The Problem of Induction and Metaphysical Assumptions Concerning the Comprehensibility and Knowability of the Universe

(Confirmation, Induction and Science, Conference at LSE, 10th March 2007, 17:30-18:15)
(**http://philsci-archive.pitt.edu/archive/00003141/**)
Nicholas Maxwell
(Emeritus Reader and Honorary Senior Research Fellow at University College London)
nicholas.maxwell@ucl.ac.uk
https://www.ucl.ac.uk/from-knowledge-to-wisdom

Abstract
Even though evidence underdetermines theory, often in science one theory only is regarded as acceptable in the light of the evidence. This suggests there are additional unacknowledged assumptions which constrain what theories are to be accepted. In the case of physics, these additional assumptions are metaphysical theses concerning the comprehensibility and knowability of the universe. Rigour demands that these implicit assumptions be made explicit within science, so that they can be critically assessed and, we may hope improved. This leads to a new conception of science, one which we need to adopt in order to solve the problem of induction.

1 Reasons for Making Implicit Metaphysical Assumptions Explicit within Science

Everyone agrees that evidence massively underdetermines theory. And yet, in scientific practice, much of the time, at most one theory is regarded as acceptable in the light of a body of evidence. Rarely does any of the infinity of rival theories able to predict the available evidence just as well as an accepted theory make their presence felt in scientific practice. It is entirely reasonable to conclude that this is because hidden, unacknowledged assumptions made by scientists, in addition to the evidence, exclude these infinitely many rivals. The obvious first step to take, in tackling the problem of induction, one would think, is to make these hidden, unacknowledged assumptions explicit. It is just this that one does if one is confronted by an invalid inference from correct premises to a correct conclusion: make explicit additional implicit premises which, once acknowledged, turn the invalid inference into a valid one. Why not take the analogous step in connection with scientific "inference" from evidence to theory (even if in this case, strictly speaking, no valid inference results)? This is the approach I argue for here. I argue that we need to make explicit implicit metaphysical assumptions concerning the comprehensibility and knowability of the universe which have the effect, when added to evidence, of tightly restricting theories that receive, and deserve, scientific attention (disunified rivals that predict the evidence being excluded).

My next point is entirely independent of the above line of thought, and has, in the first instance, absolutely nothing to do with the problem of induction. It is this. *If theoretical physics is to be rigorous, it is essential that physics makes explicit the substantial assumption that is implicit in the persistent acceptance of more or less unified theories only, even though there are always endlessly many empirically more successful but disunified rival theories available for consideration*.

This point is established in the next three sections.

2 Metaphysical Conjecture Implicit in the Methods of Theoretical Physics

Consider any accepted fundamental physical theory T – Newtonian theory, say, or Maxwellian electrodynamics, general relativity, quantum theory or the standard model. Endlessly many empirically more successful, but disunified, rival theories may be concocted as follows. First note that T, inevitably, however well-confirmed, will only have been verified empirically for a miniscule and highly atypical portion of the vast ocean of its empirical consequences. Thus, in order to concoct disunified rivals to T, at least as empirically successful as T, all we need do is specify some small region in the vast "space" of empirical consequences of T not yet put to the test, cancel the predictions of T for this small region, and put instead any predictions we please. The small region in question may be a region of space and time not yet observed. Alternatively, we may specify a small range of physical variables other than location in space and time – mass, for example, or temperature, or relative distance or velocity. Thus, given Newtonian theory, a rival theory might assert: everything occurs as Newtonian theory predicts except for systems consisting of solid gold spheres, each having a mass of a thousand tons, moving in otherwise empty space up to a mile apart, in which case the spheres attract each other by means of an inverse cube law of gravitation. Another rival asserts that everything occurs as Newtonian theory predicts until thirty tons of gold dust and thirty tons of diamond dust are heated in a platinum flask to a temperature of 500°C, in which case gravitation will instantly become a repulsive force everywhere. These last two rivals to Newtonian theory are "strictly universal" in Popper's sense (Popper, 1959, pp 62-8), in that they make no reference to any specific time, place or object. There is no limit to the number of rivals to Newtonian theory that can be concocted in this way, each of which has all the predictive success of Newtonian theory as far as observed phenomena are concerned but which makes different predictions for some as yet unobserved phenomena.

In order to concoct empirically more successful rivals to any accepted physical theory, T, we need only note the following. Almost all physical theories (a) run into some empirical difficulties and are, ostensibly, empirically refuted. Furthermore, (b) there are always repeatable phenomena, specifiable by means of low level empirical laws, L say, which T should apply to and predict, but which T fails to predict because the equations of T cannot be solved. And finally, (c) there will be repeatable phenomena, specifiable by means of empirical laws, L* say, which lie outside the range of applicability of T. In order to concoct T* - an empirically more successful rival to T – all we need to do is (a) modify T in an entirely *ad hoc* way so that the new theory successfully predicts the phenomena that refute T, and then, to this modified version of T, add (b) the laws L and (c) the laws L* (modifying the predictions of T appropriately). The resulting theory, T*, is empirically more successful than T in that it recovers all the empirical success of T and, furthermore (a) is empirically successful where T is (ostensibly) refuted, (b) is empirically successful where predictions cannot be extracted from T, and (c) successfully predicts phenomena outside the range of applicability of T. Furthermore, T* will make successful new predictions beyond the scope of T. We can, in any case, always enhance the empirical content and predictive success of T by adding onto T one or more independently testable and confirmed hypotheses, h_1, h_2, etc., to form the new theory T*. And we can combine these tricks to create a whole lot of further theories all empirically more successful than T.

None of these empirically more successful rivals to T ever get considered for a moment in scientific practice because they are all horribly *ad hoc*. They are what may be called "patchwork quilt" theories, in that they specify quite different dynamical laws for different ranges of phenomena. In order to be accepted in scientific practice a physical theory must satisfy two very different kinds of requirement. It must be (a) sufficiently empirically successful, and (b) sufficiently unified – i.e. such that just one set of laws applies to the range of phenomena to which the theory applies.

Now comes the crucial point. *In persistently only accepting unified theories, even though endlessly many empirically more successful but disunified rivals are always available, physicists make a substantial, persistent, metaphysical assumption about the nature of the universe, namely: the universe is such that no grossly disunified theory (such as those indicated above) is true.*

If scientists only accepted theories that postulate atoms, and persistently rejected theories that postulate different basic physical entities, such as fields - even though many field theories can easily be, and have been, formulated which are even more empirically successful than the atomic theories - the implications would surely be quite clear. Scientists would in effect be assuming that the world is made up of atoms, all other possibilities being ruled out. The atomic assumption would be built into the way the scientific community accepts and rejects theories – built into the implicit methods of the community, methods which include: reject all theories that postulate entities other than atoms, whatever their empirical success might be. The scientific community would accept the assumption: the universe is such that no non-atomic theory is true.

Just the same holds for a scientific community which rejects all disunified or patchwork quilt rivals to accepted theories, even though these rivals would be even more empirically successful if they were considered. Such a community in effect makes the assumption: the universe is such that no grossly disunified theory is true. Or rather, more accurately, such a community makes the assumption: "the universe is such that no disunified theory is true which is not entailed by a true unified theory – plus, possibly, true relevant initial and boundary conditions". (A true unified theory entails infinitely many approximate, true, disunified theories.)

2 Highly Problematic Character of Metaphysical Presupposition of Physics

The metaphysical assumption made by physics as a result of the persistent acceptance of unified theories only is highly *problematic* for a number of rather obvious reasons.

First, despite being substantial, influential, and a secure part of scientific knowledge (because empirically successful theories which clash with it are rejected), nevertheless there is no hope whatsoever of there being an argument in support of the *truth* of the assumption. It must remain a pure conjecture.

Second, it is uncertain as to what the assumption *is*, or *ought to be*. We may take "the universe is such that no grossly disunified theory is true" to be equivalent to "the universe is such that some as-yet-to-be-discovered more or less unified physical 'theory of everything' is true". In order to know what this asserts, we need to know what more or less "unified" means in this context. This is a fundamental problem in the philosophy of physics. Elsewhere, I have, I claim, solved this problem: see Maxwell (1998, chapter 4; 2004a, chapter 1, and appendix, section 2; 2007, chapter 14, section 2). Here is a brief sketch of this solution. The crucial insight is to appreciate that, in order to solve the problem we need to attend, not to the theory itself, its axiomatic structure or pattern of derivations, but to *the world* – or rather, to what the theory says about the world, the *content* of the theory in other words. A dynamical physical theory is unified if and only if its content, what it asserts about the world, is *the same* for all the physically possible phenomena to which it applies. A theory that is *different* in N ways, in N regions of the space of all possible phenomena to which the theory applies, is disunified to degree N. For unity, we require that N = 1. There are, however, *different* ways in which a theory may be different, some more serious that others. A theory may be different in different space-time regions. Failing that, it may be different for values of other variables – mass, temperature, relative distance or velocity. Failing that, it may be different because, in different regions of possible phenomena to which the theory applies, different forces apply or, failing that, there are different kinds of particle with different dynamical properties, such as charge or mass. In all, as I have shown elsewhere, there are at least eight different ways in which the content of a dynamical theory can be different, from *one* region to another in the space of all possible phenomena to which the theory applies. We thus have n different kinds of disunity, with n = 8, 7, … 1, disunity becoming of a less and less serious type as n goes from 8 to 1. These 8 kinds of disunity all exemplify, however, the same basic idea. The degree of disunity of a theory T is given by the value of (n, N), where n runs from 8 to 1 and N runs from ∞ to 1. For perfect unity we require that (n, N) = (1,1). The outcome of all this is that the thesis that the universe is such that the true 'theory of everything', T, is more or less unified can be interpreted to mean that the kind and degree of disunity of T, T(n,N) is such that $1 \leq n < 8$ and $1 \leq N < M$, where M is some appropriate, not too large integer. There are a large number of possibilities to choose from.[1]

Third, if we look at the metaphysical theses that have influenced acceptance of physical theories throughout the history of physics, we find that ideas have changed dramatically. In the 17th century there was the idea that everything is made up of corpuscles which interact only by contact. This gave way to the idea that the universe consists of point-particles surrounded by rigid, spherically symmetrical fields of force, which in turn gave way to the idea that there is one unified self-interacting classical field, varying smoothly throughout space and time. Today we have the idea that everything is made up of minute quantum strings embedded in ten or eleven dimensions of space-time. Nowadays everyone would agree that all but the last one are false and, given this historical record, it seems not unreasonable to hold that the last one is false as well. We have good reasons to hold, in other words, that the best metaphysical conjecture currently implicit in the methods of physics is *false*.

Fourth, this last point is reinforced by the observation that the metaphysical conjecture we are considering is about the ultimate nature of the physical universe, just that of which we are most ignorant.

Not only is the metaphysical presupposition of physics a pure conjecture. Not only is it uncertain as to what it is. Whatever it is, we have good grounds for holding it to *false*. And yet it is one of the most secure items of theoretical scientific knowledge, in that theories which conflict with this thesis are rejected however empirically successful they may be. To say that this thesis is *problematic* would seem to be a serious understatement.

4 Rigour Requires Implicit Metaphysical Thesis be made Explicit

Intellectual rigour demands that this substantial, influential, highly problematic and implicit assumption be made explicit, as a part of theoretical scientific knowledge, so that it can be critically assessed, so that alternative versions can be considered, in the hope that this will lead to an improved version of the assumption being developed and accepted. What is being appealed to here is the following:

Principle of Intellectual Rigour: An assumption that is substantial, influential, problematic and implicit must be made explicit so that it can be critically assessed, so that alternatives can be considered, in the hope that it can be improved.[2]

---

[1] The solution to the problem of the unity of theory, just outlined, emerges out of, and is linked to, the argument of section 2 demonstrating that persistent acceptance of unified theories only means physics makes a persistent metaphysical assumption.

[2] "Improved" means "made less false". For an explication of what this means, see Maxwell (2007, chapter 14, section 5, "The Solution to the Problem of Verisimilitude").

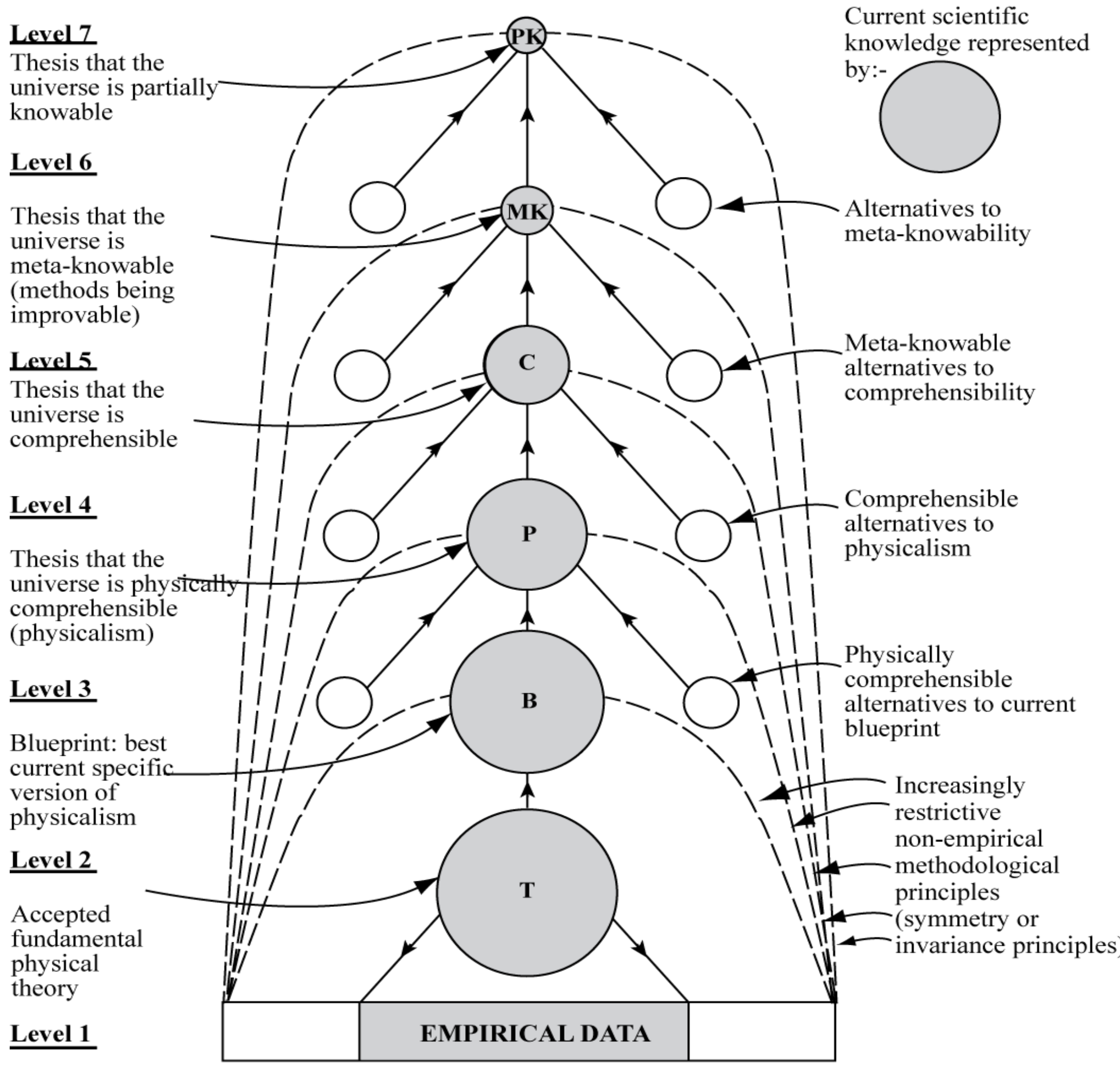
Even if this implicit metaphysical thesis was not particularly problematic, in the four ways just indicated, rigour would still require that it be made explicit within physics. The fact that the thesis is highly problematic, and profoundly influential over what theories are developed and accepted in physics, makes it a matter, not just of formal rigour, but of real practical importance for theoretical physics, that the thesis is made explicit.

At this point it may be objected that this metaphysical thesis – whatever precisely it may be – cannot be accepted as an item of authentic theoretical knowledge in physics until an argument has been produced providing grounds for holding that the thesis is *true*. This objection misses the point of the above line of argument. That has established that the thesis is already an item of theoretical knowledge, whether this is explicitly acknowledged or not. The only choice before us is (a) to acknowledge the thesis explicitly, or (b) to disavow the thesis, pretending it is not an item of scientific knowledge. I have shown that (a) results in a more rigorous physics than (b) *entirely in the absence of any argument in support of the truth of the thesis*. (a) makes it possible to criticize the thesis, develop and consider alternatives, in the hope that the thesis can be improved, whereas (b) makes this very difficult to do. Choosing (a) helps make physics conform to the above principle of intellectual rigour, whereas choosing (b) results in physics violating this principle.

5 Aim-Oriented Empiricism (AOE)

Once it is conceded that physics does persistently assume that the universe is such that all seriously disunified theories are false (or such that the true theory of everything is more or less unified), two fundamental problems immediately arise. What precisely ought this assumption be interpreted to be asserting about the universe? Granted that the assumption is profoundly problematic, a pure conjecture, substantial and influential but bereft of any kind of justification and all too likely in its current form to be false, how can rival versions of the assumption be rationally assessed, so that what is accepted by physics is improved?

Elsewhere I have argued at length that in order to solve these problems we need to adopt a conception of science that I have called aim-oriented empiricism (AOE): see Maxwell (1974; 1976; 1984; 1993; 2002 and especially 1998; 2004a; 2005a; 2006 and 2007, chapter 14). The basic idea of AOE is that we need to see physics (and science more generally) as making not one, but a hierarchy of assumptions concerning the unity, comprehensibility and knowability of the universe, these assumptions becoming less and less substantial as one goes up the hierarchy and thus becoming more and more likely to be true, and also becoming such that their truth is increasingly a requirement for science, or the acquisition of knowledge, to be possible at all: see diagram. The idea is that in this way we separate out what is most likely to be true, and not in need of revision, at and near the top of the hierarchy, from what is most likely to be false, and most in need of criticism and revision, near the bottom of the hierarchy. Evidence, at level 1, and assumptions high up in the hierarchy, are rather firmly accepted, as being most likely to be true (although still open to revision): this is then used to criticize, and to try to improve, theses at levels 2 and 3 (and perhaps 4), where falsity is most likely to be located.



Aim-Oriented Empiricism

At the top there is the relatively insubstantial assumption that the universe is such that we can acquire some knowledge of our local circumstances. If this assumption is false, we will not be able to acquire knowledge whatever we assume. We are justified in accepting this assumption permanently as a part of our knowledge, even though we have no grounds for holding it to be true. As we descend the hierarchy, the assumptions become increasingly substantial and thus increasingly likely to be false. At level 5 there is the rather substantial assumption that the universe is comprehensible in some way or other, the universe being such that there is just one kind of explanation for all phenomena. At level 4 there is the more specific, and thus more substantial assumption that the universe is *physically* comprehensible, it being such that there is some yet-to-be-discovered, true, unified "theory of everything". At level 3 there is the even more specific, and thus even more substantial assumption that the universe is physically comprehensible in a more or less specific way, suggested by current accepted fundamental physical theories. Examples of assumptions made at this level, taken from the history of physics, include those already mentioned (the universe is made up of corpuscles, point-atoms, classical field, quantum strings). Given the historical record of dramatically changing ideas at this level, and given the relatively highly specific and substantial character of successive assumptions made at this level, we can be reasonably confident that the best assumption available at any stage in the development of physics at this level will be false, and will need future revision. At level 2 there are the accepted fundamental theories of physics, currently general relativity and the standard model. Here, if anything, we can be even more confident that current theories are false, despite their immense empirical success. This confidence comes partly from the vast empirical content of these theories, and partly from the historical record. The greater the content of a proposition the more likely it is to be false; the fundamental theories of physics, general relativity and the standard model have such vast empirical content that this in itself almost guarantees falsity. And the historical record backs this up; Kepler's laws of planetary motion, and Galileo's laws of terrestrial motion are corrected by Newtonian theory, which is in turn corrected by special and general relativity; classical physics is corrected by quantum theory, in turn corrected by relativistic quantum theory, quantum field theory and the standard model. Each new theory in physics reveals that predecessors are false. Indeed, if the level 4 assumption of AOE is correct, then all current physical theories are false, since this assumption asserts that the true physical theory of everything is unified, and the totality of current fundamental physical theory, general relativity plus the standard model, is notoriously disunified. Finally, at level 1 there are accepted empirical data, low level, corroborated, empirical laws.

In order to be acceptable, an assumption at any level from 6 to 3 must (as far as possible) be compatible with, and a special case of, the assumption above in the hierarchy; at the same time it must be (or promise to be) empirically fruitful in the sense that successive accepted physical theories increasingly successfully accord with (or exemplify) the assumption. At level 2, those physical theories are accepted which are sufficiently (a) empirically successful and (b) in accord with the best available assumption at level 3 (or level 4). Corresponding to each assumption, at any level from 7 to 3, there is a methodological principle, represented by sloping dotted lines in the diagram, requiring that theses lower down in the hierarchy are compatible with the given assumption.

When theoretical physics has completed its central task, and the true theory of everything, T, has been discovered, then T will (in principle) successfully predict all empirical phenomena at level 1, and will entail the assumption at level 3, which will in turn entail the assumption at level 4, and so on up the hierarchy. As it is, physics has not completed its task, T has not (yet) been discovered, and we are ignorant of the nature of the universe. This ignorance is reflected in clashes between theses at different levels of AOE. There are clashes between levels 1 and 2, 2 and 3, and 3 and 4. The attempt to resolve these clashes drives physics forward.

In seeking to resolve these clashes between levels, influences can go in both directions. Thus, given a clash between levels 1 and 2, this may lead to the modification, or replacement of the relevant theory at level 2; but, on the other hand, it may lead to the discovery that the relevant experimental result is not correct for any of a number of possible reasons, and needs to be modified. In general, however, such a clash leads to the rejection of the level 2 theory rather than the level 1 experimental result; the latter are held onto more firmly than the former, in part because experimental results have vastly less empirical content than theories, in part because of our confidence in the results of observation and direct experimental manipulation (especially after expert critical examination). Again, given a clash between levels 2 and 3, this may lead to the rejection of the relevant level 2 theory (because it is disunified, *ad hoc*, at odds with the current metaphysics of physics); but, on the other hand, it may lead to the rejection of the level 3 assumption and the adoption, instead, of a new assumption (as has happened a number of times in the history of physics, as we have seen). The rejection of the current level 3 assumption is likely to take place if the level 2 theory, which clashes with it, is highly successful empirically, and furthermore has the effect of increasing unity in the totality of fundamental physical theory overall, so that clashes between levels 2 and 4 are decreased. In general, however, clashes between levels 2 and 3 are resolved by the rejection or modification of theories at level 2 rather than the assumption at level 3, in part because of the vastly greater empirical content of level 2 theories, in part because of the empirical fruitfulness of the level 3 assumption (in the sense indicated above).

It is conceivable that the clash between level 2 theories and the level 4 assumption might lead to the revision of the latter rather than the former. This happened when Galileo rejected the then current level 4 assumption of Aristotelianism, and replaced it with the idea that "the book of nature is written in the language of mathematics" (an early precursor of our current level 4 assumption). The whole idea of AOE is, however, that as we go up the hierarchy of assumptions we are increasingly unlikely to encounter error, and the need for revision. The higher up we go, the more firmly assumptions are upheld, the more resistance there is to modification.

AOE is put forward as a framework which makes explicit metaphysical assumptions implicit in the manner in which physical theories are accepted and rejected, and which, at the same time, facilitates the critical assessment and improvement of these assumptions with the improvement of knowledge, criticism being concentrated where it is most needed, low down in the hierarchy. Within a framework of relatively insubstantial, unproblematic and permanent assumptions and methods (high up in the hierarchy), much more substantial, problematic assumptions and associated methods (low down in the hierarchy) can be revised and improved with improving theoretical knowledge. There is something like positive feedback between improving knowledge and improving (low-level) assumptions and methods – that is, knowledge-about-how-to-improve-knowledge. Science adapts its nature, its assumptions and methods, to what it discovers about the nature of the universe. This, I suggest, is the nub of scientific rationality, and the methodological key to the great success of modern science.

6 AOE, Intellectual Rigour, and the Problem of Induction

Elsewhere I argue that AOE provides a framework within which the problem of induction can be solved, insofar as it is solvable at all: see Maxwell (1984, chapter 9; 1998, especially chapter 5; 2004a, especially appendix, section 6; 2005b; 2005c; 2007, chapter 14, section 6). It would take us too far afield for me to reproduce those arguments here. Instead I confine myself to making one or two remarks, before going on, in the next section, to consider an apparently decisive objection to the claim that the problem of induction must be solved within the framework of AOE.

Other things being equal, physics is more rigorous if it makes explicit the metaphysical assumption implicit in persistent exclusive acceptance of unified theories, than if it leaves this assumption implicit. This much has been established above. But physics is even more rigorous if it implements AOE. This is because AOE facilitates the *critical assessment* and *improvement* of metaphysical assumptions of physics in a way in which merely acknowledging some such assumption cannot do. AOE puts the above principle of intellectual rigour into practice more fully than any "one-level" view, which merely makes explicit a metaphysical thesis (or conjunction of such theses) implicit in the methods of physics, at one level.

But if AOE is more rigorous than these rival views, then attempts at solving the problem of induction must begin with AOE (in preference to these rival views). It is obviously hopeless to attempt to solve the problem of induction presupposing a conception of science that is unrigorous from the outset. That would be equivalent to attempting to justify the unjustifiable. I suggest, indeed, that the key reason why the problem of induction has remained unsolved ever since Hume brought it so graphically to our attention over two and a half centuries ago is that hitherto all attempts to solve the problem have been made in ignorance of AOE.

There is another reason why AOE is to be preferred to any "one-level" view, when it comes to tackling the problem of induction. The *hierarchical* character of AOE makes it possible to provide *different* reasons for accepting theses at the different levels of the hierarchy (or for preferring the theses of AOE to rival theses). There are no arguments for the *truth* of any of these theses. There are only arguments for *accepting* these theses (or preferring them to rival theses) *granted our aim is to improve knowledge of factual truth*.

Thus acceptance of the level 7 thesis of partial knowability is justified because the truth of this thesis is required for science, or the acquisition of knowledge, to be possible at all. Accepting this thesis as an item of scientific knowledge can only aid, and cannot harm, scientific progress.

The argument for accepting physicalism, with full unity, (n,N) = (1,1), at level 4 is rather different. This thesis deserves to be accepted because of its great empirical fruitfulness, in the sense that the whole of theoretical physics since Galileo can be regarded as one gigantic research programme progressively drawing closer to capturing physicalism as a testable "theory of everything" (a theory applicable in principle to all physical phenomena).[3] Newtonian theory, Maxwellian electrodynamics, Einstein's special and general theories of relativity, quantum theory, QED, quantum electroweak theory, chromodynamics, the standard model: all these great contributions to theoretical physics bring about far greater *unity* to the edifice of theoretical physics, in one or other of the eight kinds of unity partially indicated above, as I have shown in detail elsewhere:

---

[3] Kuhn (1962), Feyerabend (1978) and Laudan (1980) have all argued that nothing theoretical persists through revolutions. AOE makes clear that this is simply not true. As I have put it elsewhere (Maxwell, 1998, p. 181), "*Far from obliterating the idea there is a persisting theoretical idea in physics, revolutions do just the opposite in that they all themselves actually exemplify the persisting idea of underlying unity".*

see Maxwell (1998, pp. 123-140; 2004b, sections 3 and 5; 2007, chapter 14, sections 2 and 6).

It may be objected that a version of physicalism that only asserts partial unity would be just as empirically fruitful – as long as the theory of everything, T, is asserted to have just slightly greater unity than general relativity + the standard model.[4]

Even if this objection is conceded, and a version of AOE is adopted which accepts such a partially unified version of physicalism at level 4, still the argument would go through that the hierarchical structure of AOE provides us with a more rigorous conception of science than views which deny physics makes metaphysical assumptions, or which acknowledge some metaphysical thesis at one level. The argument of this paper is not tied to the specific version of AOE I have presented in section 5 above.

There are, however, grounds for preferring that version of physicalism which asserts full unity to any version which asserts partial unity only.

Fundamental to the whole argument for AOE is that physics needs to put the Principle of Intellectual Integrity into practice (AOE emerging, I claim, as the outcome of successive applications of this principle). In considering what thesis ought to be accepted at level 4, then, we need to consider what is implicit in those current methods of physics that influence what theories are to be accepted on non-empirical grounds – having to do with simplicity, unity, explanatoriness. There can be no doubt that, as far as non-empirical considerations are concerned, the more nearly a new fundamental physical theory satisfies all of the requirements for unity [with (n,N) = (1,1)], the more acceptable it will be deemed to be. Furthermore, failure of a theory to satisfy elements of these criteria is taken to be grounds for holding the theory to be false even in the absence of empirical difficulties. For example, high energy physics in the 1960s kept discovering more and more different hadrons, and was judged to be in a state of crisis as the number rose to over one hundred. Again, even though the standard model (the current quantum field theory of fundamental particles and forces) does not face serious empirical problems, it is nevertheless regarded by most physicists as unlikely to be correct just because of its serious lack of unity. In adopting such non-empirical criteria for acceptability, physicists thereby implicitly assume that the best conjecture as to where the truth lies is in the direction of physicalism(1,1). The Principle of Intellectual Integrity requires that this implicit assumption – or conjecture – be made explicit so that it can be critically assessed and, we may hope, improved. Accepting any version of physicalism that asserts partial unity only undermines rigour, because implicit assumptions are not made fully explicit and are thus not available for critical scrutiny. In addition, one may note, because physicalism(1,1) makes more definite, substantial claims than rival versions of physicalism that assert partial unity only, it is more open to critical appraisal than rival versions.

A second point to note is that it may well be that, even if some other version of physicalism(n,N) is true, with $n > 1$ and $N > 1$, nevertheless our best hope of discovering the truth may still lie in attempting to discover a theory that exemplifies physicalism(1,1), and failing in the attempt. As N becomes bigger, so the number of possible theories of

---

[4] Partially unified physicalism would assert that the true theory of everything, T, is unified to the extent that $1 \leq n < 6$ and $N < M$, where M is some appropriate integer, probably less than 50 taking into account the number of theoretically undetermined parameters of the standard model (see Maxwell, 2007, chapter 14, section 2 for details).

everything compatible with that version of physicalism rapidly increases. (If N = 2, and the universe is made up of two distinct unified, dynamical patterns, there are, nevertheless, in general, infinitely many ways in which these two distinct patterns can be fitted together to make infinitely many different possible universes exemplifying just these two dynamic patterns. The step from one specified unified pattern to two is the step from one possible universe to infinitely many!) It makes sense to seek the simplest, most discoverable possibility, and design our methodology accordingly.

My claim is, then, that AOE (a) makes explicit metaphysical assumptions implicit in persistent acceptance by physicists of unified theories only, and (b) provides the best possible framework for the *critical assessment* and *improvement* of these influential, problematic assumptions. AOE implements the above Principle of Intellectual Rigour in a more thoroughgoing way than any rival view, and thus constitutes a more intellectually *rigorous* conception of science than any rival view.

But if this is the case, then attempts at solving the problem of induction must begin with AOE (in preference to rival views). It is obviously hopeless to attempt to solve the problem of induction presupposing a conception of science that is unrigorous from the outset. That would be equivalent to attempting to justify the unjustifiable. I suggest, indeed, that the key reason why the problem of induction has remained unsolved ever since Hume brought it so graphically to our attention over two and a half centuries ago is that hitherto all attempts to solve the problem have been made in ignorance of AOE.

This assessment of the situation is, however, confronted by a difficulty. AOE seems to suffer from the fatal defect of vicious circularity: the success of science is justified by an appeal to metaphysical principles, which in turn are justified by the success of science. AOE even proudly boasts of this circularity! But, as Bas van Fraassen has put it ‘From Gravesande's axiom of the uniformity of nature in 1717 to Russell's postulates of human knowledge in 1948, this has been a mug's game’ (van Fraassen, 1985, pp. 259-60). Far from being rigorous, AOE is, it seems, hopelessly invalid. How does AOE escape this charge?

One point can be made straight away. The whole point of making the implicit metaphysical presuppositions of physics explicit is not to *justify* them but, quite to the contrary, to subject them to *criticism.* But this is not a very convincing reply. It is still the case that metaphysical theses influence what theories are accepted, and theories can influence acceptance of metaphysical theses.

I conclude by making two points in rebuttal of this charge of invalid circularity.

7 AOE not Invalidly Circular

The first point is this. The circularity inherent in AOE ceases to be a problem once we note that the thesis of “meta-knowability” at level 6 asserts that the universe is such that the circular, positive-feedback character of AOE is justified and, crucially, the reasons for accepting this thesis as a part of scientific knowledge make no appeal to the success of science whatsoever.

Permitting metaphysical assumptions to influence what theories are accepted, and at the same time permitting the empirical success of theories to influence what metaphysical assumptions are accepted, may (if carried out properly), *in certain sorts of universe,* lead to genuine progress in knowledge. Meta-knowability is to be interpreted as asserting that *this is just such a universe*.

Relative to an existing body of knowledge and methods for the acquisition of new knowledge, possible universes can be divided up, roughly, into three categories: (i) those which are such that the meta-methodology of AOE (even when modified) can meet with no success, not even apparent success, in the sense that new metaphysical ideas and associated methods for the improvement of knowledge cannot be put into practice so that success (or at least apparent success) is achieved; (ii) those which are such that AOE appears to be successful for a time, but this success is illusory, this being impossible to discover during the period of illusory success; and (iii) those which are such that AOE (perhaps when appropriately modified), can meet with genuine success. Meta-knowability asserts that our universe is a type (i) or (iii) universe; it rules out universes of type (ii).

Meta-knowability asserts, in short, that the universe is such that AOE can meet with success and will not lead us astray in a way which we cannot hope to discover by normal methods of scientific inquiry (as would be the case in a type (ii) universe). If we have good grounds for accepting meta-knowability as a part of scientific knowledge – grounds which do not appeal to the success of science – then we have good grounds for adopting and implementing AOE (from levels 5 to 2). Meta-knowability, if true, does not guarantee that AOE will be successful. Instead it guarantees that AOE will not meet with illusory success, the illusory character of this apparent success being such that it could not have been discovered by any means whatsoever before some date is reached.

If AOE lacks meta-knowability, its circular procedure, interpreted as one designed to procure knowledge to the extent that this is possible, becomes dramatically invalid, as the following consideration reveals. Corresponding to the succession of accepted fundamental physical theories developed from Newton down to today, there is a succession of severely disunified rivals which postulate that gravitation becomes a repulsive force from the beginning of 2150, let us say. Corresponding to these disunified theories there is a hierarchy of disunified versions of physicalism, all of which assert that there is an abrupt change in the laws of nature at 2150. The disunified theories, just as empirically successful as the theories we accept, render the disunified versions of physicalism just as scientifically fruitful as unified versions of physicalism are rendered by the unified theories we actually accept. The circularity inherent in AOE is invalid because it can be employed so as to lead to the adoption of *disunified* theories and metaphysical theses just as legitimately as it can be employed to lead to the adoption of *unified* theories and metaphysical theses. This is the case, at least, if AOE is bereft of meta-knowability. But if we have good reasons to accept meta-knowability as a part of scientific knowledge, then we have good reasons to reject *disunified* versions of physicalism: these clash with the level 6 thesis of meta-knowability. If we have good reasons to accept meta-knowability as an item of scientific knowledge, and these reasons make no appeal to the success of science, then the circularity inherent in AOE ceases to be invalid: meta-knowability asserts that the universe is such that empirical success achieved by implementing AOE will not be illusory *in a way which could not discovered by any means before a certain date*.

But what reasons have we for accepting meta-knowability that make no appeal to the success of science? What we have is an argument, not for the *truth* of meta-knowability, but rather for *accepting* meta-knowability as a part of scientific knowledge granted that our aim is to improve our knowledge of factual truth. We can argue that, as a result of accepting meta-knowability, we may have much to gain and little to lose. In accepting meta-knowability we decide, in effect, that it is worthwhile to try to improve knowledge about how to improve knowledge. We take seriously the possibility that the universe is such that we can discover something rather general about its nature which will enable us to *improve* our methods for improving knowledge. Not only do we hope to learn about the world; we hope to learn about how to learn about the world, and we are prepared to implement a meta-methodology which capitalizes on this possibility should it turn out to be actual. To fail to try to *improve* methods for improving knowledge on the grounds that apparent success might prove to be illusory is surely to proceed in a cripplingly over-cautious fashion. Any attempt at improving knowledge may unexpectedly fail, including the attempt to improve methods for improving knowledge. But eschewing the attempt to learn because it may fail cannot be sound: such an excuse for not making the attempt always exists. In accepting meta-knowability we do not assume, note, that the universe *is* such that AOE will meet with success. We assume, merely, that it is such that *if* AOE appears to meet with empirical success, this success will not be illusory in a way which could not have been discovered prior to the illusory character of the success becoming apparent. But this is an entirely sensible assumption to make. Nothing is to be gained from foregoing the attempt to acquire knowledge because of the fear that future, inherently unpredictable changes in the laws of nature may occur which render knowledge acquired obsolete.

Neither partial knowability at level 7 nor meta-knowability at level 6 excludes the possibility that such inherently unpredictable events occur. Even though we accept these theses, we might, nevertheless, still discover and accept that unpredictable changes in the laws of nature do occur (if they did occur). We might live, or come to live, in a world in which inherently inexplicable, unpredictable events occur quite often. Objects vanish, or abruptly appear; substances abruptly change their properties; bridges collapse, mountains vanish, houses turn into elephants, trees become daffodils. People die as a result, but life might nevertheless go on, and it might be possible, not just to improve knowledge, but to improve knowledge about how to improve knowledge. Meta-knowability asserts that, if we have had no such experience of them, such events do not occur. We are justified in ignoring the possibility that such events may occur in future in both science and life because, if they occur in the future nothing, in the nature of things, can be done to anticipate their occurrence, or evade the harm they may cause. It is this which provides the grounds for accepting meta-knowability as an item of scientific knowledge.

Hume, famously, argued that what exists at one moment cannot necessarily determine what exists at the next moment. If he is right, we may well feel that anything may happen at any moment – just because there can be nothing in existence now to *determine* (perhaps probabilistically) what will exist next. However, elsewhere I have shown that Hume is wrong, and it *is* possible that what exists at one instant necessarily determines what exists at the next moment (Maxwell, 1968; 1998, pp. 141-155). Since this is possible, it is, in my view, madness not to assume that what exists now *does* necessarily determine what exists next. Recognizing that Hume's arguments, here, are invalid is bound to affect ideas about how likely it is that utterly inexplicable, inherently unpredictable events will occur, as long as we do not seem to have had any experience of them.

Accepting meta-knowability, then, puts on record our decision to try to learn how to learn – to try to improve assumptions and associated methods in the light of improving knowledge and understanding, in the light of which seem best to promote empirical progress. This goes on, after all, in a thoroughly acknowledged and uncontroversial manner at the empirical level. New knowledge can give rise to new technology, new instruments and experimental techniques – from the telescope and microscope to the cyclotron – which are in turn employed to help create new knowledge. At the empirical level, uncontroversially and fruitfully, there is a kind of circular, positive feedback between improving knowledge and improving observational and experimental methods for the further improvement of knowledge. Something analogous has long gone on too, implicitly, in scientific practice, at the theoretical level. Science would be more rigorous, and even more successful, if this latter was explicitly recognized and acknowledged.

I have argued that we are justified in ignoring the possibility that apparent success achieved as a result of implementing AOE might turn out to be illusory in a way we could not possibly have discovered. Are we justified, however, in ignoring illusory apparent success of a less fiendish kind – apparent success which we could have discovered to have been illusory, if we had tried harder? Does not AOE always carry the danger that it will actively create the illusion of success – metaphysical assumptions and methods being chosen to promote the illusion of success in the pursuit of knowledge?

This brings me to my second point. AOE is better equipped to defeat this danger than any other rival methodology for science. That this is so provides decisive additional grounds for *rejecting* the objection that AOE suffers from invalid circularity.

Consider the best that any version of standard empiricism (any empiricist view which fails to acknowledge that scientific knowledge includes metaphysical theses) can do to defeat illusory success. First, accepted observational and experimental results can be subjected to sustained critical scrutiny. Experiments can be repeated in different laboratories by different scientists; and essentially the same experiment can be performed in different ways in an attempt to eliminate errors associated with one type of experiment. Second, accepted laws and theories can be severely tested, a variety of consequences being put to the test. Third, rival laws and theories can be developed in order to disclose crucial experiments which may falsify the accepted laws and theories, and which would not otherwise have been thought of: these crucial experiments can then be performed. These three standard empiricist procedures for detecting illusory empirical success are all important.

But AOE science can go further. In addition, it can subject the current best blueprint (at level 3) and associated methodological principles, to sustained critical scrutiny. It can actively seek to develop improved versions of this blueprint. It can even criticize and develop alternatives to metaphysical theses higher up in the hierarchy, at level 4, and even higher (see diagram). AOE comes with a framework that facilitates sustained critical scrutiny of current aims and methods, assumptions and methods; it provides meta-methodological machinery for the development of alternative possible aims and methods - alternative vantage points from which any illusory success of current aims and methods may be much more readily detected. Basic blueprint assumptions of a science do much to determine what kind of evidence is acceptable within that science. A change of blueprint may lead to a change in what constitutes acceptable evidence (which is what happened when Galileo rejected Aristotelianism). There is always the danger that a science seems to make great empirical success and fails to discover that this success is illusory because the evidence required to reveal this is declared illegitimate by the accepted blueprint. Thus the demand within physics that experimental result be repeatable prevents physics from discovering miracles – unique, unrepeated events – on empirical grounds. In order to discover the illusory character of such apparent empirical success it may be necessary to view matters from the standpoint of a modified blueprint, with modified standards for what constitutes an acceptable empirical result. AOE encourages the development of such modified blueprints, whereas standard empiricism does not even recognize the need for them. (Any view which specifies a fixed metaphysical assumption for science, on one level, is no better than standard empiricism in the respect just discussed.)

That AOE is better equipped to discover illusory empirical success than rival views provides a decisive rebuttal of the charge that there is an inherently invalid circularity in the manner in which AOE adjusts assumptions and methods in the light of empirical success and failure. On the contrary, AOE science is in a better position to detect such illusory success than science conducted in accordance with any rival view. AOE can modify its aims and methods, its assumptions and methods, in the direction of those which seem to produce the greatest empirical success – thus implementing something like positive feedback (and circularity). At the same time, AOE provides means for discovering when such apparent success is illusory in a way that is better, more effective, than any rival view.

I conclude that AOE is not invalidly circular. It accurately depicts methods in fact employed in scientific practice, and is a more rigorous conception of science than any rival view. It provides us with the proper framework within which to solve the problem of induction. Elsewhere I have gone further, and argued that the problem of induction can indeed be solved within the framework of AOE: see Maxwell (1998, chapter 5; 2004a, appendix, section 6, 2005b; 2007, chapter 14, section 6).

I have also argued elsewhere that AOE has dramatic and revolutionary implications for science and, when generalized, for academic inquiry as a whole: see Maxwell (1976; 1984; 1992; 2000; 2004a; 2007).

References

Feyerabend, P. (1978) *Against Method*, Verso, London.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*, Chicago University Press, Chicago.

Laudan, L. (1980) 'A Confutation of Convergent Realism', *Philosophy of Science 48*, pp. 19-48.

Maxwell, N. (1968) 'Can there be Necessary Connections between Successive Events?', *British Journal for the Philosophy of Science 19*, pp. 1-25.

_______ (1974) 'The Rationality of Scientific Discovery', *Philosophy of Science 41*, pp. 123-53 and 247-295.

_______ (1976) *What's Wrong With Science?*, Bran's Head Books, Frome, England.

_______ (1984) *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science*, Basil Blackwell, Oxford.

_______ (1992) 'What Kind of Inquiry Can Best Help Us Create a Good World?', *Science, Technology and Human Values 17*, pp. 205-27.

_______ (1993) 'Induction and Scientific Realism', *British Journal for the Philosophy of Science 44*, pp. 61-79, 81-101 and .275-305.

_______ (1998) *The Comprehensibility of the Universe: A New Conception of Science*, Oxford University Press, Oxford (pbk. 2003).

_______ (2000) 'Can Humanity Learn to become Civilized? The Crisis of Science without Civilization', *Journal of Applied Philosophy 17*, pp. 29-44.

_______ (2002) ‘The Need for a Revolution in the Philosophy of Science’, *Journal for General Philosophy of Science 33*, pp. 381-408.
_______ (2004a) *Is Science Neurotic?*, Imperial College Press, London.
_________ (2004b) ‘Scientific Metaphysics’, http://philsci-archive.pitt.edu/archive/00001674/
_______ (2005a) ‘Popper, Kuhn, Lakatos and Aim-Oriented Empiricism’, *Philosophia 32*, nos. 1-4, pp. 181-239.
_______ (2005b) ‘A Mug’s Game? Solving the Problem of Induction with Metaphysical Presuppositions’, http://philsci-archive.pitt.edu/archive/00002230/
_________, (2005c) ‘Practical Certainty and Cosmological Conjectures’, http://philsci-archive.pitt.edu/archive/00002259/
________ (2006) ‘A Priori Conjectural Knowledge in Physics: The Comprehensibility of the Universe’, http://philsci-archive.pitt.edu/archive/00002895/ (to be published in *New Perspectives on A Priori Knowledge and Naturalism*, ed. M. Shaffer and M. Veber, Open Court, La Salle.
________ (2007) *From Knowledge to Wisdom*, second edition, Pentire Press, London.
Popper, K. (1959) *The Logic of Scientific Discovery*, Hutchinson, London.
Van Fraassen, B. (1985) ‘Empiricism in the Philosophy of Science’, in P. M. Churchland and C. A. Hooker, eds., *Images of Science*, University of Chicago Press, Chicago.