

# Non-Empirical Requirements Scientific Theories Must Satisfy: Simplicity, Unification, Explanation, Beauty

Nicholas Maxwell

Email: [nicholas.maxwell@ucl.ac.uk](mailto:nicholas.maxwell@ucl.ac.uk)

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

## Abstract

A scientific theory, in order to be accepted as a part of theoretical scientific knowledge, must satisfy both empirical and non-empirical requirements, the latter having to do with simplicity, unity, explanatory character, symmetry, beauty. No satisfactory, generally accepted account of such non-empirical requirements has so far been given. Here, a proposal is put forward which, it is claimed, makes a contribution towards solving the problem. This proposal concerns unity of physical theory. In order to satisfy the non-empirical requirement of unity, a physical theory must be such that the same laws govern all possible phenomena to which the theory applies. Eight increasingly demanding versions of this requirement are distinguished. Some implications for other non-empirical requirements, and for our understanding of science are indicated.

**Key Words:** Simplicity, Unification, Explanation, Symmetry, Methodology, Scientific Theory, Theoretical Physics, Metaphysics

## 1 - The Problem

A scientific theory, in order to be accepted as a part of theoretical scientific knowledge, must be sufficiently:

- (1) empirically successful;
- (2) empirically contentful;
- (3) simple, unified, explanatory, beautiful, elegant, harmonious, non-*ad hoc*, conceptually coherent, invariant, symmetrical, organic, inwardly perfect (all terms used in this context by scientists and philosophers of science).

It is important to note that this third non-empirical requirement plays a crucial role in science, especially in physics, to the extent, even, of persistently over-riding empirical requirements. Given any accepted physical theory, T, however successful empirically, it will always be possible to concoct endlessly many more empirically successful theories, T<sub>1</sub>, T<sub>2</sub>, etc., if non-empirical requirements can be ignored. T will make endlessly many predictions concerning phenomena not yet observed. Rivals to T can be concocted by modifying T in ad hoc ways so that each rival makes a different prediction for some unobserved phenomenon. Then independently testable and corroborated hypotheses can be added to these rivals, the result being a series of theories, T<sub>1</sub>, T<sub>2</sub>, etc., which have all the empirical success of T, have excess empirical content over T, this excess content being empirically corroborated. T<sub>1</sub>, T<sub>2</sub>, etc., are thus empirically more successful than T. Furthermore, almost all accepted physical theories run into empirical difficulties for some phenomena and are, on the face of it, refuted. T<sub>1</sub>, T<sub>2</sub>, etc., can be further modified in an entirely ad hoc, arbitrary fashion, so that these theories predict correctly the phenomena that ostensibly refute T, so that T<sub>1</sub>, T<sub>2</sub>, etc. are, in addition empirically successful where T

is refuted. In scientific practice, of course, these rivals to T, much more empirically successful than T, are never considered at all because of their failure to satisfy non-empirical requirements. The fact that such empirically more successful theories are persistently ignored because of their unacceptably ad hoc, complex, disunified character means that non-empirical considerations are persistently over-riding empirical considerations in physics.<sup>1</sup> Non-empirical considerations thus play an irreplaceable and fundamental role in science.

But what is this mysterious non-empirical feature of simplicity, unity, etc., that any acceptable scientific theory must possess? This is the problem I tackle in this paper.

It deserves to be noted that this is an absolutely fundamental problem in the philosophy of science. The solution is required for (a) a specification of scientific method, and (b) the solution to the problem of induction. Both points are demonstrated by the point made above, namely that non-empirical considerations persistently over-ride empirical considerations when it comes to the acceptance of scientific theories.

Non-empirical considerations can have a purely pragmatic role in science: in certain contexts, we choose one formulation over another, or even one theory over another, not because we judge our choice to be more likely to be true, but because it is such that the equations are easier to solve, it is easier to extract useful predictions from the choice we make. Here, I ignore such pragmatic considerations, at least initially, and concentrate exclusively on non-empirical requirements judged to be indicative of truth or knowledge (however fallibly).

The following nine aspects of the problem can be distinguished.

- (1) The terminological problem: How can simplicity, unity (etc.) be significant notions, having methodological significance, when the question of whether a theory is simple or complex, unified or disunified, will depend crucially on how the theory in question is formulated? A change of formulation can turn a simple theory into a complex one, and vice versa.
- (2) What does it mean to assert of a theory that it is unified, simple, explanatory, etc.?
- (3) How can degrees of simplicity, unity (etc.) be assessed?
- (4) Can notions of simplicity, unity (etc.) be explicated which do justice to the intuitions and decisions of scientists about non-empirical requirements theories must satisfy, and which even clarify and improve such intuitions and decisions?
- (5) How many *different* features of theories are involved? The plethora of terms used by scientists and philosophers of science in this context does not inspire confidence that people know what they are talking about.
- (6) How can one do justice to the fact that conceptions of simplicity or unity evolve with

evolving knowledge? Three of Newton's four rules of reasoning concern simplicity (Newton, 1962, 398-400), and yet Newton's notions are different from those of a modern physicist.

(7) How can one do justice to ambiguity of judgements concerning the relative simplicity or unity of theories? Thus Newton's theory of gravitation seems in one way much simpler than Einstein's, but in another way more complex, or at least less unified.

(8) How is persistent preference for simple or unified theories in science, even against the evidence, to be justified? This is, it should be noted, the problem of induction. Solve this, and the problem of induction is solved.

(9) What implications does the solution to these problems have for science itself?

I shall concentrate initially on problems (1) to (3), as these arise in connection with unity of theory, and then make some remarks about (4) to (9).

Richard Feynman has provided the following amusing illustration of problem (1): see (Feynman et al. 1965, 25-10 - 25-11). Consider an appallingly complex universe governed by  $10^{10}$  quite different, distinct laws. Even in such a universe, the true "theory of everything" can be expressed in the dazzlingly simple, unified form:  $A = 0$ . Suppose the  $10^{10}$  distinct laws of the universe are:

(1)  $F = ma$ ; (2)  $F = Gm_1m_2/d^2$ ; etc. Let  $A_1 = (F - ma)^2$ ,  $A_2 = (F - Gm_1m_2/d^2)^2$ , etc., for all  $10^{10}$  distinct laws. Let:

$10^{10}$   
 $A = \sum_{r=1} A_r$ . The true "theory of everything" of this universe can now be formulated as:

$A = 0$ . (This is true iff each  $A_r = 0$ .)

Most scientists and philosophers of science recognize that non-empirical considerations of simplicity, etc., play an important role in science, but no one has been able so far to solve the terminological problem (problem (1)). Weyl (1963, 155) remarked correctly that "The problem of simplicity is of central importance for the epistemology of the natural sciences". Einstein (1949, 23) recognized the problem but confessed that he was not "without more ado, and perhaps not at all" able to solve it. Jeffreys and Wrinch (1921) suggested that simplicity could be identified with paucity of adjustable constants in equations, but unfortunately number of constants can be changed by changes of formulation. Popper (1959, ch. VII) proposed that simplicity is falsifiability, but unfortunately falsifiability can always be increased by adding on independently testable hypotheses which, in general, will drastically decrease simplicity. (Popper's adjunct proposal, in terms of dimension, does not work either, and is in any case subservient to falsifiability.) More recently Friedman (1974), Kitcher (1981) and Watkins (1984) have sought to identify simplicity or unity with structural, formal or

axiomatic features, but these attempts fail: see Maxwell (1998, 65-8) for a detailed criticism of these three attempted solutions to the problem. More recently, McAllister (1996), Weber (1999), Schurz (1999), and Bartelborth (2002) have tackled the problem: I comment on these contributions below. For an excellent paper on unification of theoretical physics see Maudlin (1996). Few, however, seem to hold out much hope for a general theory of unification in science. One author has declared recently of “a general ‘theory’ of unification” that “no such account is ... possible” (Morrison, 2000, 1).

## 2 The Proposal Concerning Unity of Theory

Many previous attempts at solving the problem have failed because of mistakes concerning two crucial preliminary points.

The first mistake is to formulate the problem, in the first instance, too generally as a problem about *scientific* theories. It is vital, in the first instance, to restrict the problem to fundamental, dynamical *physical* theories. Branches of the natural sciences are not independent of one another; they are interconnected. Biology presupposes chemistry, and even physics, chemistry presupposes physics, geology and astronomy presuppose physics, and phenomenological physics presupposes fundamental physics. All branches of natural science besides theoretical physics, in other words, are constrained by results from some more fundamental science which, in the end, can be traced back to theoretical physics. This is neither a pro- nor anti-reductionist thesis; it is just the simple observation that theories in non-physical branches of natural science are in the end, in general, exceptions aside, constrained by physics. Only in fundamental theoretical physics does the question of the nature of non-empirical constraints on theories arise in something like a naked, pure form. We must, in the first instance, restrict the problem to that of fundamental, dynamical *physical* theory.

The second mistake is to suppose that simplicity, unity, etc., is a feature of the *theory* itself, its axiomatic structure, its simplicity of formulation, its number of postulates, its characteristic pattern of derivations, its number of adjustable constants. But all this involves looking at entirely the wrong thing. What one needs to look at is not the theory itself, but at the world, or rather at *what the theory says about the world*, the *content* of the theory in other words.<sup>2</sup> At a stroke the worst aspect of the problem of what unity *is* vanishes. No longer does one face the *terminological* problem of unity - the problem of the formulation-dependent nature of unity. Suppose we have a given theory T, which is formulated in N different ways, some formulations exhibiting T as beautifully unified, others as horribly complex and disunified, but all formulations being interpreted in precisely the same way, so as to make precisely the same assertion about the world. If unity has to do exclusively with *content*, then *all these diverse formulations of T, having the same content, have precisely the same degree of unity*. The variability of apparent unity with varying formulations of one and the same theory, T, (given some specific interpretation), which poses such an insurmountable problem for traditional approaches to the problem, poses no problem whatsoever for the thesis that unity has to do with *content*. Variability of formulation of a theory which leaves its content unaffected is wholly irrelevant: the unity of the theory is unaffected.

So much for the first of our nine problems, in so far as it concerns unity. We now face problems (2) and (3): What exactly does it mean to assert that a dynamical physical theory has a unified content? How are degrees of unity of the *content* of a theory to be assessed?

What unity of content means is that the theory has *the same* content throughout the range of possible phenomena to which the theory applies. Unity, in other words, means that there is just *one* content throughout the range of possible phenomena to which the theory applies. If the theory postulates *different* contents, *different* laws, for different ranges of possible phenomena, then the theory is *disunified*, and the more such *different* contents there are so the more *disunified* the theory is. Thus “unity” means “one”, and “disunity” means “more than one”, the disunity becoming worse and worse as the number of different contents goes up, from two to three to four, and so on. Not only does this enable us to distinguish between “unified” and “disunified” theories; it enables us to assign “degrees of unity” to theories, or to partially order theories with respect to their degree of unity.<sup>3</sup> This is the nub of my proposed solution to problems (2) and (3).

To give an elementary example, Newton’s theory of gravitation,  $F = GM_1M_2/d^2$  is unified in that what the theory asserts is *the same* throughout all possible phenomena to which it applies (all bodies of all possible masses, constitution, shape, relative velocity, distance apart, at all times and places). An aberrant version of this theory, which asserts that  $F = GM_1M_2/d^2$  for times  $t \leq t_0$ , where  $t_0$  is some definite time, and  $F = GM_1M_2/d^3$  for times  $t > t_0$ , is disunified because what the theory asserts is *not* the same throughout the range of possible phenomena to which the theory applies.

Note that special terminology could be introduced to make Newtonian theory look *disunified*, and the aberrant version of Newtonian theory look *unified*. All we need do is interpret “ $d^N$ ” to mean “ $d^N$  if  $t \leq t_0$  and  $d^{N+1}$  if  $t > t_0$ ”. In terms of this (admittedly somewhat bizarre) terminology, the aberrant theory has the form “ $F = GM_1M_2/d^2$ ”, and Newtonian theory has the “aberrant” form “ $F = GM_1M_2/d^2$  for times  $t \leq t_0$  and  $F = GM_1M_2/d$  for times  $t > t_0$ ”. But this mere *terminological* reversal of aberrance or disunity does not affect the *content* of the two theories: the content of Newtonian theory remains unified, and the content of the aberrant version (which looks unified) remains disunified.

This almost suffices to solve the problem. A little more needs to be added, however, because in practice in physics assessments of degrees of unity are somewhat more complex than I have indicated so far because of the following consideration. In assessing the extent to which a theory is disunified we may need to consider *how* different, or *in what way* different, one from another, the different contents of a theory are. A theory that postulates different laws at different times and places is disunified in a much more serious way than a theory which postulates the same laws at all times and places, but also postulates that distinct kinds of physical particle exist, with different dynamical properties, such as charge or mass. This second theory still postulates *different* laws for different ranges of phenomena: laws of one kind for possible physical systems consisting of one kind of particle, and slightly different laws for possible physical systems

consisting of another kind of particle. But this second kind of difference in content is much less serious than the first kind (which involves different laws at different times and places).

What this means is that there are different *kinds* of disunity, different *dimensions* of disunity, as one might say, some more serious than others, but all facets of the same basic idea. We can, I suggest, distinguish at least eight different facets of disunity, as follows.

Any dynamical physical theory, T, can be regarded as specifying an abstract space, S, of possible physical states to which the theory applies, a distinct physical state corresponding to each distinct point in S. (S might be a set of such spaces.) For unity, we require that T asserts that the same dynamical laws apply throughout S, governing the evolution of the physical state immediately before and after the instant in question. If T postulates N distinct dynamical laws in N distinct regions of S, then T has disunity of degree N. For unity in each case we require that  $N = 1$ .

(1) T divides spacetime up into N distinct regions,  $R_1 \dots R_N$ , and asserts that the laws governing the evolution of phenomena are the same for all spacetime regions within each R-region, but are different within different R-regions. Example: the aberrant version of Newtonian theory (NT) indicated above.

(2) T postulates that, for distinct ranges of physical variables (other than position and time), such as mass or relative velocity, in distinct regions,  $R_1, \dots, R_N$  of the space of all possible phenomena, distinct dynamical laws obtain. Example: T asserts that everything occurs as NT asserts, except for the case of any two solid gold spheres, each having a mass of between one and two 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. Here,  $N = 2$  in a type (2) way.

(3) In addition to postulating non-unique physical entities (such as particles), or entities unique but not spatially restricted (such as fields), T postulates, in an arbitrary fashion,  $N - 1$  distinct, unique, spatially localized objects, each with its own distinct, unique dynamic properties. Example: T asserts that everything occurs as NT asserts, except there is one object in the universe, of mass 8 tons, such that, for any matter up to 8 miles from the centre of mass of this object, gravitation is a repulsive rather than attractive force. The object only interacts by means of gravitation. Here,  $N = 2$ , in a type (3) way.

(4) T postulates physical entities interacting by means of N distinct forces, different forces affecting different entities, and being specified by different force laws. (In this case one would require one force to be universal so that the universe does not fall into distinct parts that do not interact with one another.) Example: T postulates particles that interact by means of Newtonian gravitation; some of these also interact by means of an electrostatic force  $F = Kq_1q_2/d^2$ , this force being attractive if  $q_1$  and  $q_2$  are oppositely charged, otherwise being repulsive, the force being much stronger than gravitation. Here,  $N = 2$  in a type (4) way.

(5) T postulates N different kinds of physical entity,<sup>4</sup> differing with respect to some dynamic property, such as value of mass or charge, but otherwise interacting by means of the same force. Example: T postulates particles that interact by means of Newtonian gravitation, there being three kinds of particles, of mass  $m$ ,  $2m$  and  $3m$ . Here,  $N = 3$  in a type (5) way.

(6) Consider a theory, T, that postulates N distinct kinds of entity (e.g. particles or fields), but these N entities can be regarded as arising because T exhibits some symmetry (in the way that the electric and magnetic fields of classical electromagnetism can be regarded as arising because of the symmetry of Lorentz invariance, or the eight gluons of chromodynamics can be regarded as arising as a result of the local gauge symmetry of  $SU(3)$ ). 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.<sup>5</sup>

(7) If (apparent) disunity of there being N distinct kinds of particle or distinct fields has emerged as a result of cosmic spontaneous symmetry-breaking events, there being manifest unity 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. Example: Weinberg's and Salam's electroweak theory, according to which at very high energies, such as those that existed soon after the big bang, the electroweak force has the form of two forces, one with three associated massless particles, two charged,  $W^-$  and  $W^+$ , and one neutral,  $W^0$ , and the other with one neutral massless particle,  $V^0$ . According to the theory, the two neutral particles,  $W^0$  and  $V^0$ , are intermingled in two different ways, to form two new, neutral particles, the photon,  $\gamma$ , and another neutral massless particle,  $Z^0$ . As energy decreases, the  $W^+$ ,  $W^-$  and  $Z^0$  particles acquire mass, due to the mechanism known as spontaneous symmetry-breaking (involving the hypothetical Higgs particle), while the photon,  $\gamma$ , retains its zero mass. This theory unifies the weak and electromagnetic forces as a result of exhibiting the symmetry of local gauge invariance; this unification is only partial, however, because the symmetry group is a direct product of two groups,  $U(1)$  associated with  $V^0$ , and  $SU(2)$  associated with  $W^-$ ,  $W^+$  and  $W^0$ .<sup>6</sup>

(8) According to GR, Newton's force of gravitation is merely an aspect of the curvature of spacetime. As a result of a change in our ideas about the nature of spacetime, so that its geometric properties become dynamic, a physical force disappears, or becomes unified with spacetime. This suggests the following requirement for unity: spacetime on the one hand, and physical particles-and-forces on the other, must be unified into a single self-interacting entity, U. If T postulates spacetime and physical "particles-and-forces" as two fundamentally distinct kinds of entity, then T is not unified in this respect. Example: one might imagine that the quantization of spacetime leads to the appearance of particles and forces as only apparently distinct from empty space-time. Here,  $N = 1$  in a type (8) way: there is just the one self-interacting entity, empty spacetime.

For unity, in each case, as I have said, we require  $N = 1$ . As we go from (1) to (5), the

requirements for unity are intended to be accumulative: each presupposes that  $N = 1$  for previous requirements. As far as (6) and (7) are concerned, if there are  $N$  distinct kinds of entity which are not unified by a symmetry, whether broken or not, then the degree of disunity is the same as that for (4) and (5), depending on whether there are  $N$  distinct forces, or one force but  $N$  distinct kinds of entity between which the force acts.

(8) does not introduce a new kind of unity, but introduces, rather, a new, more severe way of counting different kinds of entity. (1) to (7) require, for unity, that there is one kind of self-interacting physical entity evolving in a distinct spacetime, the way this entity evolves being specified, of course, by a consistent physical theory. According to (1) to (7), even though there are, in a sense, two kinds of entity, matter (or particles-and-forces) on the one hand, and spacetime on the other, nevertheless  $N = 1$ . According to (8), this would yield  $N = 2$ . For  $N = 1$ , (8) requires that matter and spacetime are no more than aspects of one basic entity (unified by means of a spontaneously broken symmetry, perhaps).

As we go from (1) to (8), then, requirements for unity become increasingly demanding, with (6) and (7) being at least as demanding as (4) and (5), as explained above.

(1) to (8) may seem very different requirements for unity. In fact they all exemplify the same basic idea: disunity arises when *different* dynamical laws govern the evolution of physical states in different regions of the space,  $S$ , of all possible physical states. For example, if a theory postulates more than one force, or kind of particle, not unified by symmetry, then in different regions of  $S$  different force laws will operate. If (8) is not satisfied, there is a region of  $S$  where only empty space exists, the laws being merely those which specify the nature of empty space or spacetime. The eight distinct facets of unity, (1) to (8) arise, as I have said, because of the eight *different* ways in which content can vary from one region of  $S$  to another. Some of these requirements for unity are suggested by developments in 20<sup>th</sup> century physics. This is true in particular of (6) to (8). The important point, however, is that all these requirements for unity, (1) to (8), exemplify the same basic idea: a theory, in order to be unified, must assert that the same laws apply throughout the phenomena to which it applies. (1) to (8) in effect represent different, increasingly subtle ways in which a theory can fail to be unified in this sense, granted that  $N > 1$  in each case.<sup>7</sup>

### 3 - Objections

It may be objected that we never encounter the naked content of a theory, formulation free; we only encounter theories given some formulation. How, then, can we judge whether the *content* does or does not vary through the space  $S$ ? The answer is that theories are not natural objects we stumble across; *we* formulate theories, and it is for us to ensure, granted we want our theories to be unified, that the content does not change as we move through  $S$ . We can arrange, however, that formulation matches content by ensuring that the terminology, the concepts, we use to formulate a theory do not surreptitiously change as we move through  $S$ . Given invariant concepts, if the form of the theory is also invariant throughout  $S$ , its content will be too. But if, for example, we

surreptitiously change our units of length as we move through space, then a theory whose content is spatially invariant will change its form with changes of spatial position (a point which will be taken up again below).

It may be objected that, given any theory, however unified, special regularities will always arise in restricted regions of S, which means disunity. Whether or not the theory is unified is, at best, ambiguous. Thus, given NT, in some regions of S there will be solar systems with planets that rotate in the same direction and conform to Bode's law, whereas in other regions of S these regularities or "laws" will be violated. The answer is to distinguish sharply between accidental and law-like regularities; only the latter are relevant for the assessment of unity. But how is this distinction to be made? This problem has been solved elsewhere: see Maxwell (1968; 1998, 141-55). Given a true law-like statement, this is a genuine physical law (having nomic necessity) if and only if physical dynamical (or necessitating) properties exist corresponding to the law. Thus Newton's law of gravitation can be interpreted as attributing the dynamical, necessitating property of Newtonian gravitational charge to massive objects. Objects that have this property of necessity obey Newton's law of gravitation. (The empirical content of NT, on this interpretation, is concentrated in the factual assertion: all massive objects possess Newtonian gravitational charge equal to their mass.) If no such property corresponds to a true regularity, then it is merely a true accidental regularity, and not a true law. (What is a necessitating property? This is explicated in Maxwell, 1968; 1998, 141-55.) For unity we require that dynamical necessitating properties remain the same throughout S; the regularities of (some) solar systems, mentioned above, are not relevant because these regularities are not law-like, and no dynamical property exists corresponding to them.<sup>8</sup>

It may be objected that physical systems which possess symmetries that are also symmetries of the theory which determines their evolution, will evolve in accordance with a simplified version of the theory. Thus systems consisting of two spheres equal in every way rotating in a fixed circle about their centre of mass obey a simplified version of the dynamical laws of NT. This means there are regions of S where the dynamical laws are especially simple, and thus different from other regions. Does this mean the theory is correspondingly disunified? The answer is No. We need, again, to consider dynamical properties corresponding to dynamical laws. In the example just considered, if NT is true (interpreted essentialistically) then the spheres in question possess gravitational charge just like all other massive objects. It is just that, in the case of the systems possessing some rotational symmetry, the full, rich implications of the dynamical property of gravitational charge is not made manifest.

It may be objected that we may not know whether two formulations of a theory are just that, two formulations with the same physical content, or two distinct theories with distinct contents. Heisenberg's and Schroedinger's distinct formulations of quantum theory might be an example. This is correct but beside the point. The terminological problem arises when we reformulate a given theory, T, in a variety of ways, some simple and unified, some horribly complex and disunified, but we do this in such a way as to ensure quite specifically that the different formulations have precisely the same content, make precisely the same assertions about the world, this being something that we can

always do. The solution to the problem proposed above is not in any way undermined by the fact that it sometimes happens that we do not know whether two formulations of a theory have the same or different contents. Nor is the distinction between form and content undermined: form has to do with what we write down on paper, content with what is being asserted. That we sometimes do not know whether difference of formulation ensures difference of content does not in the least undermine the distinction between form and content. It deserves to be noted, in addition, that one and the same formulation of a theory may be interpreted in more than one way, and thus may have different contents associated with it - a point which, again, does not undermine the theory presented here.

It may be objected the distinction between dynamical laws which do, and do not, remain the same throughout the space  $S$  cannot be maintained. Consider the two functions (1)  $y = 3x$  for all  $x$ , and (2)  $y = 3x$  for  $x \leq 2$  and  $y = 4x$  for  $x > 2$ . It is tempting to say that (1) remains the same as  $x$  changes, but (2) does not, since what (2) asserts changes at  $x = 2$ . But given the mathematical notion of function as a rule, (2) is just as good a function as (1) and, like (1), “remains the same” as  $x$  varies. Functions corresponding to physical theories are somewhat more elaborate than this, but the above point is not affected by that consideration: it seems that the very distinction between “remains the same” and “changes” as one moves through  $S$  collapses. Clearly, in order to meet this objection, functions corresponding to physical theories need to be restricted to a narrower notion of function than the above standard mathematical one, if we are to be able to distinguish between functional relationships which do, and which do not, “remain the same” as values of variables change. We need to appeal to what may be called “invariant functions”, functions which specify some fixed set of mathematical operations to be performed on “ $x$ ” (or its equivalent) to obtain “ $y$ ” (or its equivalent). In the example just given, (1) is invariant, but (2) is not. (2) is made up of two truncated invariant functions, stuck together at  $x = 2$ . Functions that appear in theoretical physics are *analytic*; that is, they are repeatedly differentiable. Such functions have the remarkable property that from any small bit of the function, the whole function can be reconstructed uniquely, by a process called “analytic continuation”. All analytic functions are thus invariant. The latter notion is however a wider one, and theoretical physics might, one day, need to employ this wider notion explicitly, if space and time turn out to be discontinuous, and analytic functions have to be abandoned at a fundamental level.

A similar remark needs to be made about Goodman’s (1983) paradox concerning “grue” and “bleen”. Modifying the paradox slightly, an object is grue if it is green up to time  $t$ , blue after  $t$ ; it is bleen if it is blue up to time  $t$ , green after  $t$ . Sometimes it is held that there is perfect symmetry between blue and green, on the one hand, and grue and bleen on the other, especially as “emeralds are green” is equivalent to “emeralds are grue up to  $t$ , and bleen afterwards”. But this symmetry is merely terminological and, as we have already seen in connection with the aberrant version of Newton’s theory of gravitation, discussed in section 2, terminological symmetry does not mean there is symmetry of content. That there is not symmetry of content in the grue/bleen case can be demonstrated as follows. If emeralds are grue, a person convinced of this can determine

whether  $t$  is future or past merely by looking at emeralds. But if emeralds are green, a person convinced of this cannot say whether  $t$  is in the future or past by just looking at emeralds. The content, the meaning, of *grue* and *bleen* contains an implicit reference to  $t$  in a way in which that of green and blue do not. Doubtless symmetry can be created by considering two possible worlds, ours and a Goodmanesque one with special physics and/or physiology of vision so that *grue* emeralds do not appear to change at  $t$ , whereas green ones do. This, however, is to consider conditions quite different from those specified by Goodman. The crucial point to make, in any case, is that dynamical or physical properties, of the kind attributed to physical entities by physical theories (interpreted in a conjecturally essentialistic way), are like blue and green, and unlike *grue* and *bleen*, in not containing any implicit reference to specific times or places (or hypersurfaces of  $S$  that distinguish one region of  $S$  from another). Physical properties must be *invariant* in a sense that corresponds to the *invariance* of allowed functions in physics. The more general notion of property, which includes Goodmanesque properties, is excluded, just as the more general notion of function, which includes (2) above as an “unchanging” function, is excluded.

#### 4 – Further Issues

What of the other aspects of the problem of non-empirical requirements in science mentioned in the introduction? I have space, here, only for staccato remarks concerning some of these further issues.

Some of the other terms used to refer to non-empirical requirements can be straightforwardly related to unity. We have seen that this is true of *symmetry* and *invariance*. A dynamical physical theory can be held to be *explanatory* in character to the extent that it is (1) empirically contentful and (2) unified. Beauty, elegance, harmony, conceptual coherence, non-ad hocness, inward perfection, organicity, can all be interpreted as at least presupposing unity. Do these notions add to unity, so that unity is necessary but not sufficient for beauty, elegance, etc? It must be remembered that there are eight increasingly demanding *facets* of unity, (1) to (8) above. Whatever exactly “beauty”, “elegance”, etc. may be taken to mean in the context of theoretical physics, insofar as these terms mean something in addition to unity when this is interpreted to mean no more than (1) above, it is increasingly likely the additional meaning of these terms will be captured by unity as we move from (1) to (8).

Simplicity is, however, somewhat different. The simplicity of a theory can be interpreted as having to do, not with whether the *same* laws apply throughout the space  $S$ , but rather with the *nature* of the laws, granted that they are the same. Some laws are simpler than others. The problem, here, is not to say what “simplicity” means (it should be understood in its ordinary sense) but to solve the problem that the simplicity of a theory would seem to be highly dependent on its formulation (Maxwell, 1998, 157-8). In order to solve this problem it is essential, as in the case of unity, to interpret “simplicity” as applying to the *content* of theories, and not to their *formulation*, their *axiomatic structure*, etc. Theories can only, at best, be partially ordered with respect to degrees of simplicity. Even when two theories are amenable to being assessed with respect to relative simplicity, there is always the problem that a change of variables may reverse the

assessment. Let the two theories be (1)  $y = x$  and (2)  $y = x^2$ . We judge (1) to be simpler than (2). Let  $x^2 = z$ . We now have (1)  $y = \sqrt{z}$ , and (2)  $y = z$ . Now (2) is simpler than (1). Assessment of relative simplicity of two theories may only be unambiguous when restrictions are placed on the form that physical variables can take, so that only linear transformations of the type  $z = Ax + B$  (where  $A$  and  $B$  are constants) are permitted, for example. It is a further great success of the theory presented here that it succeeds in distinguishing sharply between these two aspects of physical theory, the *unity* and *simplicity* aspects, and succeeds in explicating both.<sup>9</sup>

We can use these two notions to solve the problem of ambiguity of judgment concerning the relative non-empirical merits of Newton's and Einstein's theories of gravitation. Einstein's theory is more unified in that it eliminates gravitation as a force distinct from space and time, and reformulates Newton's first law so that it becomes the assertion that bodies move along geodesics in curved space-time, curvature being caused by mass, or by stress-energy-density more generally. On the other hand Newton's theory is simpler, in that Einstein's field equations are really six independent equations which, taken together, are more complex than the single equation which determines the Newtonian gravitational field (Schutz, 1989, pp. 195-200). For the purposes of this comparison one must attend to the *content* of these equations, what they assert about physical reality, and not merely to the *form*, which could vary with changes of formulation. Incidentally, it is reasonable to expect that, as theoretical physics draws closer to capturing the true theory of everything, the totality of fundamental physical theory will become increasingly unified, and at the same time increasingly complex.

The above eight facets of unity, and this additional notion of simplicity, may be thought to capture, between them, all that is methodologically significant in the additional terms of our list that are employed in this context: beauty, elegance, etc. However, towards the end of the paper, I have one further comment to make about this issue.

So far I have stressed that terminological unity and simplicity are irrelevant when it comes to assessing unity and simplicity in a physically significant sense. In scientific practice, however, terminology is chosen so as to reflect physically significant unity and simplicity (Maxwell, 1998, 110-3). Thus if the content of a theory exhibits certain symmetries, terminology is chosen so that it too exhibits these symmetries, so that if the theory is invariant with respect to position or orientation in space, terminology is chosen which reflects this fact. Once a theory is formulated in such "physically appropriate" terminology (as it may be called), two versions of symmetry operations arise as a result, "active" (which make changes to physical systems) and "passive" (which make corresponding changes to the description of unchanged physical systems). Granted that we formulate physical theories exclusively in such "physically appropriate" terminology, then terminological unity and simplicity comes to reflect physical unity and simplicity, and is thus, to that extent, physically significant.

What of the simplicity and unity of theories in sciences other than fundamental physics? Much needs to be said on this topic; the following brief remarks can serve only as pointers to a more adequate treatment. Solutions to the equations of fundamental

physical theory, specifying precisely how increasingly complex physical systems evolve in space and time, rapidly become horrendously complex in character. In carrying out derivations, physicists invariably “simplify” results obtained by discarding variable quantities or higher order terms judged to be insignificant in the physical situations under considerations. Just this is done when NT is “derived” from Einstein’s theory, or Kepler’s and Galileo’s laws of motion are “derived” from NT.<sup>10</sup> The outcome is a range of more or less terminologically simple phenomenological laws of only approximate validity. But the simplicity is not, here, merely pragmatic, since such a law has been “approximately derived” from some fundamental physical theory formulated in a “physically appropriate” way, the “approximate derivation” showing what the range of applicability of the law is with what degree of accuracy. Even though such laws are incompatible with the fundamental physical theory from which they have been “approximately derived”, nevertheless what the “derivations” reveal is that pragmatic simplicity has been obtained by sacrificing strict derivability and precise empirical accuracy, there being nothing here to counter the underlying unity in nature postulated by fundamental physical theory (insofar as it does postulate this). Laws such as these are prevalent throughout phenomenological physics, astrophysics and parts of physical chemistry. Even where such “approximate derivations” cannot be carried through, for large parts of chemistry, and for biology, nevertheless, as I have already remarked, laws and theories of these sciences are constrained by fundamental physics, and must endeavor to be compatible with fundamental physics, at least in the qualified way just indicated in connection with phenomenological physics. Thus, much of the great explanatory power of Darwinian theory stems from the fact that it postulates mechanisms for evolution - random inheritable variation and natural selection - which are capable of designing living things able to pursue the goals of survival and reproductive success in their given environments, these mechanisms nevertheless being compatible with the purposeless cosmos depicted by physics. Biology must accord with physics in much more specific ways as well, in that the mechanisms of inheritance and development must accord with physics, and so too the multitude of processes that take place in living things.

What implications does the account of non-empirical requirements for theories, given here, have for science? How can justice be done to evolving non-empirical requirements? How is persistent preference for unified theories, even against the evidence, to be justified? I take these three problems together.

At the beginning of this paper I demonstrated that, in physics, theories that are unified, in senses (1) and (2) at least, are persistently chosen in preference to available, empirically more successful, but disunified theories. To proceed in this way is to make the permanent assumption that the phenomena under consideration are such that all theories of these phenomena that are disunified in senses (1) and (2) are false. If physicists persistently accepted theories that postulate atoms in preference to available, empirically more successful *field* theories, it would be clear that physicists are thereby assuming that all field theories are false. Just the same holds for the persistent rejection of empirically more successful disunified theories.

But rigour demands that assumptions that are substantial, influential, problematic and

implicit need to be made explicit, so that they can be critically assessed, so that alternatives can be developed and considered, the hope being that in this way such assumptions can be improved. Thus rigour demands that science makes explicit, and so criticizable and improvable, the substantial, problematic, influential and implicit assumption that the universe, or the phenomena, are such that all disunified theories are false. This assumption, M, can easily be shown to be metaphysical, as follows. Persistent acceptance of theories unified in ways (1) and (2) involves rejecting infinitely many empirically more successful disunified rivals,  $T_1, T_2, \dots T_\infty$ , because they clash with M. In effect,  $M = \text{not}T_1$  and  $\text{not}T_2$  and ...and  $\text{not}T_\infty$ . In order to verify M we would need to falsify all of  $T_1, T_2, \dots$ and  $T_\infty$ , but as there are infinitely many theories, this cannot be done. In order to falsify M we need to verify just one of  $T_1, T_2, \dots$ or  $T_\infty$ , but physical theories cannot be verified. Hence M, being neither verifiable nor falsifiable, is metaphysical. It is a permanent metaphysical assumption of science - permanent, at least, as long as all theories disunified in senses (1) and (2) are rejected whatever their empirical success might be.

At once the question arises: How is this assumption M to be critically assessed and, perhaps, improved? Elsewhere it has been argued that once the metaphysical assumption implicit in persistent preference in science for unified theories is acknowledged, it becomes apparent that we need to adopt a new conception of science, which construes science as making a hierarchy of such assumptions, these assumptions asserting less and less as one goes up the hierarchy, and thus becoming more and more likely to be true (see Maxwell, 1998). These are assumptions about the knowability and comprehensibility of the universe. As we descend the hierarchy, assumptions become more substantial and specific, and much more likely to be false, and in need of revision. Revision is, however, kept as low down in the hierarchy as possible. Those physical theories are accepted which best accord with the evidence and the best available metaphysical assumption, B say, lowest down in the hierarchy. But B may itself be revised if a rival assumption,  $B^*$ , is developed which (a) is compatible with the assumption above it in the hierarchy, and (b) supports an empirical research programme that is more successful than the one supported by B. Examples of such metaphysical assumptions, changing over time, taken from the history of physics, include theses which assert that nature is composed of: infinitely rigid corpuscles which interact only by contact; point-atoms with mass that interact by means of rigid, spherically symmetrical, centrally-directed forces which vary with distance; a self-interacting unified classical field of force; quantum fields; some physical entity which evolves in accordance with a unified Lagrangian (or Lagrangian density) which is not the sum of two or more distinct Lagrangians with distinct physical interpretations or symmetries; a superstring quantum field; branes of M-theory. Relatively unproblematic assumptions high up in the hierarchy thus form a fixed framework within which much more specific, problematic assumptions, low down in the hierarchy, can be revised in the light of empirical success and failure. As knowledge improves, assumptions and associated methods improve as well; there is something like positive feedback between improving knowledge and improving knowledge-about-how-to-improve-knowledge, the methodological key to the success of modern science. Non-empirical requirements for theory acceptance, corresponding to metaphysical assumptions, improve with improving knowledge. Newton's requirements of simplicity

evolve into the symmetry principles of modern physics. For a suggestion as to how acceptance of the hierarchy of metaphysical assumptions is to be justified, see Maxwell (1998, ch. 5).

Insofar as physics, at some stage in its development, accepts one or other of the metaphysical theses just indicated (from the corpuscular hypothesis via the classical field to M theory), physics thereby accepts quasi-non-empirical requirements which are in addition to, and more restrictive than, the general requirements of unity and simplicity, as explicated above. (I say “quasi-non-empirical” because these metaphysical theses change in part in the light of the empirical success and failure of the research programmes to which they give rise.) There are, then, unquestionably, much more demanding, if changing, requirements governing acceptance of physical theories than mere unity and simplicity. It may well be that “beauty”, “elegance”, “organicity”, etc., acquire meanings, at least for some physicists, which are such that these notions are tied to one or other of these metaphysical theses. If so, the hierarchical view indicated above can do justice to these additional “metaphysical” meanings.

## **5 - Comparison with Other Views**

I conclude by comparing the view defended here with views defended by McAllister (1996), Weber (1999), Schurz (1999), and Bartelborth (2002).

According to McAllister, aesthetic considerations that influence choice of theory in science fall into five classes of properties of theories: symmetry, invocation of a model, visualizability/abstractness, metaphysical allegiance, and simplicity (related to unity). McAllister 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.

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.

The account of unity of physical theory I have given above is not to be found in McAllister's work. Instead of stressing, as I have done, that unity (as explicated above) is a persistent requirement that physical theories must satisfy to be acceptable, at least since Newton,<sup>11</sup> McAllister stresses rather that “There are many ways in which classes of phenomena can be said to admit unification. Because of this, the prescription that scientists should choose the theory with the greatest unifying power is indeterminate” (McAllister, 1996, 110). This last remark is in sharp contrast to the view defended here, which renders the requirement of unity highly determinate (in that there will always be

infinitely many disunified rivals to any unified theory). That requirements of simplicity and unity evolve with evolving scientific knowledge and metaphysical theses is common ground between McAllister's view and the view defended here (and spelled out at greater length in Maxwell, 1974; 1984, ch. 9; 1994; and especially 1998). The view defended here also, however, upholds a conception of unity that persists in theoretical physics, at least since Newton. No such conception is to be found in McAllister's work. For a detailed critical comparison of McAllister's and Maxwell's views, see Maxwell (forthcoming).

Weber gives what he admits is only a partial explication of unification of events. Interpreted as an explication of the unity of law or theory, however, Weber's account fails. According to Weber, two events  $E_1$  and  $E_2$  are unified with one another if they are explained by the same law. A law, for Weber a statement of the form  $\forall x(Px \rightarrow Qx)$ , must not be (i) tautological or analytic, (ii) true vacuously, (iii) an accidental generalization, or (iv) such that the statement contains "irrelevant antecedent conditions", as in "Men who take birth control pills do not get pregnant" (Weber, 1999, 482). But these four requirements do not suffice to exclude laws that are disunified, in ways (1) to (3) for example. Two events,  $E_1$  and  $E_2$ , might satisfy Weber's requirements for unity, and yet intuitively be very different kinds of event, because  $E_1$  is explained by one part of a disunified law,  $E_2$  by a very different other part of the law. The predicates  $P$  and  $Q$  may, for example, be "grue-like", in that they refer to different properties at different times or places. Weber's account fails, whether interpreted as an account of unity of events, or of laws or theories.

Schurz's account of unification is very complex in that it appeals to cognitive agents with cognitive states made up of descriptive knowledge ( $K$ ), and inferential knowledge ( $I$ ),  $K$  being made up of "relevant knowledge elements" (Schurz, 1999, 103). A "unification classification" of  $K$  is a partition of  $K$  into four disjoint subsets of phenomena, actually assimilated ( $K_a$ ), potentially assimilated ( $K_p$ ), basic ( $K_b$ ), and dissimilated ( $K_d$ ). The fewer the phenomena in  $K_d$  or  $K_b$ , and the more phenomena in  $K_a$  or  $K_p$ , the greater the unification of  $K$  will be. Along these lines, Schurz claims, the unification of a cognitive agent's cognitive states, before and after acquiring an answer to a why question, can be compared.

There is not space here to expound properly, let alone criticize, Schurz's very complex account of unification. Suffice it to say that it suffers from the disadvantage of being very much more complex than the account of unification I have given above. Furthermore, I found nothing in Schurz's account which would declare theories disunified in ways (1) to (8) to be disunified, especially when such theories are formulated so as to appear unified – apart, that is, from appealing to our intuitions that such disunified theories require further explanation. Schurz does not draw the crucial distinction between form and content, and  $K$  is made up of "linguistic representations" of phenomena (Schurz, 1999, 104).

Bartelborth's account of unification differs from the one I have given here in that it is not restricted, in the first instance, to theoretical physics, but is intended to apply, not just

to natural science but to social science as well, and to contexts where phenomena cannot be derived from laws. Bartelborth bases his account on the structuralist view of theories. On his view, explanation is to be understood in terms of the notion of “embedding”. A concrete phenomenon being explained by a theory can be construed to be a case of a partial model, representing the phenomenon in question, being embedded “into a larger theoretical model of the theory that represents the theoretical patterns or mechanisms described by the theory” (Bartelborth, 2002, 96-7). There are three key requirements for a theory to be explanatory. The explanatory power of a theory is the greater as (i) the range of phenomena to which it applies is greater, (ii) the richness of the information the theory provides about the phenomena to which it applies is greater, and (iii) the fewer the number of decompositions of the theory there are.

Requirement (iii) is the crucial one as far as the problem of unification is concerned. It is not clear (to me at least) that Bartelborth has done enough to show that some theories – the unified ones - resist decomposition. He says: “We want to exclude trivial unifications by conjunctions. Unifications should be accomplished by a coherent and organic theory that is not decomposable into sub-theories. An explanatory theory should use only a few patterns to embed many phenomena and should show a surplus content compared with conjunctions of sub-theories” (Bartelborth, 2002, 101). On the face of it, however, all theories can be decomposed into sub-theories without loss of content. Bartelborth does not show how the formal machinery he deploys excludes such decompositions in the case of those theories that are “unified”. The intuition that an “explanatory theory should use only a few patterns to embed many phenomena” is excellent, but even if the patterns postulated by a theory are interpreted in terms of the *content* of the theory, and not its *form*, it is not clear that patterns can be clearly distinguished and unambiguously counted in the very broad context that is Bartelborth’s concern, ranging as it does from the natural to the social sciences. Bartelborth’s account of unification differs, in any case, from the one I have presented here, and does not issue in the eight facets of theoretical unity distinguished above.

## References

- Aitchison, I. and A. Hey (1982), *Gauge Theories in Particle Physics*. Bristol: Adam Hilger.
- Bartelborth, T. (2002) “Explanatory Unification”, *Synthese* 130, 91-107.
- Einstein, A. (1949), "Autobiographical Notes" in P. A. Schilpp (ed.) *Albert Einstein: Philosopher-Scientist*. La Salle: Open Court.
- Feynman, R., R. Leighton and M. Sands, *The Feynman Lectures on Physics vol. II*. Reading, Mass.: Addison-Wesley.
- Friedman, Michael (1974), "Explanation and Scientific Understanding", *Journal of Philosophy* 71: 5-19.
- Goodman, N. (1983), *Fact, Fiction and Forecast*. Cambridge, Mass.: Harvard University Press.
- Griffiths, D. (1987), *Introduction to Elementary Particles*. New York: Wiley.
- Isham, C. (1989), *Lectures on Groups and Vector-Spaces for Physicists*. London: World Scientific.
- Jeffreys, H. and D. Wrinch (1921), "On Certain Fundamental Principles of Scientific

- Enquiry", *Philosophical Magazine* 42: 269-98.
- Jones, H. (1990), *Groups, Representations and Physics*. Bristol: Adam Hilger.
- Kitcher, P. (1981), "Explanatory Unification", *Philosophy of Science* 48: 507-31.
- McAllister, J. (1996), *Beauty and Revolution in Science*. Ithaca: Cornell University Press.
- Mandl, F. and G. Shaw (1984), *Quantum Field Theory*. New York: Wiley.
- Maudlin, T. (1996) "On the Unification of Physics", *Journal Of Philosophy* 93, 129-44.
- Maxwell, N. (1968), "Can there be Necessary Connections between Successive Events?", *British Journal for the Philosophy of Science* 19, 1-25.
- Maxwell, N. (1974), "The Rationality of Scientific Discovery", *Philosophy of Science* 41, 1974, 123-53 and 247-95.
- Maxwell, N. (1984), *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science*, Basil Blackwell, Oxford.
- Maxwell, N. (1993) Induction and Scientific Realism: Einstein versus van Fraassen, *The British Journal for the Philosophy of Science* 44, 1993, 61-79, 81-101 and 275-305.
- Maxwell, N. (1998), *The Comprehensibility of the Universe: A New Conception of Science*. Oxford: Oxford University Press.
- Moriyasu, K. (1983), *An Elementary Primer for Gauge Theory*. Singapore: World Scientific.
- Morrison, M. (2000), *Unifying Scientific Theories*. Cambridge: Cambridge University Press.
- Newton, I. (1962), *Principia, ii*, trans. In 1729 by A. Motte and F. Cajori (first published 1687). California: California University Press.
- Popper, K. (1959), *The Logic of Scientific Discovery*. London: Hutchinson.
- Salmon, W. ((1989), *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Schurz, G. (1999), "Explanation as Unification", *Synthese* 120, 95-114.
- Schutz, B. (1989), *A first course in general relativity*. Cambridge: Cambridge University Press.
- Watkins, J. (1984), *Science and Scepticism*. Princeton: Princeton University Press.
- Weber, E. (1999), "Unification", *Synthese* 118, 479-99.
- Weyl, H. (1963), *Philosophy of Mathematics and Natural Science*. New York: Athenium (first published in German in 1926).

## Notes

<sup>1</sup> See Maxwell (1998, 47-56) for additional arguments in support of the point.

<sup>2</sup> To say that  $T_1$  and  $T_2$  are different formulations of one and the same theory, with precisely the same *content*, is to say that  $T_1$  and  $T_2$  make precisely the same assertion about the world. That is,  $T_1$  and  $T_2$  express the same proposition. They have the same truth conditions. Whatever makes  $T_1$  true (or false) also makes  $T_2$  true (or false).

<sup>3</sup> If the theory is formulated as a set of differential equations, then what is invariant throughout the possible phenomena to which the theory applies is what is asserted by the physically interpreted set of differential equations. Laws specifying precisely how diverse

---

physical states evolve in space and time may be quite diverse in character: what matters is that they are all solutions of the same set of differential equations.

<sup>4</sup> Counting entities is rendered a little less ambiguous if a system of  $M$  particles is counted as (a somewhat peculiar) field. This means that  $M$  particles all of the same kind (i.e. with the same dynamic properties) is counted as one entity. In the text I continue to adopt the convention that  $M$  particles all the same dynamically represents one kind of entity, rather than one entity.

<sup>5</sup> For accounts of the locally gauge invariant structure of quantum field theories see: Moriyasu (1983), Aitchison and Hey (1982: part III), and Griffiths (1987, ch. 11). For introductory accounts of group theory as it arises in physics see Isham (1989) or Jones (1990).

<sup>6</sup> For accounts of spontaneous symmetry breaking see Moriyasu (1983), Mandl and Shaw (1984), Griffiths (1987, ch. 11).

<sup>7</sup> This account of unity radically simplifies and improves on the account given in Maxwell (1998, chs. 3 and 4).

<sup>8</sup> I am grateful to Jos Uffink for drawing my attention to the two objections just discussed.

<sup>9</sup> For further discussion of simplicity, and how terminological simplicity can be related to unity, see Maxwell (1998, 110-3 and 157-9).

<sup>10</sup> For a discussion of such “approximate derivations”, the conclusion being strictly incompatible with the premises, see Maxwell (1998, 211-7).

<sup>11</sup> Of the eight facets of theoretical unity distinguished above, (1) to (3) have been persistent requirements acceptable physical theories must satisfy since Newton, with (4) and (5) becoming, perhaps, more prominent in the 19<sup>th</sup> century, and (6) to (8) only becoming explicit requirements in the 20<sup>th</sup> century. With hindsight, we can see that (1) to (8) are all facets of the same conception of theoretical unity. Nevertheless, requiring only that a theory satisfy (1) to (3) is less demanding than requiring that it satisfy (1) to (7), in turn less demanding than that it satisfy (1) to (8).