Comprehensibility rather than Beauty

Nicholas Maxwell

(Text of talk given at “Aesthetics of Science” Conference, Dubrovnik, 9-14 April 2001)

Posted 27 May 2004

Email: nicholas.maxwell@ucl.ac.uk

Website: http://www.nick-maxwell.demon.co.uk

Abstract
Most scientists and philosophers of science recognize that, when it comes to accepting and rejecting theories in science, considerations that have to do with simplicity, unity, symmetry, elegance, beauty or explanatory power have an important role to play, in addition to empirical considerations. Until recently, however, no one has been able to give a satisfactory account of what simplicity (etc.) *is*, or how giving preference to simple theories is to be *justified*. But in the last few years, two different but related accounts have appeared, both of which address the above issues. On the one hand, James McAllister has argued that aesthetic criteria in science reflect scientists' judgements about what kind of theory is most likely to be empirically successful, based on the relative empirical success and failure of different kinds of theories in the past. Scientists employ what McAllister dubs "the aesthetic induction". On the other hand, I have argued that we need to see science as making a hierarchy of metaphysical assumptions about the comprehensibility and knowability of the universe, these assumptions asserting less and less as one ascends the hierarchy. One of the more substantial of these assumptions is that the universe is physically comprehensible. The key non-empirical feature a body of fundamental theories in physics must possess to be acceptable is unity. The better such a body of theory exemplifies the metaphysical thesis that the universe is physically comprehensible, in the sense that it has a unified dynamic structure, so the more acceptable such a body of theory is, from this standpoint. This affects not just theoretical physics, but the whole of natural science. In this paper I compare and contrast, and try to assess impartially the relative merits of, these two views.

1 Beauty or Comprehensibility?

Most scientists and philosophers of science acknowledge that aesthetic considerations play, quite properly, an important role in influencing acceptance and rejection of theories in science, in addition to empirical considerations. A famous example is Dirac, who went as far as to declare "It is more important to have beauty in one's equations than to have them fit experiment" (quoted in McAllister, 1996, 15).

The view that beauty ought to influence choice of theory in science faces, however, a serious problem. Why should beauty be a good indication of truth? Unless the truth is beautiful, and unless we have valid grounds for holding this to be the case, there can be no good reasons, it would seem, for giving preference to beautiful theories in science.

Not only may it seem dubious that we can have grounds for holding that the truth is beautiful; there may well seem to be grounds for holding that it is wildly implausible that the truth should be beautiful, especially in theoretical physics.

Whether we find something beautiful or ugly must depend, to some extent at least, on our personal, subjective, emotional responses to that thing. Aesthetic criteria have their roots deep in the human psyche, and in human culture. But physical reality, that which theoretical physics seeks to grasp, is utterly remote from the human psyche, from human culture. It may well seem utterly implausible that something as anthropomorphic, as personal, as quintessentially human, as ideas about beauty should have anything to do with the ultimate nature of the physical universe, utterly impersonal and remote from the circumstances of human life. Beauty may seem to be the last consideration to take into account in assessing the merits of rival fundamental theories in physics.

An extremely interesting and original defence of the thesis that aesthetic considerations do quite properly influence theory choice in science has, however, been put forward recently by James McAllister (1996): see also (1989), (1990) and (1991). Quite independently, I have, over a number of years, developed a view which resembles McAllister's view in a number of striking ways, but which is also different in important respects: see Maxwell (1972a, 1974, 1984, 1993b and especially 1998; for recent summaries see Maxwell, 1999a, 2000a; see also Maxwell, 2001a, and 2001b, chapter 3 and appendix 3, 2002a, 2002b, 2004, and Smart, 2000).

In this paper I compare and contrast the two views. I begin with a sketch, first of my view, then of McAllister's. I then discuss how they resemble, and differ from, each other. And finally I discuss the question of which is to be preferred.

2 Aim-Oriented Empiricism

According to the view I defend, which I call "aim-oriented empiricism" (AOE), science is obliged to make a big, persistent, metaphysical assumption about the nature of the universe. This assumption is implicit in those methods of science which specify that theories, in order to be accepted, must be sufficiently non-ad hoc, simple, unified or explanatory.

This claim is denied by a rather widely held view, which I call "standard empiricism" (SE), which asserts that, in science, all claims to knowledge are to be assessed impartially with respect to the evidence, *no thesis about the world being upheld permanently as a part of knowledge independently of evidence, let alone in violation of evidence*. Most, if not all, versions of SE stress that questions of simplicity, unity, beauty or explanatory power play a valid, important role in influencing choice of theory in science, in addition to considerations of empirical success - although some versions of SE give to simplicity, beauty or explanatory power much more important roles in science than other versions do. The decisive point that all versions of SE agree on is that no substantial thesis about the nature of the universe can be upheld as a part of scientific knowledge independently of empirical considerations, and certainly not in violation of empirical considerations. In so far as theory choice is biased in the direction of simplicity, unity, beauty or explanatoriness, this bias must not commit science to making the permanent assumption that nature herself is simple, unified, beautiful or explainable.

This thesis of SE is common ground for logical positivism, inductivism, logical empiricism, hypothetico-deductivism, conventionalism, constructive empiricism, pragmatism, realism, induction-to-the-best-explanationism, and the views of Popper and Kuhn.[1] McAllister too, as we shall see, defends a version of SE.

SE is, nevertheless, untenable, as the following considerations demonstrate.

Given any scientific theory, however well verified empirically, there will always be infinitely many rival theories which fit the available evidence just as well, but which make different predictions, in an arbitrary way, for phenomena not yet observed. Thus, given Newtonian theory (NT), one rival theory might assert: everything occurs as NT asserts up till midnight tonight when, abruptly, an inverse cube law of gravitation comes into operation. A second rival theory might assert: everything occurs as NT asserts, except for the case of any two 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. There is no limit to the number of rivals to NT that can be concocted in this way, each of which has all the predictive success of NT as far as observed phenomena are concerned but which makes different predictions for some as yet unobserved phenomena.[2] Such theories can even be concocted which are *more* empirically successful than NT, by arbitrarily modifying NT, in just this entirely *ad hoc* fashion, so that the theories yield correct predictions where NT does not, as in the case of the orbit of Mercury for example (which very slightly conflicts with NT).[3]

One can set out to refute these rival theories by making the relevant observations or experiments, but this needs an infinitely long time to complete as there are infinitely many rival theories to be refuted, each requiring a different refuting experiment. Thus, if science really did take seriously the idea that evidence alone decides what theories are to be accepted and rejected, scientific knowledge would be drowned in an infinite ocean of rival theories, all just as empirically successful as currently accepted theories, or actually even more successful empirically. Science would come to an end.[4]

Why does this not happen in scientific practice? Because, as most versions of SE stress, in practice *two* considerations govern acceptance and rejection of theories in science: (1) considerations of empirical success and failure; and (2) considerations that have to do with the simplicity, unity or explanatory power of the theories in question. In order to be accepted as a part of scientific knowledge, a theory must satisfy *both* considerations. It must be *both* empirically successful *and* simple, unified, or explanatory in character.[5]

Scientific theories that are accepted as a part of scientific knowledge do (more or less adequately) satisfy *both* considerations. They are both amazingly successful in their capacity to predict observable phenomena, and astonishingly simple, unified, explanatory.

But the infinitely many empirically successful rivals to these accepted theories all *fail* to satisfy the second consideration. They may fit all available evidence just as well as Newton's theory does, or Einstein's theories do: but they fail, quite drastically, to be simple, unified, explanatory. For these rival theories all assert that, for some as yet unobserved kind of phenomenon, something entirely peculiar and arbitrary occurs. Where NT assures us that gravitation obeys an inverse square law and is attractive uniformly everywhere, for all time, the aberrant rivals to NT assert that for some specific kind of phenomenon or range of phenomena gravitation obeys a quite different law, an inverse cube law perhaps, or one that asserts that gravitation is a repulsive rather than attractive force.

Thus the infinitely many rivals to accepted physical theories are rejected out of hand, not on empirical grounds, but because they are grotesquely *ad hoc*, grotesquely lacking in simplicity, unity, explanatory power.

This, then, is why in practice science is not buried beneath an infinite mountain of rival theories, all of which fit all available evidence equally well, if not better. Almost all the rivals are horribly complex, disunified, non-explanatory.

But now comes the decisive point. In persistently rejecting infinitely many such empirically successful but grotesquely *ad hoc* theories, science in effect makes a big permanent assumption about the nature of the universe, to the effect that it is such that no grotesquely *ad hoc* theory is true, however empirically successful it may appear to be for a time.[6] Without some such big assumption as this, the empirical method of science collapses. Science is drowned in an infinite ocean of empirically successful *ad hoc* theories.

All versions of SE are, in short, untenable.[7] In persistently rejecting, or just ignoring, empirically successful *ad hoc* theories, science commits itself to the assumption that the universe is such that no *ad hoc* theory is true; and, devoid of some such substantial assumption, science would self-destruct.

One might think that, before this non-*ad hoc* thesis (asserting that the universe is such that no *ad hoc* theory is true) can be accepted as a part of scientific knowledge, reasons must be given for holding that the thesis is true. But this completely misses the point. What the above argument has established is that the non-*ad hoc* thesis is implicit in current methods of science (and scientific knowledge would collapse if no such methods were adopted capable of excluding empirically successful *ad hoc* theories from science). The non-*ad hoc* thesis is substantial, influential, problematic, and implicit in current methods of science. But intellectual rigour demands that assumptions that have these features be made explicit, so that they can be criticized, so that alternatives can be developed and assessed. It follows that science can only claim to be intellectually rigorous in so far as it does make explicit the substantial, influential, problematic and, at present, implicit, assumption of non-*ad hocness*. A science that openly acknowledges that this assumption is a part of current scientific knowledge is more rational, more rigorous, than a science which disavows or represses the assumption, even though no grounds whatsoever are given for holding the assumption to be true.[8]

The moment it is recognized that persistent exclusion of empirically successful *ad hoc* theories from science commits science to making a substantial metaphysical assumption about the nature of the universe, two questions arise: What ought this assumption to be? How is it to be justified?

The solution to these fundamental epistemological problem of science (the very existence of which is denied by SE) is to construe science as adopting, as a part of scientific knowledge, a hierarchy of assumptions about the comprehensibility and knowability of the universe, these assumptions asserting less and less about the universe as one ascends the hierarchy, thus being more and more likely to be true: see diagram. Corresponding to these cosmological assumptions there are methodological rules (not represented in the diagram) which govern acceptance of assumptions lower down in the hierarchy, and which, together with empirical considerations, govern acceptance and rejection of scientific theories.

According to this view, then, scientific knowledge can be represented (in a highly schematic and simplifying way) as being made up of the following ten levels (see diagram). At level 1, we have empirical data (low level experimental laws). At level 2, we have our best fundamental physical theories, currently general relativity and the so-called standard model. At level 3, we have the best, currently available specific idea as to how the universe is physically comprehensible, the best available "blueprint" as I shall call this thesis. This asserts that everything is made of some specific kind of physical entity: corpuscle, point-particle, classical field, quantum field, convoluted space-time, quantum string field, or whatever. Because the thesis at this level is so specific, it is almost bound to be false (even if the universe is physically comprehensible in some way or other). Here, ideas evolve with

Level 7 Thesis that the universe is partially knowable (PK)	A	PK	A	Current scientific knowledge represented by:- PK, MK, C, P, B, T and Empirical Data
Level 6 Thesis that the universe is meta-knowable (methods	A	MK	A	At Levels 6 to 3 there are alternative metaphysical

being improvable) (MK)				theses compatible as far as possible with thesis above represented by A
<u>Level 5</u> Thesis that the universe is comprehensible (C)	A	C	A	Ideally T implies B which implies P which implies C which implies MK which implies PK, but in our present state of ignorance T clashes with B which clashes with P
<u>Level 4</u> Thesis that the universe is physically comprehensible (physicalism) (P)	A	P	A	
<u>Level 3</u> Blueprint: best current specific version of physicalism (B)	A	B	A	As one descends from Level 7 to 3, increasingly restrictive methodological principles are associated with each metaphysical thesis, restricting what is accepted lower down in the hierarchy
<u>Level 2</u> Accepted fundamental physical theories (T)		T		
<u>Level 1</u> Experimental and observational results	Empirical		Data	

Diagram: Aim-Oriented Empiricism

evolving knowledge. At level 4 we have the much less specific thesis that the universe is *physically* comprehensible in some way or other, a thesis which I shall call *physicalism*. According to physicalism, some kind of unchanging, unified physical entity, some kind of field unifying space-time and matter, exists at all times and places, throughout all phenomena, and determines, perhaps probabilistically, the way phenomena unfold. If physicalism is true, then some yet-to-be-discovered unified physical "theory of everything" is true. At level 6 we have the even less specific thesis that the universe is comprehensible in some way or other, there being *something*, God, tribe of gods, cosmic purpose, cosmic programme, kind of fundamental physical entity, which exists at all times and places in an unchanging form and determines (perhaps probabilistically) all change and diversity. At level 6 there is the even more unspecific thesis that the universe is at least nearly comprehensible in the sense that it is such that the best conjecture that science can adopt, at this level of generality, in order to promote empirical progress at levels 1 and 2, is that the universe is comprehensible. At level 7 there is the even more unspecific thesis that the universe is roughly comprehensible, in the sense that it is such that the best conjecture that science can adopt, at this level of generality, in order to promote empirical progress, is that the universe is partially comprehensible, there being, for example, three basic forces as opposed to one which determine the way phenomena unfold. At level 8 there is the even more unspecific thesis that the universe is such that there is some discoverable thesis which, if adopted, leads to improved methods for the improvement of knowledge. At level 9 there is the still more unspecific thesis that the universe is such that whatever makes it possible for us to acquire knowledge of our local circumstances exists at all times and places, so that local knowledge can be used to acquire some knowledge of what exists non-locally. And at level 10 there is the much less substantial thesis that the universe is such that we can acquire some knowledge of our local circumstances sufficient to make action possible.

The top two assumptions, at levels 10 and 9, are such that accepting these assumptions as a part of scientific knowledge can only aid, and can never damage science (or the task of acquiring knowledge more generally) whatever the universe may be like. These are justifiably permanent items of scientific knowledge.

As we descend, from level 8 to level 3, the corresponding theses make increasingly substantial assertions about the nature of the universe: it becomes increasingly likely that these theses are false. At each level, from 8 to 3, we adopt that assumption which (a) is a more precise version of the assumption above it in the hierarchy (in so far as this is possible), and (b) holds out the greatest hope for the growth of empirical knowledge, and seems best to support the growth of such knowledge (at levels 1 and 2). If currently adopted cosmological assumptions, and associated methods, fail to support the growth of empirical knowledge, or fail to do so as apparently successfully as rival assumptions and methods, then assumptions and associated methods are changed, at whatever level appears to be required.[9] In this way we give ourselves the best hope of making progress, of acquiring authentic knowledge, while at the same time minimizing the chances of being taken up the garden path, or being stuck in a cul de sac. The hope is that as we increase our knowledge about the world we improve the cosmological assumptions implicit in our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world - the methodological key to the astonishing, accelerating progress of modern science.

This conception of science, postulating more or less specific evolving aims and methods for science within a framework of more general fixed aims and methods, I call *aim-oriented empiricism*.[10] It is a special case of a more general idea of *aim-oriented rationality*, according to which, whenever basic aims are problematic (as they usually are in science and in life) we need to display aims at distinct levels of specificity and generality, thus creating a framework of unproblematic, fixed aims-and-methods within which we have the best chance of improving more or less specific, problematic aims-and-methods as we proceed, in the light of success and failure.[11]

Ideally, the thesis at level 2 implies the one at level 3, and so on up the hierarchy until one reaches level 9 or 10. This is true for levels 4 to 9. It breaks down dramatically, however, when we come to levels 2, 3 and 4. Fundamental theories currently accepted in physics, general relativity and the standard model, clash, and thus fail to exemplify physicalism. Furthermore, instead of postulating just one kind of self-interacting entity, the standard model postulates three kinds of forces, and many different kinds of particles with diverse properties, such as mass, that are not theoretically determined. All this is a sign of our ignorance (just as failure of theories to predict phenomena successfully is). What drives physics forward is the attempt to solve the problems that arise as a result of clashes between levels 1, 2, 3 and 4. According to AOE, a basic task of theoretical physics will have been completed when a level 2 theory has been discovered which (a) in principle predicts all (physically) possible level 1 phenomena, and (b) implies a true level 3 thesis, which (c) exemplifies (and thus implies) the level 4 thesis of physical comprehensibility (physicalism).

It may be objected that AOE is exclusively about theoretical physics, and cannot do justice to the variety of methods to be found in different branches of the natural sciences. In fact just the opposite is the case; AOE predicts diversity of method throughout natural science, overlaid by unity of method at a meta-methodological level. AOE can do justice to the diversity of methods to be found in diverse sciences, without underlying unity and rationality being sacrificed.

It is important to appreciate, first, that different branches of the natural sciences are not isolated from one another: they form an interconnected whole, from theoretical physics to molecular biology, neurology and the study of animal behaviour. Different branches of natural science, even different branches of a single science such as physics, chemistry or biology, have, at some level of specificity, different aims, and hence different methods. But at some level of generality all these branches of natural science have a common aim, and therefore common methods: to improve knowledge and understanding of the natural world. All (more or less explicitly) put AOE into practice, but because different scientific specialities have different specific aims, at the lower end of the hierarchy of methods different specialities have somewhat different methods, even though some more general methods are common to all the sciences. Furthermore, all natural sciences apart from theoretical physics presuppose and use results from other scientific specialities, as when chemistry presupposes atomic theory and quantum theory, and biology presupposes chemistry. The results of one science become a part of the presuppositions of another, implicit in the aims of the other science (equivalent to the level 3 blueprint of physics, or the level 4 thesis of physicalism). This further enhances unity throughout diversity, and helps explain the need for diversity of method.

But in order to exhibit the rationality of the diversity of method in natural science, apparent in the evolution of methods of a single science, and apparent as one moves, at a given time, from one scientific speciality to another, it is essential to adopt the meta-methodological, hierarchical standpoint of AOE, which alone enables one to depict methodological unity (high up in the hierarchy) throughout methodological diversity (low down in the hierarchy). Various versions of SE, lacking this hierarchical structure, cannot begin to do justice to this key feature of scientific method, diversity at one level, unity at another; nor can it begin to do justice to the rational *need* for this feature of scientific method.

Elsewhere I have argued at some length that AOE solves a number of outstanding problems about the nature of science, such as the problem of induction, the problem of specifying precisely the nature of scientific method (just touched upon), the problem of verisimilitude, and the problem of how new fundamental physical theories can be discovered (see Maxwell, 1998, chapters 4-6). Here, I indicate how AOE solves another key problem: what it means to assert of a theory that it is simple, unified, explanatory, elegant, harmonious or beautiful.

According to AOE, the key idea is explanatory power, or unity. The totality of fundamental physical theory, T, is unified to the extent that its *content* exemplifies physicalism. The more the content of T departs from exemplifying physicalism, the more disunified T is.[12] Because what matters is content, not form, the way T is *formulated* is irrelevant to this way of assessing simplicity or unity. No version of SE can avail itself of this way of assessing unity because it involves acknowledging that physicalism is a basic tenet of scientific knowledge, something which SE denies. Within AOE, there is a second way in which the unity of T may be assessed: in terms of the extent to which the content of T exemplifies the best available level 3 metaphysical blueprint. This second conception of simplicity or unity evolves with the evolution of level 3 ideas. As we improve our ideas about how the universe is unified, with the advance of knowledge in theoretical physics, so non-empirical methods for selecting theories on the basis of simplicity or unity improve as well. Thus current symmetry principles of modern physics, such as Lorentz invariance and gauge invariance, which guide acceptance of theory, are an advance over simplicity criteria upheld by Newton.

This account of simplicity can be extended to individual theories in two ways. First, we may treat an individual theory as a candidate theory of everything. Second, given two individual theories, T_1 and T_2, and given the rest of fundamental theory, T, T_1 is simpler than T_2 iff T + T_1 is simpler than T + T_2, where the latter is assessed in one or other of the ways indicated above.[13]

It may be objected that this proposed solution to the problem of simplicity is circular: the unity of level 2 theory is explicated in terms of the unity of level 4 physicalism. But this objection is not valid. In order to solve the problem, it is not necessary to explicate what "simplicity" or "unity" mean; rather, what needs to be done is to show how theories can be partially ordered with respect to "simplicity" or "unity" in a way that does not depend on formulation. This is achieved by partially ordering theories in terms of how well their content exemplifies the content of physicalism, so that, roughly, the more the content of a theory violates the symmetries associated with the content of physicalism, the less unity it has. As long as physicalism is a meaningful thesis, and provides a formulation-independent way of partially ordering theories in the way indicated, this suffices to solve the problem. That physicalism embodies intuitive ideas of "unity" is a bonus. For a more detailed rebuttal of this objection, see Maxwell (1998), 118-23.

But how, it may be asked, does this account provide a basis for partially ordering theories with respect to unity? How can degrees of unity be assigned to theories?

This can be done as follows. A collection of level 2 theories, T, may clash with physicalism and yet exemplify physicalism to some degree, in that it is disunified to some degree in one or more of the following eight ways of being disunified.

(1) T has a different *content* in the N different space-time regions, $R_1,...R_N$,

(2) T postulates that, for distinct ranges of physical variables, such as mass or relative velocity, in distinct regions, $R_1,...R_N$ of the space of all possible phenomena, distinct dynamical laws obtain.

(3) T postulates, in an arbitrary fashion, N distinct, unique, spatially restricted objects, each with its own distinct, unique dynamic properties.

(4) T postulates N different kinds of physical entity, differing with respect to some dynamic property, such as value of mass or charge, and interacting by means of different forces.

(5) As in (4) except the distinct kinds of physical entity interact by means of the same force.

(6) Consider a theory, T, that postulates N distinct kinds of entity (e.g. particles or fields), but these N kinds of entity can be regarded as arising because T exhibits some symmetry. If the symmetry group, G, is not a direct product of subgroups, we can declare that T is fully unified; if G is a direct product of subgroups, T lacks full unity; and if the N entities are such that they cannot be regarded as arising as a result of some symmetry of T, with some group structure G, then T is disunified.

The way in which relativistic classical electromagnetism unifies the electric and magnetic fields is an example of this kind of unity. Given the electric field, then the magnetic field must be adjoined to it if the theory is to exhibit the symmetry of Lorentz invariance. Again, the way in which chromodynamics brings unity to the eight gluons, and to quarks that differ with respect to colour charge, postulated by the theory, provides another example of this kind of unity. The diverse gluons and colour charged quarks of the theory are required to exist if the theory is to have its distinctive locally gauge invariant character, in this case the symmetry group being SU(3). The electroweak theory of Salam and Weinberg is an example of partial unity of this type, in that, in this case, the symmetry group, corresponding to the locally gauge invariant character of the theory, is SU(2) X U(1) - a group that is a direct product of subgroups. The theory only partially unifies the diverse quanta of the associated fields, the photon of electromagnetism and the vector bosons of the weak force.[14]

(7) If (apparent) disunity has emerged as a result of a series of cosmic spontaneous symmetry-breaking events, there being manifest unify before these occurred, then the relevant theory, T, is unified. If current (apparent) disunity has not emerged from unity in this way, as a result of spontaneous symmetry-breaking, then the relevant theory, T, is disunified.

(8) According to general relativity, the force of gravitation is merely an aspect of the curvature of space-time. As a result of a change in our ideas about the nature of space-time, so that its geometric properties become dynamic, a physical force disappears, or becomes unified with space-time. This suggests the following requirement for unity: space-time on the one hand, and physical particles and forces on the other, must be unified into a single self-interacting entity. If T postulates space-time and physical "particles and forces" as two fundamentally distinct kinds of entities, then T is not unified in this respect.

We have here, then, eight *different* ways in which the totality of fundamental physical theory can exemplify physicalism to some degree N (with N = 1 for unity). The most severe kind of disunity is that specified by (1); (2) and (3) specify slightly less severe kinds of disunity, (4) and (5) less severe kinds of disunity still, and (8) specifies the least severe kind of disunity of all. (1) to (8) are to be understood as indicating eight different kinds of degrees of disunity, but not as *defining* disunity.

Analogously, T may clash with a blueprint, B, and yet exemplify B to some degree, in that it postulates B-type entities, forces and symmetries, but at the same time violates, to some degree, and in one or more ways, the specific kind of unity postulated by B.

This account of simplicity, or unity, can be extended so that it applies, not just to fundamental dynamical theories of physics, but to theories, and even low-level empirical laws, of the whole of natural science. In order to do this one simply needs to take note of the point made above that any branch of natural science other than fundamental physical theory invariably presupposes some part, P, of an explanatorily more fundamental science, and then proceed as above, P now having the role of physicalism, or the best blueprint. In this way, the above account does justice to persisting simplicity criteria relevant to the whole of natural science (stemming from acceptance of physicalism), *and* does justice to simplicity criteria that evolve with time (stemming from the evolving best blueprint), and to criteria that change, at a given time, as one moves from one branch of natural science to another (stemming from parts, P, of different, more fundamental sciences being presupposed).[15]

How is acceptance of physicalism and the other metaphysical theses in the hierarchy of AOE to be justified? Here I can only sketch my answer; for a more detailed response see Maxwell (1998), especially chapter 5.

The first point to note is that even our most trivial, everyday claims to factual knowledge contain implicit factual claims about the entire universe, and the ultimate nature of the universe. Thus, in claiming that I know I can walk across the room I am, implicitly, claiming to know that nowhere in the universe is there now occurring an explosion of chaos which will travel at almost infinite speed to engulf the earth, the room, and me before I can take a step. In claiming that *I* have the power to decide to walk across the room I am claiming, implicitly, that the ultimate nature of the universe is such that free will, in some meaningful sense, is both possible, and actual for me. Granted that even our most trivial common sense claims to knowledge contain cosmological and metaphysical dimensions, it should occasion no surprise that far more contentful scientific claims to knowledge do as well.

I have already argued that AOE, with its hierarchy of metaphysical theses at levels 3 to 10, is more rigorous than any version of SE, which would depict scientific knowledge as existing only at levels 1 and 2, everything above 2 being merely speculation to be considered in the context of discovery only. The AOE picture makes explicit and so criticizable and revisable assumptions that are substantial, influential and problematic, but only implicit granted SE. This in itself makes AOE more rigorous than SE, and provides a kind of justification for accepting metaphysical theses at levels 3, 4, and above.

SE tends to depict the transition from pre-science to science as the process of excluding metaphysical theses from science, only testable theories being candidates for scientific acceptance. But it is precisely this way of construing this transition that generates the unrigorous conception of science of SE, a symptom of this lack of rigour being the failure to solve the problem of induction - the problem of how theories can be confirmed by evidence. It is this, indeed, that *creates* the unsolvable problem of induction. What we need to do, instead, is to construe the transition from pre-science to science as involving, ideally, not the ejection of untestable metaphysics from scientific knowledge altogether, but instead the articulation of implicit metaphysics as explicit theses, and then the selection of those metaphysical theses which either (a) the search for knowledge cannot do without, or (b) appear to be the most fruitful in leading to the growth of empirical knowledge (at levels 1 and 2). Above all, we need to organize these metaphysical theses into the hierarchical structure of AOE, thus creating a framework of relatively unchallengeable and unproblematic assumptions within which much more specific and problematic assumptions, lower down in the hierarchy, can be revised as empirical success and failure seem to require. In this way we can focus criticism on that part of the metaphysical presuppositions of science which, we conjecture, it is the most fruitful to criticize from the standpoint of achieving progress in scientific knowledge and understanding.

In other words, the fundamental epistemological problem of science is not the problem of induction, nor even the problem of justifying the truth of metaphysical assumptions made by science. It is, rather, the problem of showing that, from diverse metaphysical, cosmological assumptions that science might make, those that science actually selects are either indispensable for any attempt at acquiring knowledge, or are more fruitful for empirical progress than any rival assumptions.

The level 4 thesis of physicalism is to be accepted as a part of scientific knowledge, in short, because, at this level, and within the general framework of AOE, there is no other thesis that has proved as fruitful in promoting scientific progress at levels 1 and 2.

But is this correct? Is it not refuted by Kuhn's point (Kuhn, 1970, chapter 13) that nothing theoretical survives a revolution, the new theory or paradigm being incommensurable with the old one?

This Kuhnian view, if valid at all, is most likely to be correct when applied to revolutions in fundamental theoretical physics, where radical discontinuity seems most marked. Ironically enough, it is above all here that Kuhn's claim fails. All revolutions in theoretical physics, despite their diversity in other respects, reveal one common theme: they are all gigantic steps in unification. Thus Newton unifies Kepler and Galileo. Maxwell's theory of the electromagnetic field unifies electricity, magnetism and optics. Quantum theory unifies chemistry, properties of matter, and ultimately, with the development of quantum electrodynamics, electromagnetic phenomena. General relativity unifies special relativity, gravitation and the structure of space-time. Quantum electroweak theory (partially) unifies the electromagnetic and weak forces. The so-called standard model (partially) unifies all known phenomena apart from gravitation. String theory, or M theory, if successful, will unify all phenomena. The very phenomenon that Kuhn holds to mark discontinuity, namely revolution, actually also reveals continuity - continuity of the search for, and the successful discovery of, underlying theoretical unity. Revolutionary developments in theoretical physics all reveal one common theme: the increasingly successful capture of physicalism, more and more adequately, as a single, precise, unified, testable, physical "theory of everything". Almost the whole of theoretical physics since Galileo substantiates the claim that physicalism is by far the most fruitful idea that science has come up with at that level in the hierarchy of assumptions. The whole way in which theoretical physics has developed points at physicalism.

But in order to appreciate this point, it is essential to adopt a generalized hierarchical view, of which AOE is a special case. The historical record may reveal discontinuity at levels 2 and 3; we need to recognize level 4 to appreciate continuity through this lower-level discontinuity. Indeed, if we take pre-Galilean, Aristotelian science into account, we would need to ascend to level 5 to see continuity through the Galilean revolution. As a result of restricting himself to levels 1 and 2, and perhaps level 3, Kuhn was unable to discern theoretical continuity across the discontinuity of revolutions (although this is manifest, even for Kuhn, in revolutions other than those in theoretical physics).

One criticism that may be levelled against AOE is that it does not just accurately *reflect* scientific practice, but has the audacity to claim to *correct* scientific practice. It does this by providing a framework for the articulation and scrutiny of level 3 metaphysical blueprints, as an integral part of science itself, thus providing a rational, if fallible, means for the development of new non-empirical methods, new symmetry principles, and new theories. AOE makes explicit what is at present only implicit, due to the current influence of SE on the scientific community. And more generally, AOE has implications for scientific practice throughout the natural sciences in depicting scientific method in a hierarchical, meta-methodological fashion. Does this not tell against AOE? No. Any new conception of science which substantially improves our understanding of science ought to enable us to improve scientific practice. It would be very odd if our ability to do science well were wholly divorced from our understanding of what we are doing. A test for a new theory of scientific method ought to be, then, that it improves scientific practice, and does not merely accurately depict current practice. AOE passes this test.

In case it should seem miraculous that science has made progress without AOE being generally understood and accepted, I should add that good science has always put something close to AOE into practice in an implicit, somewhat covert way, and it is this which has made progress possible.

3 The Model of the Aesthetic Induction

I turn now to James McAllister's account of the role of non-empirical, aesthetic factors in the selection of theories in science. I shall call McAllister's view "the model of the aesthetic induction" (MAI). Here, in summary, is his view.

According to MAI, the basic aim of science is to develop a body of theory that successfully predicts all observable phenomena. MAI holds that from this aim of "empirical adequacy", we can arrive at the following criteria for assessing theories: success in predicting existing empirical data, success in predicting new phenomena, consistency with other high-level theories, explanatory power, and internal consistency.

Many scientists have however declared that aesthetic considerations, in addition to the above, play a vital role in both the discovery and acceptance of theories in science. Dirac, Einstein, and many others have stressed the importance of aesthetic considerations, such as beauty, elegance, harmony, uniformity amidst variety, simplicity, symmetry. MAI holds that such criteria do indeed have an important role to play in deciding what theories are accepted, to the extent, even, on occasions, of over-riding empirical considerations.

But, according to MAI, in so far as such aesthetic considerations exercise a rational influence over choice of theory in science, two crucial points need to be borne in mind. First, theories must be considered to be abstract entities, distinct from this or that linguistic formulation. Second, what matters is not the (subjective) aesthetic judgements themselves, but rather objective, non-aesthetic properties that theories, construed as abstract entities, do actually possess, in virtue of which scientists make their aesthetic judgements.

There are, according to MAI, five classes of properties of theories that are relevant: symmetry, invocation of a model, visualizability/abstractness, metaphysical allegiance, and simplicity (related to unity). MAI stresses that many different properties fall under each of these headings. There are different kinds of symmetry; different theories have different kinds of models; some scientists, in some contexts, hold visualizability to be a virtue, while others, in other contexts, prize almost its opposite, namely abstractness; scientists have upheld different metaphysical views at different stages in the development of science, in terms of which they have sought to interpret scientific theories; and there are many different ways of assessing the simplicity of theories, yielding quite different results.

How, then, does the scientific community decide which of these very many different kinds of properties of theories are the relevant or important one's to employ in order to assess the acceptability of theories on non-empirical, or aesthetic grounds? And what is the justification for so assessing theories, in terms of the preferred properties? How, in particular, can MAI do justice to the fact that aesthetic criteria in science change over time?

The answer is that, at any given stage, a scientific community prefers those new theories which have properties which earlier theories, which have proved to be empirically successful, also possess. If a certain kind of theory, with characteristic aesthetic properties, has met with empirical success in the past then, understandably enough, scientists are influenced to give preference to similar kinds of theories in the future, with similar properties. This is "the aesthetic induction". At a stroke, the above three questions are answered.

In a little more detail, we can imagine that a scientific community can consider many different aesthetic properties of theories, P, Q. R,.... The community will assign a different weighting, W_P, W_Q, W_R, to each of these properties, each weighting determining how influential the corresponding property is in theory choice. The weightings are in turn determined by what kinds of theory, with what properties, have (or have not) met with empirical success in the past. W_P, W_Q, W_R... are, in other words, determined by the aesthetic induction.

According to MAI, then, two kinds of criteria are employed in science to choose theories. On the one hand there are criteria, listed above, arrived at by analysis of the basic aim of science of achieving empirical adequacy. And on the other hand, there are criteria arrived at by the aesthetic induction. The second presupposes the first.

Aesthetic criteria will tend to be conservative, based as they are on empirical performance of theories in the past. New theories, with the potential for great predictive success, may violate existing, conservative aesthetic criteria. When such a theory is developed, there is a rupture in accepted aesthetic criteria. Initially the new theory is judged to be "ugly"; but as its empirical potential becomes manifest, aesthetic criteria are changed to suit the new theory. This is what a scientific revolution amounts to, according to MAI, a conception somewhat different from Kuhn's. In terms of this conception, neither Copernicus's theory, nor Einstein's theory of special relativity, were revolutionary, because neither broke with aesthetic criteria of the past. But Kepler's laws of planetary motion, and quantum theory, were both revolutionary, in that these theories broke dramatically with aesthetic criteria generally accepted at the time.

Finally, though the aesthetic induction might one day favour some particular metaphysical world view, so far this has not happened (McAllister, 1996, 102-4).

4 Comparison of the Two Views

What is rather astonishing about AOE and MAI is that, though arrived at independently, and though giving what are, in some respects, very different pictures of the scientific enterprise, nevertheless the two views have much in common. Both seek to uphold what McAllister calls "the rationalist image of science". Both hold that (some) criteria of theory choice can be justified by an appeal to the aims of science. Both hold that non-empirical criteria of theory choice have an enormously important part to play in science. Both hold that these non-empirical criteria are, in practical applications, quite diverse in character. Both hold that they change over time, as science progresses. And there is considerable agreement as to what these non-empirical criteria are: simplicity, unity, symmetry, and compatibility with some metaphysical world view, are all important, for both views.

Both hold that these criteria apply, not to specific formulations of theories, but to what all possible formulations have in common. And both give accounts of scientific revolutions that differ substantially from Kuhn's account.

But there are also dramatic differences. MAI is, for McAllister, "a medium-level model of scientific practice, of a scope intermediate between the loftiest generalization and the historical study" (1996, 2). AOE is put forward as a "highest-level model", with implications and applications for all of natural science. (Strictly speaking, it is what I call "generalized AOE" [Maxwell, 1998, 191-2, 185, 191, 208 and 223-4], embodying the hierarchical structure of AOE, but lacking specific, lower level theses of AOE, that is a model at the highest level, applying to science throughout history; AOE is restricted to post-Galilean science.)

Again, MAI is a version of SE, whereas AOE emphatically rejects SE. That MAI is a version of SE is clear from the way the aim of science is characterized as "empirical adequacy". It is also apparent in the way science can, according to MAI, establish a metaphysical world view. This can only happen via the aesthetic induction, and has not as yet come about. According to AOE, by contrast, at levels 10 and 9 there are metaphysical, cosmological theses that are permanently accepted by science, and at levels 8 to 3, there are metaphysical theses which are a part of current scientific knowledge, but which are increasingly likely to require revision with the advance of science, as one descends from level 8 to level 3.

Whereas MAI gives to science just one aim (empirical adequacy), AOE sees science as having a hierarchy of aims, from empirical adequacy, perhaps, at the highest level, down to the aim to turn the best available level 3 blueprint into a precise, true theory of everything, at the lowest level. (And even more specific, and different, aims are assigned to different branches of natural science.)

That AOE postulates this hierarchical structure to the aims of science, whereas MAI does not, leads to different treatments of changing criteria for theory choice. According to MAI, criteria of theory choice are of two kinds: those that are justified by an appeal to the basic aim of science (empirical adequacy), and those that are justified by inductive projection - the aesthetic criteria arrived at by the aesthetic induction. These latter are weaker than the former, and presuppose, for logical reasons, the former (McAllister, 1996, 76). According to AOE, by contrast, all criteria of theory choice are arrived at by aim analysis: those that evolve do so because the level 3 aim of science evolves.

Even though MAI and AOE agree that non-empirical criteria of theory choice change with time, they disagree about what criteria change, and what this change involves. According to AOE, something close to physicalism has been implicit in the methods of theoretical physics since Galileo or Newton; the demand for theoretical unity, associated with physicalism, has been more or less unchanging. What has changed is the form that the demand for unity takes, as manifest in dramatically changing level 3 metaphysical blueprints. MAI does not claim that physicalism, and the requirement of unity associated with it, is a part of the unchanging criteria of theory choice (since Galileo, at least). Nor could MAI claim this, as long as it is a version of SE.

According to AOE, the level 4 thesis of physicalism, and the level 3 best metaphysical blueprint, are arrived at by a quasi-Popperian process of conjecture and criticism, the whole direction of progress in theoretical physics being taken into account since the birth of modern science (or since the Presocratics). The claim is that these theses make explicit what theoretical physics hopes to achieve: they are intended to be the best conjectures as to what the basic aims of theoretical physics should be, at different levels of specificity. These theses are intended to lead to criteria, to methodological principles such as symmetry principles, that will be relevant for future theories, not yet developed. Indeed, according to AOE, the activity of further articulating the best blueprint, and solving problems of unity to which it gives rise, provides science with a rational, if fallible method of discovery. All this contrasts dramatically with criteria arrived at by the aesthetic induction, according to MAI, which are almost bound to be conservative, and more or less inapplicable to revolutionary developments. AOE criteria anticipate and provoke revolution, and judge the existing body of fundamental physical theory as unsatisfactory because of its failure to comply with the demand for unity (the standard model postulates too many particles and forces, and clashes with general relativity); by contrast, MAI criteria are conservative, and are almost bound to be at odds with revolutionary developments (McAllister, 1996, 81-5 and 128-33). AOE criteria are heuristically powerful; MAI criteria are the opposite. Furthermore, AOE criteria, associated with level 3 blueprints, evolve or improve as physics makes progress, and in a way which admits some elements of continuity: see, in particular, Maxwell (1998), 80-9. MAI sees change, but no overall progress, in non-empirical criteria, and holds, in a quasi-Kuhnian fashion, that revolutions create a rupture in aesthetic criteria, there being no account of the modification and generalization of blueprints, which AOE provides.

MAI and AOE agree that non-empirical criteria apply, not to any specific formulation of a theory, but to what all formulations have in common. But there are somewhat different accounts of what this is. According to MAI, a formulation-independent theory is an abstract entity that exists in its own right, with its own properties distinct from the phenomena the theory postulates (see, for example, McAllister, 1996, 98-100). This leaves obscure what sort of thing such an abstract entity is, and what its relationship is with a linguistic formulation of the theory, and with the phenomena it predicts. According to AOE, the matter is much more straightforward: a formulation-independent theory, T, is the *content* of T, what T predicts, or asserts to be the case. AOE does not appeal to the abstract entities of MAI; it appeals only to possible phenomena, not as actually existing entities, but merely as possibilities. The claim that T exhibits a certain symmetry thus amounts to the claim that the phenomena predicted by T exhibit this symmetry. There is here no mystery about the relationship between a linguistic formulation of T, the abstract entity T, and the phenomena T predicts: the "abstract entity" is just what any linguistic formulation of T asserts to be the case, the *content* of T. This leads to an account of the importance of linguistic-dependent criteria of simplicity (Maxwell, 1998, 110-3), something which MAI does not provide.

A fundamental difference between MAI and AOE, encapsulated in the title of this paper, is that, whereas MAI holds that aesthetic criteria are important in science, AOE denies this, all non-empirical criteria for theory choice being reducible to the demand that the totality of fundamental physical theory exemplify the level 4 thesis of physicalism or, more specifically, the best available blueprint at level 3. For AOE, what matters is unity or comprehensibility, not beauty.

But this difference is terminological rather than substantial. McAllister defends a projectivist, subjectivist account of the beauty of theories. Beauty is in the eye of the beholder, rather than in the theory itself.[16] Scientists judge certain theories to be beautiful because of non-aesthetic properties that they possess objectively; it is these non-aesthetic properties that are important methodologically and epistemologically, and play the crucial role in the aesthetic induction. One of these is metaphysical allegiance. The demand that the totality of fundamental physical theory should exemplify physicalism, and the best available blueprint, are special cases of metaphysical allegiance. Comprehensibility is just one of McAllister's aesthetic properties. Comprehensibility, one might say, is beautiful. It fits perfectly Hutchinson's characterization of beauty (McAllister, 1996, 17-23) as involving "uniformity amidst variety": see the discussion of "unity through diversity" in Maxwell (1998), chapter 3.

A more serious disagreement would seem to be that whereas AOE recognizes only *one* methodologically significant non-empirical property, namely unity or comprehensibility, MAI stresses that there are endlessly many, falling under the five headings of symmetry, invocation of a model, visualizability/ abstractness, metaphysical allegiance, and simplicity (related to unity).

This disagreement is not quite as big as it might at first appear to be. Here, very briefly, are the similar, but also different, ways in which AOE and MAI treat unity, symmetry, metaphysical allegiance and simplicity.

Unity. AOE and MAI both recognize that the demand for unity takes a number of different forms, but AOE alone holds that these are aspects of just one, single conception of unity. According to AOE, dynamic unity, postulated to exist by physicalism, can be broken in thought in a number of different ways, this creating a number of different kinds of (relative) disunity, and hence a number of different ways in which degrees of unity (or disunity) can be assessed. But these different kinds of disunity all relate to just one conception of unity, namely that which is postulated to exist by physicalism: see Maxwell (1998), 89-93, and 280 note 22. MAI too stresses that there are different kinds of unification (McAllister, 1996, 110) but, unlike AOE, does not relate these to one basic conception of unity.

Symmetry. Here, again, AOE and MAI both recognize that the demand for symmetry takes a number of different forms, but AOE alone holds that these, in so far as they are methodologically legitimate within theoretical physics, all relate to the one basic demand for unity. One of the achievements of AOE is to demonstrate clearly how different kinds of symmetry relate to unity, the demand that theories exhibit symmetries itself being an aspect of the demand for unity (Maxwell, 1998, 89-103, 123-40 and 257-65). MAI recognizes that theories exhibit different kinds of symmetry (McAllister, 1996, 41-4), but fails to recognize that different kinds of symmetry, in theoretical physics at least, are aspects of unity.

Metaphysical Allegiance. Once again, both AOE and MAI recognize that an important non-empirical requirement in theoretical physics, upheld by some physicists at least, is that fundamental physical theories should accord with some metaphysical view. Both recognize that metaphysical views associated with physics since Galileo have changed dramatically over time; both recognize that different physicists espouse different metaphysical views at the same time. But AOE and MAI differ, here too, in that AOE holds that diverse, evolving level 3 blueprints, in order to be acceptable, need to accord with physicalism, whereas MAI makes no such demand. For AOE, the requirement that a theory exemplifies a metaphysical view, in so far as it is methodologically legitimate, is but an aspect of the basic requirement that the body of fundamental physical theory exemplifies the unity of physicalism (as much as possible). MAI makes no such demand.

Simplicity. Here, yet again, both AOE and MAI recognize that the demand for simplicity takes a number of different forms; both see simplicity as being related to unity, but in somewhat different ways: compare Maxwell, 1998, 111-3 and 157-9, with McAllister, 1996, 109-11. But AOE alone relates the demand for simplicity to the more basic demand for just one kind of unity, dynamic unity postulated by physicalism.

The difference that this reveals in the two views can be summed up like this. AOE postulates just one basic non-empirical requirement, unity, and relates different requirements, of different kinds of unity, symmetry, metaphysical allegiance and simplicity, to this one demand for unity. MAI, by contrast, holds that there just are many different kinds of requirements of different kinds of unity, symmetry, metaphysical allegiance and simplicity. Unlike AOE, MAI sees no unity in diverse kinds of unity, symmetry, metaphysics and simplicity. (In this respect, AOE might be said to give a more unified, and hence more beautiful, account of scientific method than MAI.)

More substantial differences arise in connection with the two remaining kinds of aesthetic properties of theories which MAI holds to be methodologically significant, which I now consider in turn.

Invocation of a Model. AOE recognizes that an important consideration in assessing a new physical theory is that it has a form similar to existing empirically successful physical theories. Thus the acceptability of quantum electroweak theory, and chromodynamics is much helped by the fact that these theories are similar in form to the highly empirically successful theory of quantum electrodynamics. All three theories, despite their differences, are locally gauge invariant quantum field theories. According to AOE, this requirement of similarity of form or structure derives, once again, from the requirement of unity (Maxwell, 1998, 112). If T_1 and T_2 have some similar structure, then some part of T_1 can be modelled by some part of T_2, and vice versa. According to AOE, having a model is only methodologically significant to this extent, and once again this requirement turns out to be derived from the demand for unity. (Of course, that physical reality is a model of a theory, in the sense that the theory is *true*, is highly significant for AOE; but this is not what MAI means by a "model".) MAI is, once again, much more open-ended in the kind of models that it is prepared to recognize as methodologically significant, and does not attempt to derive the requirement that a theory should have some kind of model from the demand for unity.

Visualizability/abstractness. According to AOE, neither visualizability nor abstractness are methodologically significant for theoretical physics. What does matter is that a theory can at least be interpreted realistically, as postulating that such and such a physical entity, (or entities), such as a field (or particles) exists, a stepping stone towards the ubiquitous, unified *something* of physicalism. (Actually, AOE demands more. It demands that fundamental physical theories are open to being interpreted in terms of conjectural essentialism: see Maxwell [1998], 141-55.) If one has acquired an intuitive understanding of a realistic theory, then one may well be able to "visualize" what the theory is about: to this extent, visualizability is methodologically significant, according to AOE, but once again derives from the demand for unity, via the demand for realism. MAI, by contrast, once again, is much more open-ended about visualizability, and makes no attempt to relate it to the demand for unity.

McAllister claims that opposition to orthodox quantum theory (OQT), by Schrödinger, Einstein and others, stemmed from the loss of visualizability and determinism associated with the new theory. But this overlooks the key, entirely legitimate objection to OQT, namely its loss of microrealism, due to the failure to solve the quantum wave/particle problem. Because it failed to specify a consistent quantum ontology, OQT had to be developed as a theory which can, at most, make predictions about the results of performing measurements on quantum systems - measurement being described classically. But this in turn meant that OQT is, quite essentially, made up of two quite different parts stitched together in a grossly *ad hoc* way, namely (1) the quantum part, and (2) some part of classical physics for a treatment of measurement. Despite its immense empirical success, OQT is still today deeply and genuinely problematic, to the point, almost, of being unacceptable, because of its grossly *ad hoc* character, due to its lack of microrealism (Maxwell, 1972b, 1976, 1982, 1988, 1998 chapter 7). The mature Einstein was well aware that this is the basic objection to OQT, not lack of visualizability or loss of determinism (Maxwell, 1993b, 290-6). Elsewhere I have argued that the grossly *ad hoc* character of OQT, stemming from its lack of microrealism, provides us with a general argument against instrumentalism and for realism (Maxwell, 1993a). I have also suggested how the quantum wave/particle problem may be solved, and how a fully microrealistic version of quantum theory may be developed, free of any reference to measurement or classical physics in its basic postulates, able to recover all the successful predictive content of OQT, but also making experimental predictions different from OQT for as yet unperformed experiments (Maxwell, 1976, 1982, 1988, 1998 chapter 7, and especially 1994). (This was done in an attempt to put the rational, but fallible, method of discovery of AOE into scientific practice.) There are, of course, other attempts at developing fully microrealistic versions of quantum theory: see Bohm (1952), Ghirardi, Rimini and Weber (1986), and Penrose (1986).

Einstein's mature objection to OQT had to do with the lack of realism of the theory, but he did also, especially earlier, object to its lack of determinism. But here, too, there is a methodologically significant issue at stake, related once again to the demand for unity. A realistic version of quantum theory must be unified, first with special relativity, and then, ultimately, with general relativity. This is a much graver problem, granted probabilistic quantum theory, than it is if quantum theory is deterministic. The demand for unity speaks against probabilistic quantum theory - but not decisively: nature may well be probabilistic, and the task may be to develop probabilistic versions of special and general relativity (Maxwell, 1985).

As for abstractness, this is, for AOE, without methodological significance except that, as physical theory draws closer to capturing physicalism, it is almost bound to specify entities increasingly remote from those of ordinary experience. We begin with corpuscles, minute billiard balls, in the 17th century; these then transmute into point-particles that interact by means of forces; these, in turn, transmute into classical fields, into quantum fields, into curved space-time, into superstrings in ten dimensional space-time - entities increasingly remote from the familiar billiard ball.

We have seen, so far, that AOE recognizes, ultimately, just one non-empirical criterion, unity or compatibility with physicalism, whereas MAI recognizes many, and makes no attempt to show that these all devolve from just one basic criterion. But I come now to a non-empirical criterion which AOE holds to be absolutely central, but which MAI does not even recognize as an aesthetic criterion at all: explanatory power.

Explanatory power is an ambiguous concept. We may hold that T_1 has more explanatory power than T_2 if (1) T_1 has greater empirical content than T_2, or if (2) T_1 has greater unity than T_2 even though the same empirical content. Let us call these type (1) and type (2) explanatory power respectively.

We need also to recognize that criteria legitimately employed in science to assess theories can be put into three categories: (a) empirical, (b) empirical-dependent, and (c) non-empirical. By (a) I mean simply the predictive success of the theory in question; by (b) I mean properties of theories that have to do with how amenable they are to being assessed empirically, such as testability and empirical content; and by (c) I mean properties of theories that have nothing directly to do with empirical success but which are deemed to indicative of truth, or of potential empirical success.

Type (1) explanatory power is a typical type (b) property of theories. But, according to AOE, type (2) explanatory power is the key type (c) non-empirical property of theories from which, as we have seen, all others, such as symmetry, simplicity or metaphysical allegiance arise. In seeking to acquire knowledge about the world, we actively hunt for clues as to the kind of universe we are in, and hence the kind of theories we need to develop. *The* big clue that we have (apparently) discovered is that the universe is more or less comprehensible in some way or other, it being possible to discover explanations for phenomena; this is enshrined in theses of near and rough comprehensibility, at levels 6 and 7. We then make the bold conjecture, at level 5, that the universe is perfectly comprehensible in some way or other - the universe being such that there is some one kind of explanation for all phenomena, couched in terms of God, a cosmic purpose (which everything is designed to fulfil), a cosmic programme, a unified physical entity, or something else. From Galileo on, science has, in effect, made the even bolder conjecture that the universe is *physically* comprehensible, at level 4, and comprehensible in terms of the best available blueprint, at level 3. Type (2) explanatory power, to repeat, is the *key* type (c) non-empirical criterion of theory choice, from which all other type (c) criteria arise. If any property of theories cried out to be *the* aesthetic property of beauty, which scientists quite properly take note of as being methodologically significant, it is type (2) explanatory power.

And yet, astonishingly, MAI does not even include explanatory power in its list of aesthetic properties of theories, despite its open-ended, all-inclusive approach to listing such properties (in such sharp contrast to AOE).

MAI holds that the requirement of type (2) explanatory power is a permanent criterion of science, one which can be arrived at by aim-analysis, taking the aim of science to be empirical adequacy (McAllister, 1996, 11). It is clear that type (2), and not merely type (1) explanatory power is intended here, for McAllister writes that a successful explanatory theory is deemed to have "identified a pattern or mechanism underlying the data" (1996, 11). But such an analysis could only, at most, justify adopting the requirement of type (1) explanatory power; it does not justify adopting type (2) explanatory power as a requirement - not unless the truth, the universe that is, is permanently presumed to have a more or less unified dynamic structure (a presumption which contradicts SE). McAllister provides no argument in support of the contention that favouring theories with type (2) explanatory power can be justified by an appeal to the aim of empirical adequacy. He does refer to an approach to the problem of induction, espoused by Braithwaite and Mellor, according to which we are justified in proceeding as if regularities or patterns exist in nature because this gives us the best hope of acquiring knowledge whatever the universe may be like (McAllister, 1996, 100-1). It is this argument, perhaps, which McAllister assumes justifies taking type (2) explanatory power as a permanent criterion for theory choice, arrived at by aim-analysis, taking the aim of science to be empirical adequacy.

But there are three things wrong with this.

First, the Braithwaite-Mellor justification of induction does not work, as I shall show in the next section.

Second, many different kinds of explanation are possible; the universe may be comprehensible (phenomena being explainable) in many different ways, and to many different degrees, as the different theses from levels 3 to 7 of AOE attest. Here, above all, science needs to be flexible and responsive, constantly modifying the kind of explanations to be sought in the light of empirical success and failure, in the kind of way in which the hierarchical methodological structure of AOE is designed to facilitate. If ever there was a role for the aesthetic induction, it would surely be here, in connection with explanatory power. But in excluding type (2) explanatory power from the list of aesthetic properties, and in making it a fixed, unchanging requirement of theory choice, MAI fails to exploit this vital need for science constantly to modify and improve the kind of explanations that it seeks. It is just this, by contrast, that is the key idea behind AOE.

Third, if McAllister's argument were successful, so that giving preference theories that exhibit type (2) explanatory power could be justified by an appeal to the aim of empirical adequacy, then this would be a disaster for MAI, for it would obviate entirely the need for science to consider aesthetic properties of theories, and to employ the aesthetic induction. As I have argued above, all aesthetic properties of theories that have any methodological significance can be derived from the demand for unity - that is, the demand for type (2) explanatory power. Once type (2) explanatory power is acknowledged to be methodologically significant, no other aesthetic properties of theories are required by science.

I conclude this section by just mentioning three further differences between AOE and MAI.

First, reasons given in defence of MAI for holding that aesthetic considerations are methodologically significant in science arise from the fact that scientists themselves have stressed their importance, and they do indeed seem influence what theories are chosen in science. Reasons given in defence of AOE for holding that type (c) non-empirical considerations are methodologically important are much stronger: science becomes impossible if such considerations are not deployed to rule out endlessly many empirically successful but grossly *ad hoc* theories.

Second, MAI, despite being a contribution to the rationalist conception of science, does not provide a basis for systematically correcting scientific practice. But AOE does. As I have already remarked, if a view genuinely increases our understanding of science, it would be surprising if it did not have implications for scientific practice. AOE passes this test, in emphasizing the need for explicit articulation of metaphysical theses at levels 3 and 4, and explicit tackling of the problems thereby generated.

Third, MAI and AOE conceive of the relationship between science and the philosophy of science differently. MAI takes the conventional view for granted: philosophy of science is a meta-discipline which seeks to spell out and justify methods implicit in successful scientific practice, but which is quite distinct from science itself. AOE upholds the unorthodox view that the philosophy of science is an integral part of science itself, influenced by and seeking to influence science, articulating and critically assessing actual and possible aims and methods for science, at various levels, the fundamental aim being to contribute to scientific progress. A new level 3 aim for physics, i.e. a new blueprint, plus associated new methods, might constitute a major contribution to theoretical physics, as well as being a contribution to the philosophy of physics. Einstein's special theory of relativity is an example. It puts forward both a modified blueprint (Newtonian space-time becoming Minkowskian space-time), and modified methods (Galilean invariance becoming Lorentz invariance): it is thus a major contribution to physics itself which is also a contribution to the aims and methods of physics - i.e to the philosophy of physics.

5 Assessment

Which is to be preferred, AOE or MAI? The two views need not, of course, be regarded as rivals. AOE is a highest level model, whereas MAI is a medium level model; one could consider accommodating MAI within AOE. This would require, however, that MAI be modified quite extensively, as the previous section has shown.

Interpreting AOE and MAI as rival rationalist accounts of science, my chief criticism of MAI is that it is a version of SE, and thus suffers from the defects that all versions of SE suffer from. Given any empirically successful theory, T, there will always be endlessly many *ad hoc* rivals to T, even more empirically successful than T, which will never even be considered within science, let alone considered and rejected. In persistently rejecting such *ad hoc* rivals, even more empirically successful than T, science makes a persistent assumption about the nature of the universe. This contradicts SE; and contradicts MAI.[17]

McAllister might seek to evade this conclusion by arguing, as he does in his book, that non-empirical, aesthetic criteria that rule out acceptance of empirically successful, *ad hoc* rival theories, are too diverse in character, too changeable over time, to amount to the implicit acceptance of any persistent assumption. But such an argument collapses the moment one takes radically *ad hoc* theories into account of the kind considered in section 2 above, and in Maxwell (1998), 47-54. Rejection (or rather complete neglect) of such radically *ad hoc* theories persists throughout revolutions and all changes in aesthetic fashions in science. The persistent rejection of such theories unquestionably commits science to making a substantial metaphysical assumption about the nature of the universe.

McAllister might, at this point, appeal to the pragmatic justification of induction of Braithwaite and Mellor, already referred to above (McAllister, 1996, 100-1). According to this argument, science proceeds, and is justified in proceeding, as if it assumes there are regularities to be discovered, but does not actually assume that regularities exist. But even if this argument is valid, it does not in any way invalidate my point above, that in persistently rejecting empirically successful *ad hoc* theories science implicitly makes a persistent metaphysical assumption about the world. It should be noted that kinds of *ad hoc* theories can be formulated that specify regularities, in that these theories are invariant with respect to position and time (no specific places or times being specified by the theories). These theories might be said to specify *ad hoc* regularities.

But in any case the Braithwaite-Mellor attempt at solving the problem of induction does not succeed. Restricting science to the search for regularities is both too narrow, and not narrow enough. Too narrow, because it is conceivable that we can live and acquire knowledge but not by searching for regularities in phenomena. God might get in touch with us, explain His purposes, keep us informed about what is going to happen. Getting in touch with God by means of prayer and meditation, and not by searching for regularities, might be the way to acquire knowledge; and various other science fiction possibilities can be imagined (see Maxwell, 1998, 185). Such possibilities are excluded by the search for knowledge as characterized by Braithwaite and Mellor; this means these possibilities are just dogmatically assumed to be false. But the Braithwaite-Mellor approach is also not narrow enough because, as I have indicated above, if science is to be possible, *ad hoc* regularities must be persistently excluded from consideration. And, as we have seen, there is no sharp distinction between the *ad hoc* and the non-*ad hoc*. In section 2 above I listed 8 kinds of disunity - in effect, 8 different ways in which regularities might be *ad hoc*, which range from the severely *ad hoc* (distinct regularities in different space-time regions) to the scarcely *ad hoc* at all (space-time and matter not being unified). What does the policy of "inductive projection" (McAllister, 1996, 101) amount to? Does it involve merely excluding permanently all theories that are type (1) *ad hoc* (distinct regularities in different space-time regions)? Again, this is both too narrow, and not narrow enough. Exactly the same objection arises wherever the line is drawn, between regularities that are too *ad hoc* to be considered by science, and those that are sufficiently non-*ad hoc* to be open to scientific consideration. We cannot, at this point, simply invoke the aesthetic induction, and declare that we discover, by induction, where the line is to be drawn between the acceptably and unacceptably *ad hoc*, because, as McAllister himself has so clearly shown, for logical reasons, the aesthetic induction can only proceed once methods have been arrived at by aim-analysis (McAllister, 1996, 76).

Another approach might be to favour permanently in science theories that are as non-*ad hoc* as possible, in all 8 senses, but not to draw a rigid line between the acceptably and unacceptably *ad hoc*. This would allow something like the aesthetic induction to proceed in science (although not in quite the open-ended way in which McAllister envisages). But even this attempt at solving the problem is both too narrow and not narrow enough. Endlessly many universes are possible in which we may live and acquire knowledge, and yet this inductive policy would not be appropriate for acquiring knowledge. It biases the search for knowledge in the direction of physicalism. But physicalism may be false; the universe may be comprehensible in some other way, or not comprehensible at all.

My claim is that the best possible way in which we can go about seeking knowledge is to do so employing the hierarchical aims-and-methods structure of (generalized) AOE. We must make some kind of guess as to what kind of universe we are in to proceed at all. At the top of the hierarchy we need to put those relatively contentless guesses which are such that their truth is required for acquisition of knowledge to be possible at all. These are justifiably permanent items of scientific knowledge. As we descend the hierarchy, we need to put increasingly contentful guesses, chosen because these seem to be the most fruitful from the standpoint of engendering methods that seem to offer the best help with acquiring empirical knowledge. As we proceed, we revise these guesses in the light of the relative empirical success and failure of rival research programmes, based on rival low-level metaphysical guesses. We try to keep such revisions as low as possible in the hierarchy when we seem to be achieving overall success, and only allow revisions to ascend higher up in the hierarchy when success is not being achieved, and higher-level revisions seem to be required.

This hierarchical conception of scientific method enables science to respond sensitively to what it seems to discover about the nature of the universe, lower-level aims and methods being adjusted in the light of apparent empirical success and failure, and within a framework of fixed, relatively unproblematic, higher-level aims and methods. All attempts at justifying induction pragmatically that are known to me, along lines advocated by Braithwaite and Mellor, fail because they fail to take note of the resources of (generalized) AOE. They all attempt to justify methods that are demonstrably not as efficient as those of AOE in enabling us to acquire knowledge of nature. They fail to encapsulate the responsiveness, the flexibility, the open-endedness and precision, of AOE.

And this is true of MAI as well. Indeed, as we saw above, in section 4, the aesthetic induction has conservatism built into it, and cannot help engender revolutionary new ideas for revolutionary new theories, whereas AOE is designed to do just that. It embodies a rational, if fallible, method of discovery for theoretical physics.

References

Bohm, D. (1952) A Suggested Interpretation of the Quantum Theory in terms of "Hidden" Variables, *Physical Review 85*, 166-79 and 180-93.

Ghirardi, G.C., Rimini, A. and Weber, T. (1986) Unified Dynamics for Microscopic and Macroscopic Systems, *Physical Review D34*, 470-91.

Goodman, N. (1954) *Fact, Fiction and Forecast*, Athlone Press, London.
Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.
Lakatos, I. (1970) Falsification and the Methodology of Scientific Research Programmes, in I. Lakatos and A. Musgrave, eds. *Criticism and the Growth of Knowledge*, Cambridge University Press, London, 91-195.
Maxwell, N. (1972a) A Critique of Popper's Views of Scientific Method, *Philosophy of Science 39*, 131-52.
_______ (1972b) A New Look at the Quantum Mechanical Problem of Measurement, *American Journal of Physics 40*, 1431-5.
_______ (1974) The Rationality of Scientific Discovery, *Philosophy of Science 41*, 123-53 and 247-95.
_______ (1976) Towards a Micro Realistic Version of Quantum Mechanics, *Foundations of Physics 6*, 275-92 and 661-76.
_______ (1982) Instead of Particles and Fields: A Micro Realistic Quantum "Smearon" Theory, *Foundations of Physics 12*, 607-31.
_______ (1984) *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science*, Basil Blackwell, Oxford).
_______ (1985) Are Probabilism and Special Relativity Incompatible?, *Philosophy of Science 52*, 23-44.
_______ (1988) Quantum Propensiton Theory: A testable Resolution of the Wave/Particle Dilemma, *The British Journal for the Philosophy of Science 39*, 1-50.
_______ (1993a) Does Orthodox Quantum Theory Undermine, or Support, Scientific Realism?, *The Philosophical Quarterly 43*, 139-57.
_______ (1993b) Induction and Scientific Realism: Einstein versus van Fraassen, *The British Journal for the Philosophy of Science 44*, 61-79, 81-101 and 275-305.
_______ (1994) Particle Creation as the Quantum Condition for Probabilistic Events to Occur, *Physics Letters A 187*, 351-355.
_______ (1998) *The Comprehensibility of the Universe: A New Conception of Science*, Oxford University Press, Oxford.
_______ (1999a) Has Science Established that the Universe is Comprehensible?, *Cogito 13*, 139-45.
_______ (1999b) Are There Objective Values?, *The Dalhousie Review 79*, No. 3, 301-17.
_______ (2000a) A new conception of science, *Physics World 13*, No. 8, 17-8.
_______ (2000b) Can Humanity Learn to become Civilized? The Crisis of Science without Civilization, *Journal of Applied Philosophy 17*, 29-44.
_______ (2001) *The Human World in the Physical Universe: Consciousness, Free Will and Evolution*, Rowman and Littlefield, Lanham.
_______ (2002a) Is Science Neurotic? *Metaphilosophy 33*, no. 3, April, 259-299.
_______ (2002b) The Need for a Revolution in the Philosophy of Science, *Journal for General Philosophy of Science 33*, 381-408.
_______ (2004) Popper, Kuhn, Lakatos and Aim-Oriented Empiricism, *Philosophia 32* (forthcoming).
McAllister, J. (1989) Truth and Beauty in Scientific Reason, *Synthese 78*, 25-51.
_______ (1990) Dirac and the Aesthetic Evaluation of Theories, *Methodology and Science 23*, 87-102.
_______ (1991) The Simplicity of Theories: Its Degree and Form, *Journal for General*

Philosophy of Science 22, 1-14.
_______ (1996) *Beauty and Revolution in Science*, Cornell University Press, Ithaca.
Penrose, R. (1986) Gravity and State Vector Reduction, in Isham, C. and Penrose, R., eds., *Quantum Concepts in Space and Time*, Oxford University Press, Oxford, 129-46.
Smart, J.J.C. (2000) Review of "The Comprehensibility of the Universe", *British Journal for the Philosophy of Science 51*, 907-11.

Notes

1. For discussion of the claim that SE is upheld by a wide range of views about science see Maxwell (1998, 37-45). Bayesianism might seem to reject SE, in acknowledging both prior and posteriori probabilities. But Bayesianism tries to conform to the spirit of SE as much as possible, by regarding prior probabilities as personal, subjective and non-rational, their role in theory choice being reduced as rapidly as possible by empirical testing: see Maxwell (1998, 44).

2. All the possible phenomena, predicted by any dynamical physical theory, T, may be represented by an imaginary "space", S, each point in S corresponding to a particular phenomenon, a particular kind of physical system evolving in time in the way predicted by T. In order to specify ad hoc rivals to T that fit all available evidence just as well as T does, all we need do is specify a region in S that consists of phenomena that have not been observed, and then replace the phenomenon predicted by T with anything we care to think of. Given any T, there will always be infinitely many such *ad hoc* rivals to T.

3. For a more detailed discussion of empirically successful *ad hoc* rivals to accepted theories, see Maxwell (1998, 51-4).

4. This argument generalizes Goodman's (1954) argument concerning bleen and grue.

5. Induction-to-the-best-explanation gets this part right!

6. This is where "induction-to-the-best-explanation" goes wrong. It tries to make persistent preference for explanatory theories in science compatible with SE, even though such persistent preference is not based on empirical considerations.

7. For a much more detailed presentation of this refutation of SE see Maxwell (1998), chapter 2.

8. See Maxwell (2001a) for an elaboration of this point.

9. It may be asked: But how can acceptance of a level 3 assumption both influence, and be influenced by, acceptance of level 2 theories? The answer is that, at any stage in the development of science, rival level 3 ideas can contend; these lead to rival research programmes (Lakatos, 1970), which can be assessed with respect to their relative empirical growth. Within a research programme, theories are rejected that clash with the basic level 3 idea; this idea is rejected if a rival research programme meets with greater empirical success over a period of time. Level 3 ideas are also assessed in terms of how well they exemplify the accepted level 4 thesis. (But this too is open to revision, if such a revision leads to a more empirically progressive research programme.) For details see Maxwell (1998), chapters 4 and 5.

10. Corresponding to each cosmological thesis, at level 3 to 10, there is a more or less problematic *aim* for theoretical physics: to specify that cosmological thesis as a true, precise, testable, experimentally confirmed "theory of everything". Aims corresponding to levels 9 and 10 are relatively unproblematic: circumstances will never arise such that it would serve the interests of acquiring knowledge to revise these aims. As one descends the hierarchy of cosmological assumptions, the corresponding aims become increasingly problematic, increasingly likely to be unrealizable, just

because the corresponding assumption becomes increasingly likely to be false. Whereas upper level aims and methods will not need revision, lower level aims and methods, especially those corresponding to level 3, will need to be revised as science advances. Thus lower level aims and methods evolve within the fixed framework of upper aims and methods.

11. For the generalization of aim-oriented empiricism to form aim-oriented rationality see Maxwell (1984), (2000b) and (2001b, chapter 9).

12. Dynamical theories are partially ordered with respect to the extent that they exemplify physicalism, with respect to their degree of unity, in other words. For further details see Maxwell (1998), chapter 4.

13. For a very much more detailed exposition of this solution of the problem of simplicity, together with an account of the way in which great unifying theories of physics illustrate the solution, see Maxwell (1998), chapters 3 and 4.

14. For more detailed accounts of locally gaunge invariant theories see Maxwell (1998), 131-2, 135-9, and further texts referred to therein.

15. For a discussion of how this account can be extended to incorporate terminological simplicity, see Maxwell (1998), 110-3. For further details, see Maxwell (1998), chapters 3 and 4.

16. Elsewhere I have defended an objectivist, realist account of value: see Maxwell (1984) chapter 10, (1999b) and (2001b) chapter 2. This does not, however, affect the present argument.

17. McAllister might, of course, reject SE, and defend MAI in such a way that MAI acknowledges that science makes a substantial, permanent metaphysical assumption about the nature of the universe - namely that the universe is such that no *ad hoc* theory is true. But at once two major problems arise. What precisely is this assumption in view of the fact that there is no sharp distinction between the *ad hoc* and the non-*ad hoc*? What is the justification for making this assumption? In order to answer these questions satisfactorily, it is necessary to adopt AOE, which involves abandoning those parts of MAI which clash with AOE.