppose, Carnot says, that more work W' is produced in some cycle C' than the amount of work W produced in a Carnot cycle C. Were this so, it would be possible to create perpetual motion. For one could first subject the substance to the transformations of cycle C', *then* direct the excess motive power $W'-W$ from the condenser (the cold body) to the boiler (the hot body), and finally subject the substance to the transformations of the Carnot cycle C. But '. . . this would be not only perpetual motion, but an unlimited creation of motive power without consumption of either caloric or of any agent whatever. Such a creation is entirely contrary to ideas now accepted, to the laws of mechanics and of sound physics' (1824: 12/69). So, Carnot establishes that $W' - W$ must be negative or zero in order to avoid perpetual motion. Hence, W is the maximum work that can be produced in a reversible cycle.

(2) *The work produced in a cycle is independent of the substance used and, for a given quantity of heat, depends only on the difference in temperature of the bodies between which the cycle works.*[6]

For Carnot the crucial factor in the process of generating mechanical work is the difference in temperatures between the boiler and the condenser of a steam engine. Hence, he correctly suggests that the work produced in a cycle is independent of the working substance involved. Carnot suggests also that the work produced in a cycle is *a function of the quantity of heat* transferred from (the hot) body A to (the cold) body B during the process. That is, the work produced in a complete cycle C is $W(C) = g(Q^{tr}, T_f - T_i)$. The demonstration of the second theorem appears to be tied to the wrong hypothesis that heat is conserved in a Carnot cycle. For, despite his doubts concerning the conservation of heat, we have seen Carnot assuming that the work produced in his cycle is due to the *redistribution* of caloric between bodies A and B. This can be taken to mean that the quantity of heat Q_A released from body A is equal to the quantity of heat Q_B absorbed from body B and that, therefore, all heat gets transferred from body A to body B. So, one may assume that Carnot's proof rests on the equation: $Q_A = Q_B = Q^{tr}$. This is a conservation statement, and it might appear that it is essentially employed in Carnot's derivation. Yet, Carnot was again very careful. In presenting his cycle, he never explicitly said that the quantity of heat released by body A was *absorbed* by body B. In the crucial step (4) of his cycle (see Figure 6.1 and the text preceding it), Carnot said only that 'the compression is continued till the air acquires the temperature of the body A' (1824: 18/75). This is correct, and by no means does it entail that $Q_A = Q_B = Q^{tr}$. Hence, Carnot did not appeal to any assumptions about the conservation of heat in order to establish his law.[7]

In order to make this last point more forceful we must jump slightly ahead and see Émile Clapeyron's account (1834) of Carnot's cycle. Clapeyron was the first to put Carnot's theory in its well-known diagrammatic form. But during the crucial step (4), where the gas, after being compressed isothermally in contact with the cold body B, is allowed to compress adiabatically, Clapeyron stated that 'the compression continued till the heat released by the compression of the gas and absorbed by the body B is exactly equal to the heat communicated by the source A to the gas, during its expansion in contact with it in the first part of the operation' (1834: 76–77). This *is* a clear conservation statement. So, in *interpreting* Carnot's cycle, Clapeyron demanded that $Q_A = Q_B = Q^{tr}$, i.e. he demanded that heat was conserved in a Carnot cycle.

In light of the foregoing analysis of Carnot's theorems and their demonstrations, it transpires that they do not depend on the hypothesis that heat is conserved in a Carnot cycle. In fact, in his posthumously published notes – which were written not long after his memoir – Carnot suggested that the caloric theory should be abandoned. He stressed that within the caloric theory, it 'would be difficult to say why, in order to develop motive power by heat, a cold body is required; why motion cannot be produced by consuming the heat in a heated body' (1986: 187). Carnot suggested that the hypothesis of the conservation of heat broke down when it was called upon to explain the production of work by heat.[8] He also stressed that the caloric theory of heat was undermined by a series of experimental results, mostly related to the production of heat by friction (1986: 185–186). From his posthumously published notes, one can also see that sometime between 1824 and his early death in 1832 Carnot countenanced a dynamical theory of heat.[9]

Localising relations of evidential support

Stated in an anachronistic way, the attitude of the most eminent scientists towards the caloric theory of heat, in the light of the well-founded laws of experimental calorimetry, the law of adiabatic change and Carnot's theory of work, was this: the probability of these laws, given the hypothesis that heat is a material substance, is not high, and moreover it is not overwhelmingly greater than the probability of these laws, given the falsity of this hypothesis.

My account thus far has rested on the premiss that it is both in principle and in practice possible to *localise* the relations of evidential support, and to show which parts of a theory are supported by the evidence at hand, or at any rate, which parts are better supported than others.[10] However, Laudan (1981: 26–27) has commented that realists must be holists in confirmational matters, for otherwise they cannot maintain that the deep-structural claims of a theory are well supported. He also seems to think that realists must accept the view that observational evidence for a theory is evidence for everything that a theory asserts. Laudan's allegations about realist commitments seem to rest on a rather misleading account of evidential support, according to which empirical evidence cannot give support to some of the theoretical claims involved; instead empirical evidence supports a theory as a whole and, therefore, it supports each and every one of its theoretical claims.

Laudan's claim stems from a bad reading of Boyd (1981), according to whom the support which empirical evidence lends to a theory extends all the way to the deep-structural claims of the theory. Boyd's point, however, is meant to deny the empiricist contention that the empirical evidence supports only the empirical claims made by the theory. He rightly stresses that evidence for the empirical adequacy of a theory can be evidence also for the truth of a theory, and in particular for the truth of its theoretical claims. Boyd's position, however, does not commit the realist to holistic confirmation. All it says is that confirmation extends all the way to the theoretical claims, and does not just stay at the observational level. Yet, there is no reason to think that empirical evidence cannot lend a different credence to the several theoretical constituents of the theory. Nor is there any reason to think that all parts of a theory are equally well supported by the evidence. Empirical evidence may well extend to the theoretical elements of a theory, and yet support some of them better than others, or remain silent about yet other theoretical claims. As this study of the caloric theory of heat has shown, in actual scientific theories there are those deep-structural claims which are warranted by the evidence, and others which are not.

Let me highlight some ways in which the evidence supports some theoretical claims only weakly.

- Some piece of evidence may be in conflict with some particular theoretical claims.
- In the light of recalcitrant experience, some theoretical claims are modified in an ad hoc way in order to conform to the unfavourable new evidence.
- Some theoretical claims are such that the evidence does not make them any more likely than alternative and incompatible claims.
- Some theoretical claims are 'neutral' with respect to sound background beliefs, in that the latter do not increase their probability of being true.

As I stressed in Chapter 5, not all the deep-structural claims of a theory play the same role in the derivation of predictions and in providing well-founded explanations of observable phenomena. Some theoretical claims may be essential to the derivation of predictions and explanations of the phenomena; some others may be 'idle'. Some theoretical claims may be mere visualisations of underlying causes, and as such unusable in the generation of testable predictions or, at any rate, in specifying circumstances under which they can be thoroughly tested. Given that deep-structural claims may be supported by evidence to different extents – conferring probabilities that range from high to low – it is a good empirical constraint on any confirmation theory to localise the praise and blame for the successes and the failures of a theory, and to differentiate the degrees of support of the several theoretical constituents. So, it is entirely consistent to stress that empirical evidence sends its support all the way up to the theoretical level, while recognising that it does not do so indiscriminately and without differentiation.

In sum, the realist answer to Laudan's allegations about holistic confirmation should be this: if scientists entertain some theoretical beliefs it is because empirical evidence, together with other sound background beliefs, renders them well confirmed. This position leaves space for a localised theory of confirmation. For empirical evidence surely extends right to the deep-structural claims. Yet, on its way there, it may confirm them differentially. Realists need not commit themselves to unwarranted theoretical claims; yet they have good reason to commit themselves to theoretical claims, insofar as the latter are well supported by evidence and other sound background beliefs. Evidence can be such that it shows which theoretical claims are likely to be true, and which we must discard or suspend our judgement about. So, scientific realists need not accept a theory in its entirety. Instead, realism requires and suggests a *differentiated attitude to*, and *differentiated degrees of belief in*, the several constituents of a successful and mature scientific theory. The degree of belief one has in a theory is, in general, a function of the extent of its support by the available evidence. Since different parts of a theory can be supported to different degrees, realists should place their bets on the truth of a theory accordingly. So, let me just emphasise that belief can be the right epistemic attitude towards scientific theories, but belief admits of degrees. Hence belief in a theory, and in its several theoretical constituents, is often a matter of degree.

From the caloric theory to thermodynamics

One main conclusion of the case-study thus far is that the laws of the caloric theory can be deemed to be approximately true independently of the referential failure of 'caloric', i.e. irrespective of the absence of a natural kind as the referent of the term 'caloric'. So, a point worth highlighting is that when the laws established by a theory turn out to be independent of assumptions involving allegedly central theoretical terms, it can still make perfect sense to talk of the approximate truth of this theory.

The existence of a significant truth-content in the caloric theory is not a conclusion that we draw by hindsight. I shall now turn my attention to Clausius, one of the founders of modern thermodynamics, in order to show the sense in which the caloric theory of heat was taken to be approximately true by the proponents of the new theory of thermodynamics.

Rudolf Clausius concentrated his research on the capacity of heat to produce work. He made the following observations:

1. Joule's experimental principle of the equivalence of heat and work, i.e. the principle that a certain quantity of heat must be consumed in the production of a proportional amount of work, strictly contradicts Carnot's 'subsidiary statement' that no heat is lost in a thermal cycle where work is produced.
2. Joule's principle is strictly compatible with Carnot's 'essential principle' that heat always flows from a warm to a cold body (Clausius 1850: 112).

According to Clausius, during the production of work it may be the case that both a quantity of heat is consumed in the generation of work *and* a quantity of heat passes from the warm to the cold body, so that both quantities stand in a definite relation to the work produced. So, in place of the one hypothesis of the caloric theory which, as such, contradicts Joule's experimental finding that heat is consumed during the production of work, Clausius issued two distinct but compatible hypotheses.

Analysing the Carnot cycle, Clausius introduced the new concept of the 'internal energy' of a gas, which 'has the properties which are commonly assigned to the total heat, of being a function of V and T, and of being therefore fully determined by the initial and final conditions of the gas . . .' (1850: 122). The internal energy is a function of the macroscopic parameters

of the gas and, therefore, it is conserved in a complete cycle. Clausius suggested that the so-called 'total quantity of caloric' absorbed by the gas (or the working substance in general) consists, in fact, of two parts: (i) the internal energy of the gas which has the properties which the advocates of the caloric theory erroneously attributed to the 'total quantity of heat' and (ii) the quantity of heat consumed for the generation of work, the amount of which depends on the course of change the gas undergoes. So, it is important to note that according to Clausius, 'caloric' was a partially referring term. It did not refer to any material substance, but, under its mature formulation, it could be seen as referring partially to the internal energy of a substance.

Clausius went on to derive the first law of thermodynamics, which asserts that the quantity of heat received by a gas during very small (infinitesimal) changes of volume and temperature is equal to the increase in the internal energy of the gas plus the heat consumed for the work done by the gas. He notes (1850: 133–134) that despite the fact that Carnot was far from proving the first law of thermodynamics, his theorems were independent of the assumption that no heat was lost in a Carnot cycle. They follow from the physical impossibility of perpetual motion.[11] Clausius concluded:

> It seems therefore to be *theoretically* admissible to retain the first and the really essential part of Carnot's assumptions [i.e. that 'the equivalent of the work done by heat is found in the mere transfer of heat from a hotter to a colder body'] . . . [And it is similarly admissible] to apply it as a second principle in conjunction with the first [i.e. the first law of thermodynamics]; and the correctness of this method is, as we shall soon see, established already in many cases by its *consequences*.
>
> (1850: 132, 134)

The reader will have noted that Clausius' derivation from Carnot's theory rests on a distinction between an *essential* and a *subsidiary* part. But, what is the justification for this distinction? I shall not repeat what I have already said about the alleged centrality of the assumption that the cause of heat is a material substance. The point I want to make is that by pointing to the reasons why the community upheld this distinction, we can see why this distinction was justified. Let us then see what these reasons were.

1. A shared desideratum of the community was to keep as much as possible of Carnot and Clapeyron's mathematical machinery and successful predictions.
2. The (for Clausius) essential parts of Carnot's theory were those which were best supported by the evidence.
3. Helmholtz, Clausius and William Thomson showed that the disputed principle of the conservation of heat was unnecessary for the derivation of Carnot's law.

4. The sound laws which had been established within the caloric theory were readily deduced and accounted for in the new theoretical framework of thermodynamics.
5. No alternative theory was ever produced which dictated the total rejection of Carnot's theory.

Hence, we may conclude that Clausius' justification for the distinction between essential and subsidiary principles of Carnot's theory reflected the *theoretical and methodological desiderata* of the scientific community.

Having thus completed my brief account of the transition from the caloric theory to thermodynamics, I must stress one last point: the development of the dynamical representation of heat was constrained by the successes of the caloric theory. The latter were such that any alternative account of heat should have been able to accommodate them. Not only did the dynamical representation of heat after 1850 provide a truer account of the causal mechanisms involved in the thermal processes, but it also succeeded in accommodating the sound parts of the previous theory within the bounds of the new causal account of the nature of heat. The important point here is that this was the practice of the main scientists working in the field: they located and preserved the well-supported content of the caloric theory of heat by replacing the erroneous hypothesis of conservation of heat by two independent and compatible hypotheses and by retaining the rest of the sound laws. It is in this sense that the caloric theory can be said to be approximately true, despite the referential failure of 'caloric'.

One may even suggest that if the term 'caloric' was not so loaded, it could have been retained in order to refer to the internal energy of a substance. As we saw, the latter, like caloric, is a function of the macroscopic properties of a substance even within the new theory of heat. Hence, there is a sense in which 'caloric' may be seen as referring to the internal energy.[12] Be that as it may, the relevant moral about the reference of abandoned theoretical terms is that: not all cases of abandoned terms are troublesome. The serious cases concern terms which were indeed central in some genuinely successful theory; *central* in the senses that

- descriptions of the putative referent of the terms were indispensable in the derivation of predictions and in the well-founded explanations of phenomena; and
- the advocates of a theory took the theory's successes to warrant the claim that there were natural kinds denoted by these terms.

It is only about such terms that the issue of preservation of reference is pressing. If such terms turn out to be vacuous, then there seems to be no connection between empirical success and the successful reference of a theory's theoretical terms. But not all abandoned terms have been this central. When some abandoned term had not been central, realists should not be

required to show how it can possibly be referential. 'Caloric', simply, was not such a central term.

NINETEENTH-CENTURY OPTICS: THEORIES AND MODELS

Abstract dynamics versus concrete models

One of the prime objectives of theoretical research in optics during the nineteenth century was the formulation of a dynamical theory of light-propagation, which aimed to yield the laws of the behaviour of light from general dynamical principles concerning the carrier of light-waves, known as the 'luminiferous ether'. This research programme was developed by Augustin Louis Cauchy, George Green, James McCullagh and George Gabriel Stokes. Within the framework of the new electromagnetic conception of light, it was pursued further by James Clerk Maxwell and his followers.

Although, thanks to the pioneering research of Augustin Fresnel, the luminiferous ether was known to be a conservative system which sustained transversal waves, it is important to stress that its physical constitution and its internal connections were unknown. In view of this, theoretical research in optics was developed on the basis of an interplay between general dynamical theories and concrete models of the constitution of the ether.

The theoretical framework that scientists adopted was Lagrangian dynamics. They considered the carrier of the light-waves as a dynamical system whose general behaviour could be studied by Lagrangian dynamics and aimed to derive, within this framework, the most general laws of light-propagation. This was taken to be sufficient for the development of a dynamical account of light-propagation. The use of Lagrange's method enabled the scientific community to investigate the general dynamical properties and functions of the carrier of the light waves, leaving 'out of account altogether the details of the mechanism, whatever it is, that is in operation in the phenomena under discussion' (Larmor 1893: 399). The subsumption of light-propagation under Lagrangian dynamics required the specification of the kinetic-energy function and the potential-energy function. While the form of the dependence of the kinetic energy on the velocity of the moving bodies is in all cases the same and can be known, the form of the dependence of the potential energy on the position of bodies cannot be generally stated: it depends on the special nature and characteristics of the system under consideration. Hence, the prime task of theorists was to specify a potential-energy function which could adequately describe the behaviour of the ether. To this end, they had to employ several modelling assumptions about the nature and characteristics of the ether.

It was exactly at this point that particular theoretical models of the ether proved to be very useful. As we shall see in detail in the sections that

follow, these models aimed to specify the potential-energy function of the ether. Having formulated such a potential-energy function, the next task was to correlate it with some of the known properties of light – amplitude, intensity and others. Then, the resulting theory was put to the test by examining whether it yielded the known laws of light-propagation. For the purpose of offering a dynamical basis for light-propagation, no further specification of the nature of the carrier of light waves was needed. For the specification of the potential-energy and kinetic-energy functions was sufficent to subsume light propagation under the domain of dynamics; and then it was possible to examine whether the resulting laws of motion could yield the known laws of light-propagation. For the purposes of this investigation, the significant issue here is that the advancement of dynamical theories of light-propagation did not require scientists to *believe* that the ether was constituted in the way implied by the specific model in use.[13]

However, the models employed for the specification of an energy function did also stand for *possible* candidates for the constitution of the carrier of light-waves. For instance, the model that Green (1838) and Stokes (1849; 1862) employed rested on the assumption that the energy function of the otherwise unknown ether could be associated with that of an ordinary elastic solid. Then a model based on the dynamics of an elastic solid (henceforth, an elastic-solid model) was used in an *heuristic* way to investigate whether the constitution and internal connections of the ether could be mapped on those of an elastic solid. Such a procedure was heuristically valuable for the discovery of what ether could be, and what ether is not.

The heuristic value of an elastic-solid model – as opposed, for instance, to models based on the dynamics of liquids – was based on certain *positive analogies* between an elastic solid and the otherwise unknown carrier of light-waves. In particular, after Fresnel's work, scientists settled for the view that light-waves were uniquely transversal. This fundamental discovery suggested that the carrier of light waves had to possess properties in virtue of which it could sustain transversal waves.[14] A model of such an otherwise unknown carrier of light-waves could be constructed on the basis of the propagation of a disturbance through an elastic solid. For the latter exhibits properties, such as capacity to sustain transversal waves, which are analogous to the known properties of light-propagation. In view of this fact, most scientists started attacking the problem of the dynamical foundations of light-propagation 'through the analogy with the propagation of elastic waves in solid bodies' (Larmor 1893: 392). They used the features of the propagation of a disturbance in elastic solids as a set of assumptions about the constitution of the carrier of light-waves.

Despite its usefulness, the elastic-solid model of the constitution of the ether was not taken to reveal the real constitution of the ether. Here part of the problem lies with the fact that an elastic solid can also transmit longitudinal waves. In fact – and this was the touchstone for the elastic-solid model – it follows from the laws of mechanics that when a transversal wave

Philosophical issues in science
Edited by W. H. Newton-Smith
Balliol College, Oxford

Real History
Martin Bunzl

Brute Science
Hugh LaFollette and Niall Shanks

Verificationism
Cheryl Misak

Living in a Technological Culture
Mary Tiles and Hans Oberdiek

The Rational and the Social
James Robert Brown

The Nature of the Disease
Lawrie Reznek

The Philosophical Defence of Psychiatry
Lawrie Reznek

Inference to the Best Explanation
Peter Lipton

Time, Space and Philosophy
Christopher Ray

Mathematics and Image of Reason
Mary Tiles

Evil or Ill?
Lawrie Reznek

The Ethics of Science: An introduction
David B. Resnik

Philosophy of Mathematics: An introduction to a world of proofs and pictures
James Robert Brown

Theories of Consciousness: An introduction and assessment
William Seager

Psychological Knowledge: A social history and philosophy
Martin Kusch

Is Science Value Free? Values and scientific understanding
Hugh Lacey

Scientific Realism: How science tracks truth
Stathis Psillos

Scientific Realism

How science tracks truth

Stathis Psillos

ROUTLEDGE
Taylor & Francis Group
London and New York

First published 1999
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

Typeset in Times by Florence Production Ltd, Stoodleigh, Devon
Printed and bound in Great Britain by Biddles Ltd, Guildford
and King's Lynn

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Psillos, Stathis, 1965–
Scientific realism: how science tracks truth/Stathis Psillos.
p. cm.
Includes bibliographical references and index.
1. Realism. 2. Science–Philosophy. I. Title.
Q175.32.R42P75 1999 99-29721
501–dc21 CIP

ISBN 0-415-20818-1 (hbk)
ISBN 0-415-20819-X (pbk)

**For my parents Maria and Demetris,
without whom not,
and for Athena**