STATHIS PSILLOS

2. IS THE HISTORY OF SCIENCE THE WASTELAND OF FALSE THEORIES?

1. INTRODUCTION

These instructions are intended to provide guidance to authors of Imagine you live in 1823 and you are about to design an advanced course on the theory of heat. About fifty years ago, Lavoisier and Laplace had posited caloric as a material substance —an indestructible fluid of fine particles— which was taken to be the cause of heat and in particular, the cause of the rise of temperature of a body, by being absorbed by the body. No doubt, you rely on the best available theory, which is the caloric theory. In particular, meticulous and knowledgeable as you are, you rely on the best of the best: Laplace's advanced account of the caloric theory of heat, with all its sophistication, detail and predictive might. You really believe that the best science teaching should be based on the best theories that are available. But you also believe that the best theory that is available is not *really* the best unless it has a claim to truth (or truthlikeness, or partial truth and the like). For what is the point of teaching a theory about the deep structure of the world unless it does say something or other about this deep structure?

The course goes really well. Your notes are impressive. They are soon turned into a textbook with lots of explanatory detail and fancy calculations. Alas! The world does not co-operate. There are no calorific particles among the things there are in it. Heat is destroyed when work is produced. The advanced theory is challenged by alternative theories, anomalies and failed predictions. There is agony, but in your lifetime, the caloric theory gets superseded and is left discredited in the wasteland of false theories. Decades come by. You are not around anymore. Your grand-children go to school and then to the university; they follow some new-fangled courses on the history of science. And there it is. The once powerful caloric theory of heat is now only a chapter in the history of science textbook.

Why is this not the fate of all (or most) of the theories we come up with? Why aren't current theories, despite their explanatory and predictive successes, just chapters in the hitherto unwritten history of science books? Why is science education not just future history of science education plus some problem-solvers? This might well be a fate we have to live with. Or, we might be able to say something different, viz., that science is a mixture of continuity and change and that there is reason to believe that parts of current scientific theories, like parts of past scientific theories, will survive radical theory-change and form (and keep on forming) a stable network of theoretical principles and explanatory hypotheses that

P.V. Kokkotas et al., (eds.), Adapting Historical Science Knowledge Production to the Classroom, 17–36.
© 2011 Sense Publishers. All rights reserved.

constitute the backbone of our evolving, but by and large true, scientific image of the world.

The aim of this paper is to motivate this alternative, especially in connection with issues related to science education. It is an appeal to render science education sensitive to the philosophical issues that can be drawn from a close look at the history of science. Section 2 is a brief outline of the caloric theory of heat. Section 3 is a little note on a methodological principle by means of which theories are judged — use-novelty. Section 4 offers a rather detailed exploration of Laplace's advanced caloric theory of heat and explains its shortcoming in light of the foregoing methodological principle. Section 5 shows that this kind of criticism of Laplace's theory has had an actual historical actor —the self-taught physicist John Herapath— and is not, therefore, available only by hindsight. Section 6 raises the question: where is the caloric theory now?; to which it offers the simple but painful answer: *in the history books*. It then paves the way for the discussion of the Pessimistic Meta-Induction, whose proper analysis and significance are given in Section 7. Section 8 draws on the material presented above to raise another important question: what is wrong with science education? To which it offers the answer that science education seems blind to the fact of theory-change in science and this obscures the importance of change as well as of continuity. History and philosophy of science can certainly help science education to avoid this blindness.

2. THE CALORIC THEORY OF HEAT

In the last quarter of the eighteenth century, French scientists, most notably Pierre Simon Laplace and Antoine Lavoisier posited caloric as a material substance — an indestructible fluid of fine particles— which was taken to be the cause of heat (Lavoisier, 1789, p. 1–2). Despite the theory's success in giving qualitative explanations of several heat phenomena[1], the caloric theory faced important experimental anomalies, most notably that caloric seemed to have no weight, and the generation of heat by friction, which contradicted the fundamental assumption of the caloric model, viz., that caloric is an indestructible fluid and that heat *per se* is a conservative quantity (cf. Davy, 1799, p. 9–23; Thompson (Count Rumford), 1798).

Moreover, the caloric theory was not the one and only theory of heat available. According to the proponents of the rival dynamical theory —most notably Humphry Davy and Count Rumford— the cause of heat was not a material fluid but rather, the very motion of the molecules that constitute a substance. In this sense, heat was nothing over and above the motion of the constituents of a body. In fact the dynamical conception of heat was able to explain both major foregoing anomalies that the caloric theory faced (cf. Thompson (Count Rumford), 1799).

The caloric theory could cope with these anomalies —for instance, by positing that the calorific particles were superfine and weightless. But as Joseph Black — a Scott advocate of the caloric theory— pointed out, all these attempts were rather ad hoc: their only justification was that they could save the caloric theory from refutation. In fact, Black (1803, p. 46) gave one of the first elegant accounts of ad hocness in his following remark:

> 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 (emphasis added).

In his lectures, Black presented both then available theories of heat and, although he stressed that "the supposition" that heat is a material fluid appeared the "most probable", he (1803, p. 44) 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.

Interestingly enough, Lavoisier and Laplace had an attitude similar to Black's. After presenting both current theories of heat, they suggested that the theory of experimental calorimetry was independent of both theoretical considerations concerning the nature of heat. They noted:

> We will not decide at all between the two foregoing hypotheses [i.e. material vs. dynamical theory of heat]. Several phenomena seem favourable to the second, [i.e. the mechanical theory] 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 [i.e. 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, p. 152–153).

3. A NOTE ON AD HOCNESS

Recall what Black said above: "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". This, for all practical purposes, can be taken to be what makes a theory (or a modification of a theory) *ad hoc* vis-à-vis a set of phenomena that theory is meant to explain. The charge of ad hocness is an epistemic charge. It is meant to illustrate a cognitive shortcoming of a theory —what Black captures by saying that an ad hoc theory "does not advance our knowledge". An ad hoc theory is not a well-supported theory despite the fact that it may entail the laws that it is meant to explain.

Clearly, there are two ways in which a known fact E can be accommodated in a scientific theory T.

(1) Information about E is used in the construction of a theory T and T predicts E.
(2) 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.

Although the Lakatosian school has produced a fine-grained distinction between levels of ad hocness, (cf. Lakatos, 1970, p. 175; Zahar, 1973, p. 101), I shall concentrate on 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:

(a) A body of background knowledge B entails the existence of E. Information about E is used in the construction of a theory T and T accommodates E.
(b) A body of background knowledge B entails the existence of 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.

The key point here is that though theories do get support by explaining already known and established empirical laws, this support is a function of the way the theory is constructed and of the way it is related to the known laws. Simply put, if a known phenomenon E is accommodated within T in the way suggested by (1) above, E does not support T, whilst if it is accommodated in the way suggested by (2) above, E does support T. Following Earman (1992, chapter 4, section 8) we can speak of "use novelty", where, simply put, a prediction P of a known fact E is use novel relative to a theory T, if no information about E was used in the construction of the theory which predicted it. So use-novelty is sharply distinguished from, and contrasted to, ad hoc accommodation.

4. ENTER LAPLACE

From the early 1780s until his death in 1827, Laplace was the dominant figure in theoretical physics in France. His programme, inspired and guided by Newton's work, was the provision of a theoretical account of all natural phenomena in terms of attractive and repulsive (central) forces exerted between the particles (cf. Fox, 1974).

In early 1820s Laplace was embroiled in a research project, aiming to give a theoretical basis and a quantitative explanation of the empirical laws of gases within the caloric theory of heat. This was a fine test for the caloric theory. Until then, the caloric theory had not been fully articulated mathematically and had not offered quantitative derivations and explanations of the empirical laws of heat. Not only did Laplace's attempts aimed to show that Newtonianism could conquer one more territory —the thermal phenomena— but also to establish that the caloric theory of heat could offer adequate theoretical explanations of heat phenomena.

Laplace first presented his mathematical theory before the French Academy of Sciences in September 1821 and came back to it in December 1822. He then published his researches in two articles in the *Connaissance des Temps* and reproduced them (with minor revisions) in the 12th book of his *Traite de Mécanique Céleste* in the early 1820s.

The central assumption of Laplace's account was that the so-called 'repulsive power' of heat —the power of heat in virtue of which a gas expands when heated— is due to repulsive forces among the particles of caloric. In particular, each molecule of ordinary matter attracts particles of caloric that form a caloric atmosphere around it. Yet, these caloric atmospheres repel one another. These repulsive forces tend to detach some quantity of caloric from each molecule and to create radiant caloric, which generates the repulsive power of heat (cf. 1823, p. 111–112)[2]. Contrary to these repulsive forces act the attractive forces between the molecules of matter, which are inversely proportional to the distance between two molecules. However, as we are about to see, Laplace took it that these attractive forces are insensible in gases and vapours.

Using these central assumptions Laplace suggested that the force law between two molecules of a gas is

$$H\,c^2\,\varphi(r)$$

where c is the quantity of caloric retained by each molecule, H is a gas-specific constant depending on the repulsive force of heat and $\varphi(r)$ is the attractive force exerted between the two molecules, where $\varphi(r) \propto 1/r$ (1821, p. 278). He then calculated the repulsive force exerted on an envelope of a gas and equated it with the pressure P exerted by this envelope on surrounding layers of the gas. He found that

$$P = 2\pi HK\rho^2 c^2 \qquad (1)$$

where $2\pi HK$ is a constant and ρ is the density of the gas (op.cit., p. 280).

Laplace had thereby managed to correlate the quantity of caloric contained in a gas with the macroscopic parameter of pressure and hence to provide a potential mechanism that connects variations in the macroscopic quantity of pressure with variations in the microscopic structure of heat and matter.

The next problem was to specify a connection between the quantity of caloric contained in a gas with the macroscopic parameter of temperature (op.cit., p. 281). Laplace suggested that the quantity of caloric rays received at a surface, at a given instance, is solely a function of the temperature of the gas, and independent of the nature of surrounding bodies. Call this function $\Pi(T)$. The quantity of radiant caloric detached from a molecule m —due to the repulsive forces between the caloric c of the molecule m and the caloric atmospheres of neighbouring molecules— is ρc^2, that is, it is proportional to the quantity ρc of the caloric of surrounding molecules and the quantity c of the caloric retained by molecule m. Since at any given moment, there is thermal equilibrium in the gas, it follows that

$$q\Pi(T) = \rho c^2 \qquad (2)$$

where q is a proportionality constant depending on the molecules of the gas. Incidentally, in arriving at this equation, Laplace neglected the quantity of free caloric emanated by surrounding bodies, since as he noted, its extreme velocity renders it insensible (1821, p. 281). Be that as it may, by means of (2) Laplace had managed to connect the macroscopic parameter of temperature with the microscopic structure of caloric.

Given that temperature and pressure determine the macroscopic behaviour of gases, Laplace could now show how the observable behaviour of gases is caused by the micro-structure of caloric. Using (1) and (2), Laplace was ready to derive — within the framework of the caloric theory of heat— the laws of gases' and in particular the Boyle-Marriotte's law, Gay-Lussac's law and the equation of the state.

So far, so good. But there is a catch, which is relevant to the philosophical conclusions we might draw from this case. The catch is that there are certain respects in which Laplace's derivation was ad hoc. Let us see why.

Laplace's derivation of the laws of gases rested on two explicit assumptions: *First*, the attractive force between two molecules of a gas located at insensible distances from each other is very small; in fact, negligible. *Second*, the *only* operative force is the repulsive force between the caloric atmospheres of the molecules of the gas (cf. 1821, p. 285). The first assumption enabled Laplace to get rid of the factor $\varphi(r)$ and hence to derive equation (1) with no problem. This assumption is relatively uncontroversial. The second assumption however is by no means innocent.

According to Laplace's theory, the action between two molecules of a gas is actually the product of the following four forces:

1. The mutual repulsion of the quantities of caloric contained in caloric atmospheres around each and every molecule.
2. The attraction between the caloric atmosphere of the second molecule and the first molecule.
3. The attraction between the caloric atmosphere of the first molecule and the second molecule.
4. The mutual attraction between the two molecules.

Yet, the derivation (and explanation) of the laws of gases rested only on the first force. Even though neglecting the attractive force between molecules may have been reasonable, excluding the other two forces (two and three above) was not obvious. Laplace (1821, p. 185) admitted this when he said:

> Yet, I do not dare assure that the second and third forces are insensible, especially concerning vapours, when a light compression reduces them to the liquid state.

To make his model more realistic, Laplace went on to take into account the attractive forces exerted between the caloric atmosphere of a molecule m and the surrounding molecules of the gas, or vapour. However, here is the point where the ad hocness of Laplace's attempts becomes rather transparent.

Imagine, Laplace said, a cylindrical vase with indefinite height, containing a gas (1821, p. 285–286). Suppose also that the gas is pressed by a weight W put on the superior surface of the cylinder. Take, then, an infinitely thin horizontal plane A, at a distance from the superior surface of the cylinder, and suppose that the molecules of the gas are situated above this plane, at fixed positions. Let *m* be such a molecule, *r* its distance from the horizontal plane *A*, *f* its distance from another molecule *m'* situated underneath the plane *A* at distance *R* from it, and *s* the distance between the points at which the perpendiculars from the two molecules cross plane *A* (cf. Figure 1).

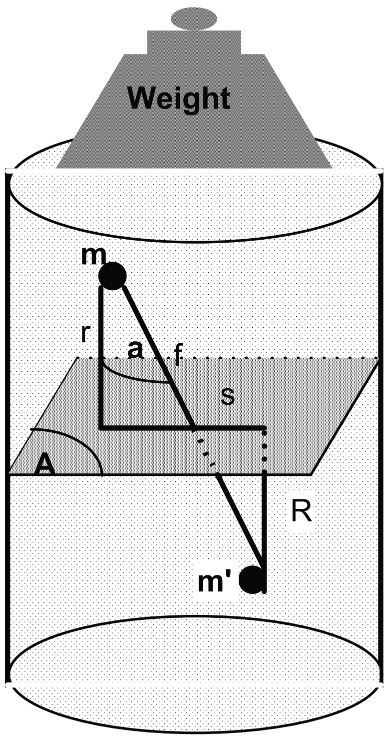

Figure 1. Laplace's model of caloric.

It is then evident that $f = \sqrt{(R+r)^2 + s^2}$. Generally, Laplace said, the repulsive action between the quantities of caloric retained by the two molecules *m* and *m'* is $H\varphi(f)$, while the attractive action between the caloric atmosphere of *m* and molecule *m'* is $N\varphi(f)$. The y-component of the total action between the two molecules will then be

$$(Hc^2 - Nc)\ \varphi(f)\ [(R+r)/f]$$

where $(R+r)/f$ is $cos(a)$. Laplace was then able to calculate the repulsive action of the whole gas situated under the plane *A* on the molecule *m* and, moreover, the whole action of the gas above plane *A* on the superior surface of the cylinder.

This action is counterbalanced by the pressure P of the weight placed on top of the superior surface. Hence, he derived

$$P = 2\pi\rho^2 (Hc^2 - Nc)K \tag{1'}$$

which is similar to (1) above, except that it also takes into account the attractive forces between the caloric atmosphere of a molecule m and the surrounding molecules.

Laplace then invented an analogous equation for temperature. Take, he said, the action between two molecules m and m' at a distance r. If all forces are taken into account, this action will be $Hc\varphi(r)–N\varphi(r)$. Suppose that the calorific radiation of molecule m is proportional to the number of surrounding molecules, their forces —except the negligible $\varphi(r)$— and the quantities of caloric contained in each molecule. Then, this radiation will be proportional to

$$\rho Hc^2 - \rho Nc \tag{A}$$

In a state of thermal equilibrium quantity (A) will be equal to the quantity of caloric received at a surface; that is,

$$\rho(Hc^2 - Nc) = q\Pi(T) \tag{2'}$$

This is similar to (2), except that it also takes into account the attractive forces exerted between the caloric atmospheres of molecules and the surrounding molecules. Then, by means of (1′) and (2′), Laplace was able to derive the laws of gases in the more realistic case where the attractive forces exerted between the caloric atmospheres of molecules and the surrounding molecules are taken into account.

The similarity between equations (1′) and (1) and (2′) and (2) seems to suggest that the attractive forces between the caloric atmospheres of molecules and the surrounding molecules could be safely neglected as very weak compared to the repulsive forces between caloric atmospheres.

However, two points are worth making:

1. In the derivation of (1′), Laplace *used* the assumption that the attractive forces between the caloric atmospheres of molecules and the surrounding molecules are very weak. As we have seen, he took it that the *total* force that the molecule m is subjected to when the attractive force between the caloric of m *and* the molecule m' is taken into account is repulsive. This means that the attractive forces between the caloric of a molecule and the surrounding molecules are very weak, and in fact negligible compared to the repulsive forces between caloric atmospheres —hence, practically they do nothing to modify or weaken these repulsive forces.
2. In arriving at equation (1′) Laplace neglected —without any reason— the effect on the pressure P of the molecules under plane A. As we shall are about to see, Laplace admitted this in his 1822 article. In fact, the only reason for formulating

the equation of pressure as he did seems to be that (1') could yield, together with equation (2'), the laws of gases *only if* it had this particular form.

Laplace's attempt to derive the laws of gases from the more realistic set of assumptions that *both* the attractive forces between the caloric of a molecule *and* the surrounding molecules are operative were ad hoc, and with no independent justification: the very fact that the attractive forces between the caloric of a molecule and the surrounding molecules must be negligible in order for the derivation to go through was *used* in showing that these forces were weak and negligible; and the very fact that the law of pressure must have a specific mathematical form if the laws of gases were to be derived, was used in the construction of this law.

As noted already, Laplace returned to his theory a year later (cf. 1822). There, he explained again how equations (1) and (2) are constructed and, therefore, how the laws of gases can be derived within the caloric theory. But he made it clear that the derivation works *only* on the assumption that the repulsive forces due to the caloric atmospheres are the *only* forces that operate (1822, p. 291). More interestingly, he remarked that in his own derivation of the laws of gases when the attractive forces between the caloric of a molecule *and* the surrounding molecules are taken into account, he neglected the action of the molecules under the plane A and hence his equation (1′) of the pressure P of the gas was not correct (1822, p. 296). He stressed that if the correct law of pressure is formulated, i.e. the one that, unlike (1'), takes also into account the pressure of the molecules under the plane A, then the three laws of gases cannot be derived (ibid.).

How were, if at all, the laws of gases to be derived within the caloric theory? Laplace admitted that the only way to carry out the derivation was to admit beforehand that "the attraction of each molecule of a gas on other molecules and their caloric is insensible" (ibid.). Therefore, Laplace's conclusion was, in effect, that unless the theory is modified in an ad hoc way, so that some forces are rendered negligible beforehand, the laws of gases could not be proved and explained within the caloric theory.

5. HERAPATH'S CRITICISM

The foregoing observation that Laplace's constructions were ad hoc is not one merely drawn by hindsight. John Herapath (1790–1868)[3], a then unknown physicist and self-taught schoolmaster from Bristol, in a paper that appeared in *Philosophical Magazine* in 1823, examined in detail Laplace's constructions, argued against their fundamental assumptions, and criticised them, explicitly, for being ad hoc. In this paper, Herapath gave one of the first clear-cut formulations of what it is for a theory to be ad hoc with respect to a set of laws, as it is clear from his following statement:

> (...) the equations [Laplace] has produced are more the offspring of a previous knowledge of what they should be from the phenomena, than of that sound reason which his other works usual manifest (1823, p. 65).

Herapath noted that Laplace's equation (2) which connects the quantity of caloric emanated from each molecule with the macroscopic quantity of temperature, is not correct. Laplace, as we have seen, took it that the calorific radiation of a molecule is ρc^2 Yet, Herapath observed, in calculating the calorific radiation of a molecule one must also take into account the *intensity* of the repulsion of the surrounding caloric. Therefore, the calorific radiation of a molecule must be $\rho c^2 \propto \psi(r)$, where $\psi(r)$ is a function of the intensity of repulsion of a particle of caloric, depending on the distance between the molecules[4]. In particular, the intensity of calorific radiation in a spherical envelope of radius r surrounding the radiating molecule will be $\psi^3\surd(\alpha/\rho)$, where α is a constant.

Then, instead of Laplace's equation (2), Herapath suggested that the correct equation should have been

$$q\Pi(t)=\rho c^2\ \psi^3\surd(\alpha/\rho) \qquad (2'')$$

It is obvious that (1) and (2”) cannot yield the laws of gases, and hence the latter cannot be derived —nor be explained— within the caloric theory of heat, unless some important assumptions are dropped, in an unjustified way. Herapath stressed that Laplace was not justified in neglecting the intensity of calorific radiation $\psi^3\surd(\alpha/\rho)$. In Laplace's theory the calorific radiation is due to the repulsive forces between the caloric atmospheres of neighbouring molecules. Then, it is obvious that these forces must depend on the distance r between these caloric atmospheres— in fact, on the distance r between molecules. Laplace, Herapath added, did consider the function $\psi(r)$ (cf. Laplace, 1821, p. 287; Herapath, 1823, p. 64). But he subsumed it under the constant q in the equations (2) and (2') (Herapath ibid.). However, this contradicted Laplace statement that the constant q is a factor dependent only on the nature of the molecules of the gas (1821, p. 281).

Herapath concluded that

> Laplace's principal and fundamental equations are erroneously deduced form his principles; and consequently that his subsequent conclusions [i.e. the laws of gases] are not consequences of what he first assumed (1823, p. 65)[5].

Herapath suggested that Laplace's theory was ad hoc with respect to the known laws of gases. In effect, Laplace knew what he wanted to derive —that is, the known laws of gases— and he 'cooked up' the principles of the caloric theory so that these laws would follow suit. The known laws of gases were not use-novel vis-à-vis Laplace's theory; they were accommodated within it in an ad hoc way. Herapath put this complaint in the following lengthy, but nice, quotation:

> Had the principles he [i.e. Laplace] sets out with been given him, namely, that there is such a thing as caloric, which, while strongly repulsive of its own, attracts and is attracted by other matter; which by some means radiates in extremely minute portions with great velocity; which attaching itself in considerable quantities to particles of mater overcomes their mutual attraction, and occasions them to stand at the greatest distance the envelope admits from each other; —had, I say, these things been given him [i.e. Laplace] without any knowledge of what the phenomena require, I would enture to appeal to himself, whether, with his mind so unacquainted, unbiased, and unprejudiced

> with the facts in question, his results would not have been very different from what they are (1823, p. 65).

Herapath challenged Laplace that had he not known in advance the laws he wanted to derive, the principles of caloric theory would not have been able to yield them. Laplace, in effect, *used* these laws in the construction of his theory, in the sense that he modified its principles in such a way that they, eventually, yield the laws of gases. As noted earlier, Laplace was aware (in his second paper on the subject) that the attractive forces between the caloric of a molecule and the surrounding molecules had to be rendered negligible if the derivation were to go through. Herapath's further point was that even if this were granted, the laws of gases could not be derived within the caloric theory, unless of course the latter was 'forced' to do so.

6. WHERE IS THE CALORIC THEORY NOW?

The fact is that the caloric theory of heat has long been abandoned. Its replacement with what came to be known as thermodynamics —pioneered by Rudolf Clausius and William Thomson (Lord Kelvin) and foreshadowed by Sadi Carnot— was an intricate and prolonged development. The key episode in this development was the admission that, contrary to what was implied by the caloric theory; heat was not a conservative quantity. After Clausius's work in thermodynamics, it was established that heat is not a state-function of the macroscopic properties (volume, temperature and pressure) of a gas. On the contrary, 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. If heat is not a conservative quantity, its representation cannot be based on an indestructible fluid, as caloric was supposed to be.

I have related this story elsewhere (cf. my 1994 and 1999, chapter 6). The point here is not to repeat it, but to answer the question in the section-heading in a straightforward manner: *the caloric theory is currently in the history of science books and not in the science textbooks*. The caloric theory is not part of the present corpus of established scientific theories; not an element in our evolving scientific image of the world. The world has simply no room for the caloric, despite the fact that a theory about it was the dominant theory for quite some time in the nineteenth century and despite the fact that it enjoyed explanatory and predictive success.

Is this case atypical? Is it an one-off case in the history of science? If it were, there would be no cause for concern. If the advanced caloric theory of heat was a historical oddity, its consignment to the history of science books would present no problem to either philosophy of science or to science education. But it is far from a typical. In fact, a well-known argument in the philosophy of science, known as the Pessimistic Meta-Induction on the history of science, suggests that current theories too are likely to be abandoned later on and be replaced by others, which are radically discontinuous with the extant theories. If this is so, there is a special problem for science education —apart form any other philosophical problem they might arise. This is that current science will turn out to be chapters in future history of science books and hence that the teaching of current scientific theories is not the teaching of a relatively stable and, by and large true, image of the world and of its deep structure, but rather the teaching of born-to-be-abolished failed explanations and hypotheses. Before we examine in this problem for science education, let us take a closer look at the Pessimistic Meta-Induction (PMI).

7. THE PESSIMISTIC META-INDUCTION

Larry Laudan has argued that the history of science is full of theories which were once empirically successful and yet turned out to be false. Laudan's argument can be summarised as follows (cf. 1981, p. 32–33):

> The history of science is full of theories which had been empirically successful for long periods of time and yet were shown to be false about the deep-structure claims they had 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, more likely to be false than true).

Laudan has substantiated his argument by means of what he has called "the historical gambit": the following list —which, Laudan says, "could be extended *ad nauseam*"— gives theories which were once empirically successful and fruitful, yet just false.

Laudan's list of successful-yet-false theories:

- 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

What is the target of Laudan's argument? It is the realist explanation of the success of scientific theories in terms of the (approximate) truth of these theories. In particular, the target of PMI is the epistemic optimism associated with scientific realism, viz., the view that science has succeeded in tracking truth. One key view associated with scientific realism is the claim that mature and predictively successful scientific theories are well confirmed and approximately true of the world; hence, the entities posited by them, or entities very similar to those posited, inhabit the world (see my 1999 and my 2009 for a defence). Part of the defence of this epistemic optimism realist has come to be known as the 'no-miracles' argument[6].

Briefly put, this is an argument that mobilises the successes of scientific theories (especially their successful novel predictions) in order to suggest that the best explanation of these theory-driven successes is that the theories that fuelled them were approximately true —at least in those respects that were implicated in the generation of the successes. But 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 current successful theories are approximately true.

There has been some discussion of the exact structure of PMI. I have argued in detail elsewhere that it is a kind of *reductio*. The target is the realist thesis that:

(A) Current successful theories are approximately true

Laudan does not directly deny that current successful theories may happen to be truthlike. His argument aims to discredit the claim that there is an explanatory connection between empirical success and truthlikeness 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 current successful theories are truthlike, then past theories cannot be

Past theories are deemed not to be truthlike 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 truthlikeness and truthlikeness cannot explain success: the realist's potential warrant for (A) is defeated.

No-one 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 —as well as everybody else— have to concede. However, a false theory can still be approximately true or truthlike. These are notions that have resisted a formal explication, but we can say, intuitively, that a theory is truthlike if it describes a world which is similar to the actual world in its most central or relevant features. So, the realist needs to show that past successful theories, although strictly speaking false, have been truthlike.

An obvious strategy that realists can follow is to try to reduce the size of Laudan's list. If indeed only very few past theories make it to Laudan's list of false-but-successful theories, the historical gambit loses much of its putative force. One way to reduce the size of the list is to impose stringent criteria as to what theories should count as mature and genuinely successful. It has been argued (see my 1999, chapter 5) that 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 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 should be happy with should such that it includes the generation of novel predictions which are in principle testable. Consequently, it is not at all clear that all theories in Laudan's list were genuinely successful.

The case of the advanced caloric theory of heat we have already discussed in some detail is a case in point. Despite its great sophistication, Laplace's mature theory enjoyed empirical success only by being, ultimately, tailored to fit the empirical laws. Not only were there no novel predictions issued by the theory, but even the already known facts that it managed to accommodate, it accommodated them in an ad hoc way.

The best strategy for blocking PMI is try to meet it head-on, by attacking its crucial premise (B). Without this premise the pessimistic conclusion does not follow, irrespective of the size of Laudan's list. But how can premise (B) be defeated?

In my (1999), I proposed the *divide et impera* strategy. The key idea is this. To defeat (B), it is enough to show that the genuine successes of past theories did not depend on what we now believe to be fundamentally flawed theoretical claims. Positively put, 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. Accordingly, when a theory is abandoned, its theoretical constituents, i.e., the theoretical mechanisms and laws it posited, should not be (and were not) rejected en bloc. Some of these theoretical constituents are inconsistent with what we now accept, and therefore they have to be rejected. But not all are. Some of them, instead of having been abandoned, 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, a substantive version of scientific realism can still be defended. So for the *divide et impera* move to work, we need to

(i) identify the theoretical constituents of past genuine successful theories that essentially contributed to their successes; and
(ii) show that these constituents, far from being characteristically false, have been retained in subsequent theories of the same domain.

The success of the *divide et impera* strategy is in the details. One should look at specific past theories that meet the stringent standards of empirical success and show in detail how those parts of them that fuelled their empirical successes were retained in subsequent theories. In my (1999, chapter 6) I engaged in two detailed case-studies concerning the several stages of the caloric theory of heat and of the theories of the luminiferous ether[7].

A claim that has emerged with some force is that theory-change is not as radical and discontinuous as the opponents of scientific realism have suggested. It has been shown that there are ways to identify the theoretical constituents of abandoned scientific theories which essentially contributed to their successes, to separate them from others that were 'idle' —or as Kitcher (1993) has put it, merely 'presuppositional posits'— and to demonstrate that the components that made essential contributions to the theory's empirical success were those retained in subsequent theories of the same domain. Given this, the fact that our current best theories may be replaced by others does not, necessarily, undermine scientific realism. All it shows is that **a)** we cannot get at the truth all at once; and **b)** our judgements from empirical support to truthlikeness should be more refined and cautious in that they should only commit us to the theoretical constituents that do enjoy evidential support and contribute to the empirical successes of the theory. Realists ground our epistemic optimism on the fact that newer theories incorporate many theoretical constituents of their superseded predecessors, especially those constituents that have led to empirical successes. The substantive continuity in theory-change suggests that a rather stable network of theoretical principles and explanatory hypotheses has emerged, which has survived revolutionary changes, and has become part and parcel of our evolving scientific image of the world.

Both Hasok Chang (2003, p. 910–912) and Kyle Stanford (2006) have challenged the move from substantive continuity in theory-change to approximate truth. It is argued that there is no entitlement to move from whatever preservation in theoretical constituents there is in theory-change to these constituents' being truthlike. But that's not quite right. What *is* right to say is that the mere demonstration of continuity in theory-change does not warrant the realist claim that science is 'on the right track'. Claiming convergence does not, on its own, establish that current theories are true, or likely to be true. Convergence there may be and yet the start might have been false. But the convergence in our scientific image of the world puts before us a candidate for explanation. The generation of an evolving-but-convergent network of theoretical assertions is best explained by the assumption that this network consists of truthlike assertions. So there is, after all, entitlement to move from convergence to truthlikeness, insofar as truthlikeness is the best explanation of this convergence.

Stanford has also claimed that the *divide et impera* move cannot offer independent support to realism since it is tailor-made to suit realism. According to him, it is the fact that the very same present theory is used *both* to identify which parts of past theories were empirically successful *and* which parts were (approximately) true that accounts for the realists' wrong impression that these parts coincide. He (2006, p. 166) says:

> With this strategy of analysis, an impressive retrospective convergence between our judgements of the sources of a past theory's success and the things it 'got right' about the world is virtually guaranteed: it is the very fact that some features of a past theory survive in our present account of nature that leads the realist both to regard them as true and to believe that they were the sources of the rejected theory's success or effectiveness. So the apparent convergence of truth and the sources of success in past theories are easily explained by the simple fact that both kinds of retrospective judgements have a common source in our present beliefs about nature.

This objection is misguided. The problem, as I see it, is this. There are the theories scientists currently endorse and there are the theories that were endorsed in the past. Some (but not all) of them were empirically successful (perhaps for long periods of time). They were empirically successful irrespective of the fact that, subsequently, they came to be replaced by others. This replacement was a contingent matter that had to do with the fact that the world did not fully co-operate with the then extant theories: some of their predictions failed; or the theories became overly ad hoc or complicated in their attempt to accommodate anomalies, or what have you. The replacement of theories by others does not cancel out the fact that the replaced theories were empirically successful. Even if scientists had somehow failed to come up with new theories, the old theories would not have ceased to be successful. So success is one thing, replacement is another. Hence, it is one thing to inquire into what features of some past theories accounted for their success and quite another to ask whether *these* features were such that they were retained in subsequent theories of the same domain. These are two independent issues and they can be dealt with (both conceptually and historically) independently. One should start with some past theories and —bracketing the question of their replacement— try to identify, on independent grounds, the sources of their empirical success; that is, to identify those theoretical constituents of the theories that fuelled their successes. When a past theory has been, as it were, anatomised, we can *then* ask the independent question of whether there is any sense in which the sources of success of a past theory that the anatomy has identified are present in our current theories. It's not, then, the case that the current theory is the common source for the identification of the successful parts of a past theory *and* of its (approximately) true parts. Current theories constitute the vantage point from which we examine old ones —could there be any other?— but the identification of the sources of success of past theories need not be performed from this vantage point.

8. WHAT IS WRONG WITH SCIENCE EDUCATION?

These instructions are intended to provide guidance to authors of Bluntly put, it is that it is oblivious to the complex philosophical lessons that can be drawn from the history of science. Unless we resolve for the view that current science teaching is future history-of-science teaching, science education should be sensitive to the fact that science as we know it is a mixed bag of continuity and change. Science education seems blind to the fact of theory-change in science and this obscures the importance of change as well as of continuity.

The responses to PMI outlined above suggest that the current scientific image of the world (which is what science teaching is about) is a hard-won image which has emerged out of a clash between truth and falsity; continuity and break. The continuity depicted in the current scientific image of the world is indeed hard-won, amidst false starts, failed hypotheses, idle and ad hoc explanations. This continuity represents whatever elements of past theories have a right to be called truthlike by having essentially contributed to the successes of otherwise abandoned theories and by having been retained in subsequent theories. This continuity signifies (at least on the realist reading suggested above) that what nowadays is taught in science courses is not destined to be part of the history of science books in two or three centuries from now.

What are science students being taught now? It's not enough to say they are being taught our best current guess about the deep-structure of the world. Laplace did not think of his theory as his best guess! He, like us today, thought of his theory as unveiling the deep and unobservable structure of the world. Guesses come and go. Theories are based on evidence and are meant to describe the world as it, more or less, is. An alternative would be to think of what is now taught as a set of practical recipes or problem-solvers; a rack filled with tools, as Pierre Duhem once put it. But this instrumentalist approach to science faces a number of problems, most of which are well-known. For one, it does not tally with the very idea that science has pushed back the frontiers of ignorance and error; for another, it does not even start to account for the fact that theories yield successful novel predictions and are used as premises in explanations of singular events. The question remains: what is taught now? Is it practical recipes + future chapters of the history of science books? Or is it chapters of an evolving-but-changing scientific image of the world?

This kind of question (or dilemma) was first raised in a serious way in the dawn of the twentieth century, just before the two major revolutions that shook up physics. It took the form of the 'bankruptcy of science' debate in France. In his address to the 1900 International Congress of Physics, Henri Poincaré (1902, p. 173) put the point thus:

> The man of the world is struck to see how ephemeral scientific theories are. After some years of prosperity, he sees them successively abandoned; he sees ruins accumulated on ruins; he predicts that the theories in vogue today will in a short time succumb in their turn, and he concludes that they are absolutely in vain. This is what he calls the bankruptcy of science.

But he went on to correct the view of 'the man of the world'. Poincaré says: "His scepticism is superficial; he does not understand none of the aim and the role of scientific theories; without this he would understand that ruins can still be good for something".

What then are ruins good for, apart from reminding us the glorious past and days of bygone splendour? There are two options, really.

> *One*: If theories are merely instruments for the co-ordination of empirical laws and the prediction phenomena, it is no problem that their theoretical parts might well be mere speculations which subsequently get abandoned and are destined to be chapter in hitherto unwritten history-of-science books. As Poincaré put it, after all, "Fresnel's theory enables us to [predict optical phenomena] as well as it did before Maxwell's time".
>
> *Two*: There is continuity in theory-change and this is not merely empirical continuity; substantive theoretical claims that featured in past theories and played a key role in their successes (especially in novel predictions) have been incorporated (perhaps somewhat re-interpreted) in subsequent theories and continue to play an important role in making them empirically successful.

This is the option that Poincaré himself favoured[8]. The key point is that option number two is not only living, but actually the one that renders science teaching intelligible and compelling.

How exactly science education should accommodate the philosophical lessons drawn from the history of science is itself a complex matter that I cannot discuss here. I will only suggest that part of the very idea of science education should be the cultivation and development of what might be called *scientific conscience*. This is not more theoretical or practical knowledge, but rather a set of methodological skills that constitute the scientific spirit: critical appraisal of one's own theory; sensitivity to the strengths and limitations of scientific inquiry; openness to criticism and correction; responsiveness to epistemic values and theoretical virtues; sensitivity to the historical complexity and the philosophical implications of the scientific enterprise.

Here is a case where scientific conscience becomes transparent. When we think about scientific theories and what they assume about the world we need to balance two kinds of evidence. The first is whatever evidence there is in favour (or against) a specific scientific theory. This evidence has to do with the degree on confirmation of the theory at hand. It is, let us say, *first-order evidence*, say about electrons and their having negative charge or about the double helix structure of the DNA and the like. First-order evidence is typically what scientists take into account when they form an attitude towards a theory. It can be broadly understood to include some of the theoretical virtues of the theory at hand (parsimony and the like) —of the kind that typically go into plausibility judgements about theories. The second kind of evidence (let's call it *second-order evidence*) comes from the past record of scientific theories and/or from meta-theoretical (philosophical) considerations that have to do with the reliability of scientific methodology. It concerns not particular scientific theories but science as a whole. (Some) past theories, for instance, were supported by (first-order) evidence, but were subsequently abandoned; or some scientific methods work reliably in certain domains but fail when they are extended to others. This second-order evidence feeds claims such as those that motivate the Pessimistic Induction. Actually, this second-order evidence is multi-faceted —it is negative (showing limitations and shortcomings) as well as positive (showing how learning from experience can be improved).

This is a philosophical problem. But science education won't educate good scientists unless it makes them aware that in judging scientific theories, they should try to balance these two kinds of evidence. And this means that science education won't train good scientists unless it trains them in history and philosophy of science.

NOTES

1 For a detailed account of the causal role that the caloric was called to play, see Fox (1971).

2 According to Chang (2004, 72) this Laplacian assumption modified considerably Lavoisier's original picture of the caloric.

3 For a brief account of Herapath's contribution to the kinetic theory of gases, see Mendoza (1961).

4 Herapath uses $\varphi(r)$ for this function, but this notation has been also used for the attractive force between two molecules of the gas.

5 Herapath did also object to Laplace's equation (1), which connected the pressure of a gas with the quantity of caloric upheld by its molecules. His chief point was that (1) unjustifiably neglects the attractive forces in virtue of which each molecule attracts its own caloric (1823, p. 62).

[6] It is based on Putnam's claim that realism 'is the only philosophy of science that does not make the success of science a miracle'.

[7] Chang (2003) has challenged some of the details of my case-study of the caloric theory. The discussion of Laplace's advanced theory presented above is meant, among other things, to meet some of Chang's criticisms concerning the actual historical development of the caloric theory.

[8] Though Poincaré took it that that there is an inherent limitation in what of the world can be known: its structure as opposed to how things are in themselves. This limitation was the child of Poincaré's adherence to some form of empiricism and some form of neo-Kantianism. It has been known as structural realism and need not concern us here (see my 2009).

REFERENCES

Black, J. (1803). *Lectures on the elements of chemistry* (J. Robison, Ed.). Edinburgh — all page references are from Roller (1950).

Chang, H. (2003). Preservative realism and its discontents: Revisiting caloric. *Philosophy of Science*, *70*, 902–912.

Chang, H. (2004). *Inventing temperature*. Oxford: Oxford University Press.

Davy, H. (1799). An essay on heat, light, and the communication of light. In *The collected works of H. Davy* (Vol. 2, pp. 1–86). London: Smith, Elder, and Co. Cornhill (1839).

Earman, J. (1992). *Bayes or bust? A critical examination of bayesian confirmation theory*. Cambridge MA: The MIT Press.

Fox, R. (1971). *The caloric theory of gases*. Oxford: Clarendon Press.

Fox, R. (1974). The rise and fall of Laplacian physics. In R. McCormmach (Ed.), *Historical studies in physical sciences*. Princeton, NJ: Princeton University Press.

Herapath, J. (1823). Observations on M. Laplace's communication to the Royal Academy of Science, "Sur l'Attraction des Sphères et sur la Répulsion des Fluides Élastiques". *Philosophical Magazine*, *62,* 61–66.

Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge*. Cambridge University Press.

Laplace, P. S., & Lavoisier, A. (1780). *Mémoire sur la Chaleur. Ouevres Complètes de Laplace* (Vol. 10). Paris: Gauthier-Villars.

Laplace, P. S. (1821). Sur l'Attraction des Sphères et sur la Répulsion des Fluides Élastiques. In *Connaissance des Temps pour l'année 1824* —reprinted in *Ouevres Complètes de Laplace* (Vol. 13, pp. 273–290). Paris: Gauthier-Villars.

Laplace, P. S. (1822). Développement de la Théorie des Fluides Élastiques et Application de Cette Théorie a la Vitesse du Son. In *Connaissance des Temps pour l'année 1825* —reprinted in *Ouevres Complètes de Laplace* (Vol. 13, pp. 291–301). Paris: Gauthier-Villars.

Laplace, P. S. (1823). Sur l'Attraction des Sphères et sur la Répulsion des Fluides Élastiques. In *Traite de Mécanique Céleste* (Livre XII, Chapitre II) —reprinted in *Ouevres Complètes de Laplace* (Vol. 5). Paris: Gauthier-Villars.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, *48*, 19–49.

Lavoisier, A. (1789). *Traite Elémentaire de Chimie*. Paris —English trans. as *Elements of chemistry*, by R. Kerr (1790), reprinted by Dover (1965).

Mendoza, E. (1961). A sketch for a history of the kinetic theory of gases. *Physics Today*, *14*, 36–39.

Poincaré, H. (1902). *La science et L'Hypothèse*. (1968 reprint). Paris: Flammarion.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.

Psillos, S. (2009). *Knowing the structure of nature*. London: Palgrave.

Roller D. (1950). The early development of the concepts of temperature and heat: The rise and the decline of the caloric theory. In J. B. Conant (Ed.), *Harvard case histories in experimental science*. Harvard University Press.

Stanford, P. K. (2006). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford: Oxford University Press.

Thompson, B. (Count Rumford) (1798). An inquiry concerning the source of the heat which is excited by friction. *Philosophical Transactions of the Royal Society*, *88*, 80–102.

Thompson B. (Count Rumford) (1799). An inquiry concerning the weight ascribed to heat. *Philosophical Transactions of the Royal Society*, *89*, 179–194.

Zahar, E. (1973). Why did Einstein's programme supersede Lorentz's. *British Journal for the Philosophy of Science*, *24*, 95–123, 223–262.

Stathis Psillos
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis (University Campus), Athens 15771, Greece
e-mail: psillos@phs.uoa.gr