The Completeness of Physics

David Jon Spurrett
Durban, 1999

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Philosophy, University of Natal (1999).

"Let me now re-emphasise the extreme looseness of the structure of all objects."
Joseph Conrad.

Abstract

The present work is focussed on the completeness of physics, or what is here called the Completeness Thesis: the claim that the domain of the physical is causally closed. Two major questions are tackled: How best is the Completeness Thesis to be formulated? What can be said in defence of the Completeness Thesis? My principal conclusions are that the Completeness Thesis can be coherently formulated, and that the evidence in favour if it significantly outweighs that against it.

In opposition to those who argue that formulation is impossible because no account of what is to count as physical can be provided, I argue that as long as the purpose of the argument in which the account is to be used are borne in mind there are no significant difficulties. The account of the physical which I develop holds as physical whatever is needed to fix the likelihood of pre-theoretically given physical effects, and hypothesises in addition that no chemical, biological or psychological factors will be needed in this way. The thus formulated Completeness Thesis is coherent, and has significant empirical content.

In opposition to those who defend the doctrine of emergentism by means of philosophical arguments I contend that those arguments are flawed, setting up misleading dichotomies between needlessly attenuated alternatives and assuming the truth of what is to be proved. Against those who defend emergentism by appeal to the evidence, I argue that the history of science since the nineteenth century shows clearly that the empirical credentials of the view that the world is causally closed at the level of a small number of purely physical forces and types of energy is stronger than ever, and the credentials of emergentism correspondingly weaker.

In opposition to those who argue that difficulties with reductionism point to the implausibility of the Completeness Thesis I argue that completeness in no way entails the kinds of reductionism which give rise to the difficulties in question. I argue further that the truth of the Completeness Thesis is in fact compatible with a great deal of taxonomic disorder and the impossibility of any general reduction of non-fundamental descriptions to fundamental ones.

In opposition to those who argue that the epistemological credentials of fundamental physical laws are poor, and that those laws should in fact be seen as false, I contend that truth preserving accounts of fundamental laws can be developed. Developing such an account, I test it by considering cases of the composition of forces and causes, where what takes place is different to what is predicted by reference to any single law, and argue that viewing laws as tendencies allows their truth to be preserved, and sense to be made of both the experimental discovery of laws, and the fact that composition enables accurate prediction in at least some cases.

Acknowledgements

I have run up an extraordinary collection of debts of various kinds in the course of researching and writing this thesis. It is a pleasure to acknowledge the following selection of them.

I would have been unable to begin this research without the University of Natal Doctoral Research Scholarship (1994 - 1996), and would have been unable to continue without the Centre for Science Development Doctoral Scholarship (1995). A significant portion of the more specialised material was researched in London during a visit made possible by a Centre for Science Development Overseas Doctoral Scholarship (1998). A great deal more research and philosophical engagement was permitted by the award of a twelve month scholarship by the Association of Commonwealth Universities which enabled me to spend an entire academic year (1998 to 1999) at King's, and to travel to and present papers at a number of conferences in England and Europe. Finally, many of the arguments which appear here were field-tested and refined at conferences around South Africa which I attended by means of grants from the research office at the University of Natal. The financial assistance of the University of Natal, the Centre for Science Development (South Africa) and the Association of Commonwealth Universities towards this research is hereby acknowledged. Opinions expressed here and conclusions arrived at are those of the author, and do not necessarily represent those of any of these funding bodies. I am similarly grateful to King's College London for making me a Visiting Research Fellow for a three year period beginning with my first visit in April 1998.

On a more personal note I should begin by singling out Daniel Herwitz, my head of department in Durban, who encouraged me when I needed it, read what I wrote when I gave it to him, and left me in supportive peace for the rest of the time. I especially appreciate his forbearance in allowing me to use students as guinea pigs by developing some of my research through graduate courses and seminars. David Papineau of King's College London has been helpful and energetically critical at various stages as well. I am especially indebted to him for proposing me as a speaker at the one-day conference on the completeness of physics held in London on 28 May 1999. Derek Wang at the University of Natal has been my major partner in philosophical discussion for nearly a decade, as well as being one of the supervisors of this thesis. During this time there are very few topics we have not discussed, always beneficially. While at King's I was supervised for some of the time by John Milton, whose understanding and knowledge of the history of science contributed vastly to my grasp of the historical backdrop to the topics discussed here, and for the rest by Keith Hossack whose standards of argument and precision were demanding and energising. Many discussions of bizarre counterfactual worlds (my favourite had quarks being made of waltzing angels) which have been removed from the current version of this thesis occurred in the course of discussions with Keith. In Durban I was also supervised by Andries

Gouws, whose advice at every stage I needed it was always finely judged and profoundly helpful, and who made many helpful and detailed comments on a very late version of the text.

Various parts of this thesis were presented at the King's College senior seminar group, and it gives me pleasure to acknowledge the lively questions, criticisms, suggestions, and much needed solidarity during the consumption of certain pints, from those in the group. I am especially grateful to Paul Sheehy, Pierre Cruse, Guy Longworth, Sophie Allen, Daniel Hill, and Eilert Sundt-Ohlsen.

Thanks also go to David Davies (McGill) and Alan Chalmers (Sydney) who kindly sent me copies of relevant papers, Gabe Segal (King's) for lively discussion on several key points, and to Deepak Mistrey (Natal, Durban) for comments on parts of a draft version.

Vital non-academic support came from Inger Nilsen and Jody Fraser (Natal, Durban), Stephen Dunnett (British Council) and Anna Gane (Association of Commonwealth Universities). Thanks to all of them for taking care of innumerable matters, leaving me free to work.

On an even more personal note I must thank my family, especially my parents and paternal grandparents, for their support and encouragement.

My good friends Catherine Whitfield, Adrian Bellengere, Neal Crankshaw, Carey-Ann Jackson, Kim Jones, Nell Aubrey, Andrea Kropf and Caron Angelbauer kept me sane at various times. While in London I stayed happily with Rosemary Creedy, John and Julia Spencer, not to mention Stuart, Maria, Sabine and Frank, the unspeakable degenerates at number 14. Very special thanks to Guy Morpuss and Julie Davis, who spoilt me rotten at various times and in highly agreeable ways.

Finally, the beverage vending machine in the ground floor of the philosophy department at King's for the substance which it insisted was espresso and which I came to know simply as number 56. *Whatever* it was I drank heinous quantities of it, and it always got me through the day.

Note

Parts of what follows have already been published in various forms, or have been presented at conferences. In all of the latter cases I am indebted to those present for comments and criticism. The details are as follows:

Chapter One includes material presented under the titles '*Could* Physicalism be True?' at the Spring Colloquium of South African analytic philosophers, Drakensburg, South Africa on 15 September 1997, 'For the Completeness of Physics' at a one day conference on the Completeness of Physics at Senate House, University of London on 28 May 1999, and relates to material appearing in the paper 'A Note on the Completeness of 'Physics'' written with David Papineau (Spurrett and Papineau 1999).

Chapter Two includes material presented under the titles 'Bhaskar on Open and Closed Systems', at the 25th annual congress of the Philosophical Society of South Africa (PSSA), University of the Orange Free State, South Africa on 20 January 1999, 'Transcendental Realism and Quantum Mechanics', at the 22nd annual congress of the PSSA in Stellenbosch, South Africa in January 1996, and material appearing in the paper 'Beyond Determinism' (Spurrett 1997).

Chapter Three develops some material appearing in a review article of Cilliers, P. (1998) *Complexity and postmodernism*, London: Routledge (Spurrett 1999b).

Chapter Four includes material presented under the title 'Fundamental Laws' at the 25th annual philosophy of science conference at the Inter University Centre, Dubrovnik, Croatia on 14 April 1999, and published as 'Fundamental Laws and the Completeness of Physics' (Spurrett 1999c), and develops aspects of material presented under the title 'Lyotard and the Postmodern Misunderstanding of Physics' at the Sceptic Tank in May 1996, at the First Colloquium of the Wits Transdisciplinary Study Group, Johannesburg, South Africa, and later at the Warwick University Philosophy of Science seminar in May 1998, then published under the same title (Spurrett 1999a).

Declaration

This thesis is entirely the work of the author, David Jon Spurrett.

Table of Contents

Introduction

Physics is complete. That, to come right out with it, and shorn of various clarifications and the inevitable qualifications and hedgings to be detailed in due course, is the eventual conclusion of the present work. The term 'completeness of physics' in this context is due to Papineau (1993: 13) and can also be formulated as the claim that physics is causally closed, that is, roughly, that all physical effects have physical ancestries sufficient to fix their likelihoods, given purely physical laws.

In what follows I pursue two major questions. First: how is the thesis that physics is complete, which from here on I call the Completeness Thesis, and sometimes simply the Thesis, properly to be formulated? Second, given the formulation: how can we go about considering whether the Completeness Thesis is true or otherwise philosophically credible? Both questions, especially the second, divide into a collection of subsidiary problems. Not only that, the two questions are intimately related: some forms of the Completeness Thesis are more credible than others, and the task of defending the Thesis is to a large extent the task of establishing which commitments are and are not essential to it, and how those which are essential should best be conceived. Later in this introduction I give an overview of the most significant of these distinctions and divisions. For the present, though, I want to spend some time looking at the philosophical uses to which the Completeness Thesis is typically put. It is the role of the Thesis in these debates which makes it an interesting question to what extent the Thesis is defensible. The two areas in which the assertion that the domain of the physical is causally closed plays an especially significant role are physicalism and the unity of science.

1.1. Physicalism

The Completeness Thesis is one of a small number of things about which those who call themselves physicalists are in broad agreement. There is a great deal about which they disagree, since some but not all physicalists are reductionists, and those who are have a selection of types of reductionism available for debate and endorsement.[1] Some think the mental is 'anomalous' in the sense of Davidson (1970) while others don't, some are eliminativists,[2] some are functionalists,[3] some connectionists[4] and so on. Those who endorse the notion of the supervenience of the mental on the physical are as likely to view some other supervenience theorist as their mortal foe as they are a more obvious enemy such as a substance dualist, if for no other reason

---

1 Various forms of reductionist thesis are discussed in more detail in Chapters Two and Three below. See especially Chapter Two section (6) and Chapter Three sections (3.1) and (7).

2 e.g. Churchland, P. S. (1986) and Churchland, P. M. (1981, 1985).

3 e.g. Fodor (1974, 1990).

4 e.g. Bechtel, W. and Abrahamsen, A. (1991) and Cummins (1989).

than that there is such a variety of forms of supervenience available. Popular flavours include 'global', 'strong', and 'weak' supervenience, as well as more exotic brews such as 'superdupervenience'.[5]

Through all these debates and squabbles, though, physicalists are united in endorsement of the Completeness Thesis, which is one of the things which makes it interesting to investigate the possible grounds for its acceptance or rejection. These thinkers may not refer directly to either the Thesis or the issue of the causal closure of physics, but the key notion can clearly be discerned in various ways in what they do say. So, for example, Davidson (1970) has it that physical theory 'promises to provide a comprehensive closed system guaranteed to yield a standardized, unique description of every physical event couched in a vocabulary amenable to law', Armstrong (1980) that the 'cause of all human (and animal) movements lies solely in physical processes working solely according to the laws of physics'. Besides these and many other cases which clearly show that causal closure is at stake, examples of enthusiasm for physics abound: Blackburn (1993) takes it as proper that we recognise what he calls the 'sovereignty of physics', Hellman and Thompson (1975), regard mathematical physics as the 'most basic and comprehensive of the sciences' and so on. Despite some variation in how these and other physicalists regard the Completeness Thesis as holding true, variations largely related to their differing commitments regarding laws, causality and ontology, they are unanimous in their acceptance of the key notion of the causal closure of the physical. Kim makes this aspect of the relationship between physicalism and the Completeness Thesis quite clear when he contends that it is 'safe to assume that no serious physicalist could accept' the prospect of a nonphysical cause of a physical outcome (1998: 40).

The thesis that physics is complete is a more modest claim than that associated with any particular version of physicalism. It is important that I make clear that in what follows here I am not specifically concerned to defend *any* form of physicalism. Indeed, for reasons to be detailed shortly, I would like to avoid discussing physicalism as much as possible, and focus rather on the Completeness Thesis itself. The way in which the completeness issue is more limited in scope than that of physicalism is simple: if physics was complete then that fact would not by itself decide how things were with anything else, including the mental.[6] So there might be nothing else, or whatever else there was might be epiphenomenal, or relate to the physical along the lines of one of the many theories of reduction, identity or supervenience presently on offer, and so on. Irrespective of that, though, *any* version of physicalism has to be committed to the completeness of physics, since without that premise mental events would not be 'anomalous' (Davidson 1970), mental states would not

---

5 For a useful survey of different types of supervenience see McLaughlin (1995). Superdupervenience was floated in Horgan (1993). Kim (1993) contains a range of papers both energetically defending and developing supervenience, also including some more recent and critical papers which point partly away from supervenience and towards the more reductionist and basically Lewisian line defended in Kim (1998).

6 On this point I am disagreeing with Blackburn (1993: 233) and anyone else who holds that the physics should also fix the supervenience relations as well. Although the point is not critical for my argument here, I do not regard it as a *physical* question what relation (supervenience or otherwise) obtains between the physical and anything else, including the mental. That requires a separate argument.

run the risk of being 'nomological danglers' (Feigl 1958), and there would be no arguments for the identity of the mental and physical via the rejection of overdetermination (e.g. Papineau 1993).

Another way of making this point is to say that physicalism is typically a thesis about *two* classes of properties, one of them physical, where the non-physical class is supposed to be determined in some way by the physical class. This is the core idea of physicalism, the notion of 'one realm of facts determining another' (Hellman and Thompson 1975: 557). For such a position to be plausible at all requires that the determining class of properties, in this case the physical ones, be suitable for that role. And a crucial aspect of this suitability is causal isolation from the second set of properties: some class of non-physical properties (mental ones say) cannot be supposed to be determined by physical ones unless the physical properties are not themselves determined by any mental ones. The 'best case scenario' for the set of determining properties is that it be causally self-sufficient, that is to say 'complete' in the sense at issue here.

All physicalists think, then, that the physical is causally closed. This agreement is, unsurprisingly, accompanied by a great deal of controversy and disagreement. Physicalist confidence that physics is complete is, for the most part, attended by a fairly laconic attitude to the question what is to count as physical, and what the relation between the expected physical causal closure and present physical knowledge might be.[7] Not only that, although relied on for a range of argumentative purposes the Completeness Thesis is rarely the object of anything which deserves the name of a defence.

In a recent book Kim (1998), for example, uses the completeness of physics, or what he calls the 'causal closure of physics' as a major premise in a number of arguments (e.g. 1998: 37-38, 40) but does not attempt to argue for the truth of the Thesis in any way, nor does he devote any significant attention to the question what should be counted as 'physical'. Kim's near silence on these questions places him at one extreme, since some physicalists (e.g. Papineau 1990, 1993, Poland 1994) are far more concerned with them. Whether or not the confidence of Kim and others turns out to be appropriate, though, it seems clear that some kind of enquiry into the defensibility of the Completeness Thesis is an important component in any comprehensive appraisal of physicalism.

This involves something of a shift of emphasis: as well as not generally being explicitly defended, the Completeness Thesis is very likely to appear as a premise in arguments for any particular version of physicalism. This is another way of stating something which has been said already, which is that physicalism is typically a thesis about two classes of properties to the effect that the one (physical) in some way determines another, where the other set (mental or biological for example) is the object of the physicalism in question. The relation of determination may be any of those (reduction, supervenience, etc.) mentioned above, or others not

---

[7] Smith (1993: 244) goes so far as to say that it is 'astonishing how little is said by physicalists themselves on the presumably crucial question of what it is to be physical.'

yet thought of, but unless physics is taken to be complete it is unlikely that such an argument can get off the ground for the simple reason that a set of causally closed properties is more likely to form a stable foundation.

It is worth noting that completeness need not be restricted to physics. A particularly common form of argument for physicalism starting from completeness is the 'causal' argument assuming that physics is complete, noting that some mental events have physical effects, and arguing that on pain of overdetermination the identity of the mental with the physical should be accepted. This type of argument is discussed in Spurrett and Papineau (1999), where it is argued that the completeness premise need not be one about the physical, but that any plausible completeness thesis will do, and will enable particular forms of the causal argument to be run. The general form of the causal argument is to move from (a) the completeness of some X, and (b) the fact that some Y have X effects to (c) the identity of X and Y, since the alternative is to accept the overdetermination of those X outcomes having Y causes.

In particular Spurrett and Papineau consider instances of the argument using the completeness of the non-mental and of the quantitative, and note that it is possible to conceive of worlds where one is complete but not the other, and hence where specific identity theses can be defended with the causal argument in each case. They conclude that it may not be especially important to settle what physics means as long as there are plausible candidates for major premises in the causal argument.

Although the Spurrett and Papineau approach has some virtues, it is also worth pointing out that it does not clear away the question of whether physics itself is, or could be, complete, but rather indicates that that particular question may well be bypassed for certain purposes. More to the point, it seems reasonable to say that the apparent plausibility of the candidate completeness premises considered by Papineau and Spurrett is perhaps lower than that of 'physics' even *without* a particularly definite sense of what physics might be.[8] Indeed it is possible that such plausibility as the premises they consider have is partly a consequence of the credibility of the notion that physics itself is complete. In any event, my interest in what follows is the completeness of physics, rather than anything else.

If the Completeness Thesis turns out to be indefensible, then, this would do serious damage to the credibility of any form of physicalism. If on the other hand it turns out to be strongly justifiable this fact would be a source of support for the wider physicalist project, no matter what outstanding problems remained to be solved for that project.[9]

---

[8] I owe this point to Keith Hossack at King's College London.

[9] I make no attempt at a general survey of these problems. Some, which are closely related to the discussions of reductionism in Chapters Two and Three are returned to in the conclusion of the present work.

1.2. The Unity of Science

Besides its role in physicalism, the notion of the completeness of physics plays a significant role in some accounts of the unity of science. The notion that science in its modern[10] form could be unified is not one which commands widespread contemporary support. Although some relatively recent work[11] continues to focus on the problem, it is something of a minority interest debate compared to a few decades ago, in the heyday of Nagel and others. It is as well to bear in mind, then, that not too long ago the expectation that the whole of science could come to form a single, connected and consistent body of knowledge was widely shared by scientists and philosophers.

There have been various different programmes which could be described as seeking the unity of science. These include Descartes' mathematisation of matter coupled with his 'mathematical' epistemology,[12] Kant's unification of the phenomenal world under the intuitions of space and time, and the categories (including causation) all in conformity with the idea of law,[13] Mill's ideal of a complete and deductive system of science as the culmination of a programme of inductive research.[14]

It is important to note that not all programmes for the unification of science rely on or entrench the notion of the completeness of physics. Not one of the three examples just listed do: Descartes was a dualistic interactionist and Mill an emergentist. Without getting side-tracked by Kant exegesis here, one way of reading Kant is as holding that a unified, and deterministic whether or not purely physical, science is possible with respect to the phenomenal world, but that this is compatible with the operation and real effectiveness of non-physical 'rational' causes in the noumenal world.

Even though there is no necessary connection between the unity of science and the completeness of physics, it is fair to say that the dominant twentieth century version of the unity ideal is one closely connected with the completeness of physics. The relevant ideal is, in a nutshell, Mill stripped of the emergentism, which is replaced with a general commitment to a programme of reductionism concerning the relationship between

---

10 That is, roughly, its post-renaissance form. Whether there is any useful or interesting sense in which Aristotelean physics, for example, could be said to be complete is a question which is not relevant to my concerns, which are primarily focussed on contemporary and even future physics.

11 e.g. Causey (1987).

12 Descartes' most detailed account of science is in his *Principles of Philosophy* (Descartes 1991). See also Garber (1992).

13 For general features of Kant's account of the epistemological science see the *Critique of Pure Reason* (Kant 1929) and the collection *Philosophy of Material Nature* (Kant 1985). For the ways in which Kant sees a regulative role for the idea of law see the latter parts, especially, of the *Critique of Judgement* (Kant 1987).

14 See Mill (1972). Mill is discussed at various points in what follows, especially Chapter Two section (2.1) and Chapter Four section (5.4).

non-physical empirical knowledge and physics. Examples of this way of thinking include Nagel (1961) and Oppenheim and Putnam (1958).[15]

Although the Completeness Thesis is closely related to these versions of the unity of science, the two are independent. Just as some types of unity (Mill's, Descartes', Kant's) do not require the completeness of physics, and actually imply its falsity, so too completeness is possible without the automatic implication of unity. This is so for the simple reason that there is more to unity than completeness, and physics could well be causally closed without the required reductions of other theories and domains of science being possible. The association between completeness and unity, then, is weaker than that between completeness and physicalism. Nonetheless completeness and unity are often packaged together, and some criticisms of the unity of science make a point of also targeting the Completeness Thesis, sometimes on the supposition that the two stand or fall together.[16]

1.3. Parameters

Although the significance of the Completeness Thesis can, I hope, clearly be seen from the considerations just discussed, the question of how best to approach the issue of its truth is a thorny one. Given the importance of the Completeness Thesis for physicalism it is not surprising that a significant portion of the attacks on it are directed from within the philosophy of mind. My approach in what follows avoids as far as possible this way of doing things. On the face of it this may seem somewhat idiosyncratic or even perverse, but I believe a compelling justification can be given.

To begin with, it appears to me as though many of the mind-motivated criticisms of the Completeness Thesis rely for their functioning upon considerations which may themselves be either less well understood or more contentious than those relating to the Thesis itself. A personal example may make more clear what I mean. My Masters thesis (Spurrett 1994) was a treatment of free will and the compatibilism debate. In it I argued for an uncompromisingly incompatibilist position, and also for the thesis that the principal threat to freedom was not determinism *per se*, as was traditionally perceived to be the case, but rather physical closure. To a significant extent I made physicalism my target, objecting that the versions of it which I considered were unable to prevent a slide into epiphenomenalism, hence being unable to sustain any reasonable account of agency. My conclusions included the suggestion that some form of emergentism was the most likely prospect for salvaging agency. (While still speaking in this personal register, I might as well admit that when I began the research which culminates in this doctoral thesis *defending* completeness my project was to follow up the

---

15 Both of which are discussed at more length in Chapter Three below, especially sections (1.2) and (6.1).

16 Dupré (e.g. 1993) is a striking example. See Chapter Three below.

argument of the earlier thesis with a direct assault on the Completeness Thesis, thus striking physicalism at its foundations.)

While having no specific complaint with the arguments I developed on the subject of free will, except for my views on emergentism which are largely reversed in the second chapter of the present work, I now see the order of business as being the other way around. The analyses of freedom, intentionality, representation, rationality, consciousness and so on are difficult enough as it is. The range of available positions is too large and the ground too shifting for it to seem optimal to extract conclusions about such a matter as physical closure from the central considerations of those debates, even though there are those, including myself, who have been willing to do just that. Other examples abound: transcendental arguments against physicalism starting with considerations about rationality (Walker 1993), from our knowledge of set theory (Hale 1993) and more. One anti-physicalist line (Lucas 1961) starts from Gödel's incompleteness theorem and contends (loosely) that since the theorem shows that no algorithmic system with a finite set of axioms can identify all the true theorems in its domain, and also that a purely physical algorithm system would be restricted in the same way, that the processes by which we think cannot be based entirely on purely physical conditions, since we are not so limited (being able to appreciate Gödel's theorem). I have no intention of engaging with this argument here. This has nothing to do with whether I think that it is good or bad. What bothers me is that the road from one end of the argument to the other crosses problems in the foundations of algebra, computer science, logic, psychology, and demands engagement with the problem of the nature of truth and mathematical knowledge. Given that all these areas are contested it seems as though by selecting a suitable combination of loyalties in each area it would be possible to defend, with reputable arguments at each point, a very large number of different theses, and that the same flexibility would be available to one's opponents.

I do not mean to disparage this kind of philosophical work, merely to suggest that it seems unlikely to be decisive. Furthermore arguments of this sort are typically directed at some or other specific version of physicalism, whereas my interest is with the Completeness Thesis, which need not be tied to any of them. In what follows my approach is, as far as possible, to consider the question of the truth of the Completeness Thesis as a matter in its own right, independently of consideration in the philosophy of mind. This is not to say that the present enquiry is entirely isolated from physicalism more widely or the philosophy of mind. Not at all. The point is rather that I hope that the order of priority goes the other way: if the Completeness Thesis is not defensible, then the ongoing attempts to refine, extend and modify physicalism are less interesting and important than they would be otherwise. If, on the other hand, there is a strong case to be made in favour of the Thesis then defects of particular versions of physicalism should be regarded as pointing to the need for further refinement, extension, modification, rather than suggesting the abandonment of the wider physicalist project.

It turns out, within the scope of the arguments and evidence which are considered here, that the Completeness Thesis is defensible, although in the course of the defence it is necessary to make various changes to the ways in which the Thesis is typically framed, most notably as far as the notion of how a law of nature is to be understood goes. That such changes have to be made at all indicates that the approach I take has its virtues. Put another way, working out the best way to understand a defensible notion of physical closure seems like a worthwhile contribution to physicalism in its own right.

1.4. Overview

This work is divided into four chapters, the first mostly focussed on problems of formulation, and the following three dedicated to the consideration of major lines of criticism of the Completeness Thesis. The three directions of criticism which are discussed are not exhaustive of the possibilities, but they do relate to each other in significant ways, and collectively represent the most serious challenges to the Thesis within the terms of the restrictions adopted above. In slightly more detail the content of the following chapters is as follows.

The principal difficulty which I address in the first chapter is that of formulation. In order to pursue the question of the truth or falsehood of the Completeness Thesis it is necessary to defend some position on how 'physical' is best to be understood. A variety of approaches to that question, as well as sceptical and critical responses, are discussed, and a strategy proposed for the purposes of the chapters which follow. The major conclusions of this chapter are that a negative characterisation of 'physical' in terms of what it does *not* include is suitable for my purposes, and that thus formulated the Completeness Thesis has significant *prima facie* plausibility.

Chapter Two is devoted to emergentism. This doctrine, after a period of relative unpopularity around the middle of the twentieth century, had enjoyed a recent revival of support and scholarly attention. In the form which I call 'strong' emergentism, it is incompatible with the truth of the Completeness Thesis, since strong emergentism is committed to the existence of non-redundant and non-physical causal antecedents of some physical outcomes. In this chapter I argue that one supposed philosophical argument for emergentism (due to Bhaskar) is deeply flawed, and that the supposed evidence for strong emergentism does not represent a serious challenge to the Completeness Thesis.

Emergentists typically allow that physical laws hold to varying extents for *all* phenomena, and that they also form the foundation for the causal order of the world. Emergentist rejection of the Thesis is made in the name of a hierarchy of laws, only the bottom level of which is physical. Chapters Three and Four are concerned with positions which reject both the Completeness Thesis *and* emergentism, in this case in the name of a commitment to general epistemic and ontological disorder.

Dupré's defence of disorder is discussed in Chapter Three. Dupré's principal thesis is that, given the manifest disorder of our taxonomic practices, we should recognise the force of the inference to the existence of widespread and varied disorder in the ontology of the world. I argue against Dupré that his inference is a *non sequitur*, and that completeness is entirely compatible with a great deal of taxonomic disorder. Although further argument on this point may seem unnecessary given the past few decades of attention to non-reductive physicalism, this chapter emphasises that the Completeness of Physics in no way implies the truth of any general reductionist thesis.

Finally in Chapter Four I consider Cartwright's philosophy of science, and especially her contention that the laws of physics are lies, and that we should regard the world as being an ontological patchwork on the basis of a similar inference to Dupré's, in her case from the patchwork of laws with which we attempt to describe the world. Since the responses to both emergentism and Dupré rely to some extent on the notion of the truth of fundamental physical laws, Cartwright's challenge is a serious one. I argue that what is most dangerous in her position can be dealt with in a way which saves completeness, but in a significantly changed form. Cartwright may not be correct about everything to do with laws, but she is right enough to have shown some versions of the Completeness Thesis to be utterly untenable.

Chapter One: Formulation and Preliminary Defence

> Mathematical Physics, as the most basic and comprehensive of the sciences, occupies a special position with respect to the all-over scientific framework. In its loosest sense physicalism is a recognition of this special position. (Hellman and Thompson 1975: 551).

1. Introduction

The problem of clearly formulating either physicalism or the Completeness Thesis is a significant one, and it has even been suggested at times that the task is impossible, or at least that it can only be carried out at philosophically unacceptable costs. Examples of positions claiming either that formulation is impossible, or at least hitherto unsuccessful include Chomsky (1968), Stroud (1987), Crane and Mellor (1990), Daly (1998).[1]

In what follows I attempt to develop a reasonable formulation of the Completeness Thesis, noting some of the dangers which beset that project, and also discussing a selection of pessimistic and critical suggestions. The criticisms are considered in order both to see how the proposed formulation bears up, and to modify it where necessary. It is my contention that the requisite senses of both 'complete' and 'physical' can be provided, in such a way that the claim that physics is, or could be, complete is empirical and significant.

After a formulation of the Completeness Thesis has been settled on, I offer some brief observations on the *prima facie* plausibility of the thesis, and also on the question of what might reasonably count as either argument or evidence for or against it. The most natural way to proceed is to begin by considering the notions of completeness and physics separately. I begin with completeness.

2. Completeness

Of the two key notions of physics and completeness here, that of completeness is the more easily clarified. Completeness is, in the present context, a matter of causal closure. Papineau (1993: 13) refers to the 'internal' completeness of physics, which makes clear that completeness at issue is a matter of the self-sufficiency of a science. As a preliminary gloss we can say that a science would be complete in the sense required if occurrences falling within its scope could be predicted, or have their probabilities of occurring ascertained, solely by reference to considerations also within its scope. For the time being I deliberately leave the notion of the 'scope' of a science relatively unspecified. Although I shortly turn to the problem of the scope of physics specifically, let me now say simply that the scope of a science will be a characteristic of the science

---

[1] For the most part these positions are explicitly opposed to physicalism (typically physicalism about the mental) rather than the Completeness Thesis, but the issue which cuts across these debates is that of determining the reference of 'physical' as will become clear in what follows.

itself, and generally be a consequence of the class of phenomena which its practitioners seek to describe, explain and perhaps predict.

The suggestion just offered lays down a criterion for completeness on the part of a *science*. Completeness, though, can also be understood in a way independent of science or knowledge, as a property of some domain of reality. In this case we might say that some category of reality would be complete if outcomes which fell within the relevant category were necessitated or had their likelihood of occurring fixed solely by factors also within the same category. Such a state of affairs could obtain or fail to obtain whether or not it was known to do so.

When it is important to mark the distinction between these possibilities I will refer to epistemological and ontological completeness respectively. In what follows my primary interest is in ontological completeness, although the plausibility of the ontological Completeness Thesis is inevitably tangled up with questions about epistemological completeness. This is because the two aspects of completeness are related: when it is suggested that some science holds out the prospect of being complete it is usually because there are taken to be reasons to expect that some domain of the world is itself causally closed in such a way that a complete science of it is 'in principle' possible.[2]

The claim of completeness is rarely made on behalf of any science or domain of reality. Meteorology, for example, concerns phenomena which may have antecedents in vulcanology or economics, and any reasonable account of the antecedents of, say, the build-up of greenhouse gasses (if we allow that this effect can be classified as a meteorological phenomenon) could be expected to refer, as well as to meteorological factors themselves, to chemistry, the history of technology, economics, geology and so on. It is worth noting, though, that completeness claims *have* been made from time to time on behalf of various sciences. Durkheim's dictum that the explanation of any sociological fact is another sociological fact (1964: 95, 110) is, at least on the face of it, an example of this, and Leibniz's metaphysics of preëstablished harmony can be seen as simultaneously making the claim on behalf of both the physical and the mental.[3] Most usually, though, at least in contemporary discussion (and on and off since the seventeenth century) the claim is made with respect to *physics*. So, for example, Papineau says, of physics specifically, that:

> I take it that physics, unlike the other special sciences, is complete, in the sense that all physical events are determined, or have their chances determined, by prior *physical* events according to *physical* laws. In

---

2 The cheerful optimism of the 'in principle' clause here is subjected to vigorous criticism in the chapters which follow.

3 This clearly contrasts his position with the standard reading of Descartes' dualism, where neither the mental nor the physical realms would be complete at least as long as there was one thinking thing united with a body. Leibniz's allegations against Descartes in the name of the causal closure of the physical are critically discussed in my 'What's Really Wrong with Descartes' Dualism' (in preparation).

> other words we need never look beyond the realm of the physical in order to identify a set of antecedents which fixes the chances of every physical occurrence. A purely physical specification, plus physical laws, will always suffice to tell us what is physically going to happen, insofar as that can be foretold at all (1993: 16).

Papineau's account of completeness in the passage just quoted here combines elements of epistemological and ontological completeness, referring to 'physical specifications' and the possibility of prediction, but also to the 'realm of the physical'. It will be useful to distinguish the two variants more precisely for the purposes of what follows. As already noted epistemological completeness involves prediction. This need not be deterministic, as long as the science is capable of fixing the *likelihood* of the occurrences it seeks to predict. In order to be complete in this way a science would clearly need resources sufficient to describe the phenomena within its scope. Not only that, the most likely way of producing predictions would be by means of laws, which is to say that completeness would also require that some set of laws be known which were sufficient to fix the likelihood of the occurrences in question, given suitable data concerning initial (or prior) conditions. So we can characterise the epistemological form of the Completeness Thesis as follows.

> **ECT1**: A science **X** is complete iff **X** is capable of fixing the likelihood of **X**-type events solely by reference to prior **X**-events and **X**-laws.

This formulation does not attempt to embrace the possibilities of either backward or simultaneous causation. Although this may seem like a shortcoming it is worth noting that the criticisms of the Completeness Thesis considered in the following three chapters make no reference to either possibility, so the above version should be sufficient for present purposes.[4] The epistemological Completeness Thesis says that given the state of a system at some time and a complete science of that system in terms of laws, it will be possible to determine states of the system at later times. An especially well known image of epistemological completeness can be seen in the famous passage from Laplace's *Philosophical Essay on Probabilities* where Laplace considers what would be possible for a suitably powerful computing intelligence assuming that the intellect in question could be supplied with sufficiently precise data. Laplace states that if given 'all the forces by which nature is animated and the respective situation of the beings who compose it,' his hypothetical 'vast intellect' could, 'embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it nothing could be uncertain; and the future as the past would be present to its eyes' (Laplace 1951: 4).

It is important to note that this formulation of epistemological completeness does not rule out the possibility of *other* accounts of the likelihood of **X**-type events. In the event that physics, say, is complete, the

---

[4] Some of the proposals concerning holism and emergent causation in quantum mechanics discussed in Chapter Two (see especially section 7.2) are akin to simultaneous causation.

truth of the Completeness Thesis on its own should not rule out the nomological overdetermination of physical outcomes, as would occur if there were, for example, both physical and mental law-like accounts of the ancestry of some physical event. Even if physics is complete, it requires further argument to arrive at any particular form of physicalism, a point noted in the introduction to this thesis.

Ontological completeness on the other hand needs to be understood in terms of causal closure. In fact the Completeness Thesis is often described in terms of causal closure, especially by those not favouring the use of the term 'complete' here (e.g. Kim 1998: 40)[5]. Ontological completeness will be a fact about the relation between certain types of events and laws, and for the same reasons as were just given in connection with **ECT1** this version of the Completeness Thesis does not rule out other causal ancestries for any given **X**-type event:

OCT1: A domain of reality **X** is complete iff **X** events always have sufficient **X**-event causal ancestries.

ECT1 and **OCT1** are both formulated in terms of events. To varying extents the major criticisms of the Completeness Thesis which are considered in the following three chapters take exception to aspects of what could, with conscious oversimplification for the time being, be called the 'Humean' analysis of causality and in particular to the notion that causation is best understood as a relation obtaining between events, typically conceived atomistically.[6] Rather than attempt to anticipate the results of those discussions, I will, for the purposes of the present chapter, take the above two formulations as paradigmatic, and modify them where necessary in the course of the specific discussions which follow, especially those about the nature of laws. For the time being, if we adopt the deliberately non-specific term 'outcomes' to refer to *whatever* events, effects, states, etc., might be the typical phenomena for a science or domain of reality, we could offer the following alternative formulations, which are not framed specifically in terms of events:

ECT2: A science **X** is complete iff **X** is capable of fixing the likelihood of **X**-type outcomes solely by reference to **X**-type entities, powers, states, laws, etc.

OCT2: A domain of reality **X** is complete iff **X** occurrences always have sufficient **X**-type causal ancestries.

---

5 Lewis (1983: 361) uses the term 'complete' to refer to physics in a way going beyond the present issue of causal closure. For Lewis a complete physics is a 'comprehensive theory of the world'. Whether a 'complete' physics in the sense the term is used in the present work would also be 'comprehensive' in Lewis's sense depends on the scope of 'physics' at the relevant world.

6 See especially Chapters Two, section 4 and Chapter Four, section 3. For Hume on causality see, e.g., his *Treatise* Book I, Part III and the first *Enquiry,* section VII (Hume 1978, 1975). Hume exegesis is not a critical factor here, and the view of causality I have in mind might just as well be called 'actualist', which is Bhaskar's (1978) term for positions which have causality as a relation holding between events, and laws as statements of event regularities, which in turn is similar to what Cartwright (1980, 1983) calls the 'facticity' view of laws.

Clearly, for us even to begin to ask whether any version of the Completeness Thesis is defensible, we need to do something about filling in **X**. Specifically, for present purposes, we need a way of doing so where **X** is some way of understanding *physical*. It is to this problem that I now turn.

3. Physics I

3.1. Properties

It will be helpful to adopt an idiom for marking the distinction between physical and non-physical for the purposes of the remaining discussions in the present chapter. I propose to describe the distinction as one between classes of properties. This approach is the one used by Daly (1998) in his critical treatment of attempts to give content to the distinction between physical and non-physical. I discuss Daly's arguments below in section (5) of the present chapter, so adopting the terminology from this point on will be doubly convenient.

Properties are ways an object is or might be. So we could think of the distinction between physical and non-physical as one which demarcates two classes of properties. Daly specifically demands a rationale for drawing a distinction within the assumed class of contingently existing properties which consistently divides the set of such properties into physical and non-physical sub-sets and which applies in all possible worlds (Daly 1998: 197-200). Let us assume, temporarily and for the purposes of argument, that there is indeed such a set of physical properties and call the set so defined ***PP***. This is by no means an easy concession to make,[7] but I make it with a view to seeing what else is wrong with Daly's arguments besides his beginning with this demand. While we're busy naming sets, let us also say that the set of physical properties at the *actual* world will be a sub-set of ***PP***, which we can call $\boldsymbol{PP}_{@}$.

Now, it is by no means self-evident that in order to have *any* debate over whether some property was or was not physical, or whether some proposition was true of physics at the actual world, we would be required to produce criteria for membership of ***PP***. There are, in fact, I suggest, a number of ways in which one might set about trying to distinguish the physical from the non-physical for the purposes of advancing some thesis concerning either category, or both. Most ambitiously of all one might attempt to set down principled criteria for membership of ***PP***, which is the only possibility Daly considers. For some purposes, though, such as those where what is at stake is a thesis about the *actual* world it could well be sufficient to give some definiteness to

---

[7] It is far from clear that what is a physical property at one world will be a physical property at any world in the sense which Daly seems to have in mind. Suppose two worlds, one in which Cartesian physics reigns and one where nature is Aristotelian. What considerations could we refer to in deciding whether to call a lump of Arisotelean earth matter 'physical' if it cropped up in the Cartesian world? Or what if some of Boyle's inter-particulate void popped up in a Cartesian world? Rather than argue that the entire notion of ***PP*** is suspect, though, I prefer to concentrate on the difficulties which arise when attempting to advance theses about the actual world. While this may seem somewhat parochial to those with heads for metaphysics, it is the actual world which is on centre stage here.

the boundaries of $PP_@$. Indeed, depending on the particular thesis in question it could well be sufficient to specify some suitable *sub-set* of the actual physical properties. (As we will see in a moment Smart advances a form of physicalism which asserts a relation between mental properties and *known* physical properties of the brain.) Alternatively one might attempt to characterise the physical in a way which gave criteria, which might be in part provisional, for membership of some set of which the actual physical properties would form a sub-set, but which is itself a sub-set of PP,[8] thus being in a position to consider at least some *alien* physical properties.[9]

So much for naming classes of properties. What can be said about task of fixing their boundaries? And just how much do we need to know in order to pursue a significant and interesting debate? It seems beyond doubt that the boundaries to all of these sets will have something to do with the findings of physics, but exactly what is no simple matter.

3.2. Allegations of Impossibility

There are two major difficulties facing any attempt to give a basis for a distinction between physical properties and non-physical properties. The first arises because there is little or no reason to think that present-day physics has discovered all there is to discover concerning physical properties. Consequently any attempt to develop a distinction along lines suggested by present physical knowledge seems doomed to be inadequate, and any claim about the scope or status of physics based on such a distinction seems equally doomed to be false.[10] A short look over the history of physics shows that this is, on inductive grounds at least, likely, since over the course of that history new entities and forces are added to physics from time to time, and others sometimes abandoned. This means that having identified *physical* with the set of properties admitted by physics at any given earlier time would have meant disqualifying subsequent discoveries from indicating the existence of formerly unknown physical properties.

On the other hand approaches which stipulate that physical properties are just those properties which would 'in principle' need to be included in ideal explanations or descriptions of processes presently recognised as physical, or in accounts of the composition of entities presently regarded as physical, seems to allow

---

8 If we label any sub-set of $PP_@$ which we might have reason to refer to PP_i (for example knowable actual physical properties) and any set intermediate between PP and $PP_@$ the criteria for membership of which we specify PP_j then we can summarise this set of alternatives by noting that $PP_i \subseteq PP_@ \subseteq PP_j \subseteq PP$. The alternatives noted here are not the only options which might lead to significant theses, since, for instance, there will be sub-sets of PP of which $PP_@$ is not itself a sub-set. Whether any of these might be useful depends on whether there is some particular debate in which they could play a role.

9 Lewis (1983: 364) explains that an alien property at some world is one which cannot be instantiated at that world, and cannot be *analysed* in terms of the natural properties of that world.

10 This point is frequently made, and is usually conceded by physicalists. See, for example, physicalists Poland (1994) and Papineau (1993), anti-physicalists Crane and Mellor (1990) and Stroud (1987). See also Putnam (1978).

whatever properties turn out play the required role to be counted as physical. This is hardly a way of discriminating between kinds of properties, and tends to make physics complete and perhaps even some forms of physicalism true by definition. This is, broadly, the danger noted by Chomsky in connection with the question whether mental phenomena could be given physical explanations:

> We can, however, be fairly sure that there will be a physical explanation for the phenomena in question, if they can be explained at all, for an uninteresting terminological reason, namely that the concept of "physical explanation" will no doubt be extended to incorporate whatever is discovered in this domain, exactly as it was extended to accommodate gravitational and electromagnetic force, massless particles, and numerous other entities and processes that would have offended the common sense of earlier generations (Chomsky 1972: 98).

So there are two dangers which threaten the project of saying what is to count as physical, which are those of making either physicalism or the completeness thesis obviously false or trivially true. This is only a dilemma given the implausible assumption that these are the only options, and there are also various ways of attempting to deal with the dangers just noted. Two examples will illustrate this point.

Smart, on the one hand, accepts that physical science is a work in progress, and that exactly which categories future physics will use to describe the world is presently unknown. In opposition to Chomsky's contention, though, he argues that 'for the purposes of biology and philosophy of mind we can tie "physicalism" to the principles of *present day* physics' (Smart 1978: 339).[11] Smart maintains that while there is no reason to think that the work of physics is over, there is also no significant reason to expect that physics will change in ways important for our understanding of the brain. I have no interest in defending Smart's physicalism here, the point is simply that his move regarding what is to count as 'physical' is one way of making precise what is at stake in particular debates over the relationship between physical properties and others, in this case properties peculiar to living and thinking entities, while leaving some of the content of the distinction between physical properties and non-physical ones open. That Smart's version of physicalism could turn out to be false is hardly an objection to the point that I am trying to make, which is simply that he successfully gives significant content to the notion of being 'physical' for the purposes of advancing a specific thesis. Smart's approach bites the bullet of relating physical properties to present physics, but does not make the thesis of physicalism trivially false by making the connection in a qualified way. It may in fact be false that current physics is unable to describe and explain brain processes, but it is not trivially false in the way asserting that current physics is equal to all descriptive and explanatory tasks would be. Effectively Smart attempts to deal with the problem at hand by giving criteria for membership of some sub-set of $\boldsymbol{PP}_{@}$, but a sub-set which is significantly related to the thesis he is advancing.

---

[11] See also Smart (1989: 80) where he asserts that he 'shall not deny that there will be revolutionary changes in physics. Nevertheless ... [they] ... will have no impact on biology and psychology.'

In contrast Papineau (1993) can be seen to take another tack, biting the other bullet so to speak, and defining physics in terms of *whatever* categories might be needed to explain some group of paradigmatic effects taken as physical, thus making physics trivially complete. A crucial characteristic of his approach is that he further stipulates that no psychological categories are to count as physical (1993: 30). Just as Smart tied *physical* to present physics without making his claims trivially false, Papineau grants physics a great deal of open-endedness without making his physicalism trivially true, closing off Chomsky's line of objection in another way. Papineau takes the claim that physics is complete *and* free from psychology to be an empirical one, and in the event that this combined claim is true his criteria for *physical* would end up delimiting something like $\boldsymbol{PP}_{@}$, or, at least, that sub-set of $\boldsymbol{PP}_{@}$ containing those properties which are in fact discoverable in the course of scientific investigation.[12]

I return to Papineau's proposal in more detail below. For the time being I merely note, having considered both Smart and Papineau, that there manifestly are strategies for delimiting *physical* in such a way as to enable significant and discriminating theses to be advanced. Neither proposal is *prima facie* wrong-headed, and these approaches are not the only ones available, so the would be physicalist or defender of completeness is simply not confronted with an unavoidable 'triviality or falsehood' dilemma. A notable characteristic of both Smart's and Papineau's approaches is that neither seeks to lay down general criteria for membership of ***PP***. Since both appear to succeed in stating non-trivial theses, the burden of argument that they *need* to do so rests elsewhere, with Daly and those who take a similar line.

My point here is not, simply by producing two examples of approaches which seem to succeed at saying how 'physical' should be understood for the purposes of advancing a specific thesis, to claim that all allegations to the effect that it is not possible to pin down the notion of being 'physical' have been vanquished. That would be ridiculous. Rather the purpose of this brief discussion has been twofold: to note the danger, and to suggest that it is not as alarming as it might at first seem. I take it only that I have done enough to justify not getting bogged down in an attempt to argue for the general claim that it is possible to draw the kind of distinction which is needed, in favour of getting directly to business with specific proposals. In what follows I discuss a range of attempts to formulate the physical/non-physical distinction, and also consider some critical challenges in more detail. On towards 'physics' then.

---

[12] It could be pointed out at this point that in the event there were biological categories which were needed to account for some of the paradigm physical effects, Papineau's approach would call 'physics' a science comprising both physics as conventionally understood *and* what we would call biology. Quite so. For a discussion of why this feature need not provoke any objections see section (6) below. Daly's criticisms of Papineau's proposal are discussed in section (5) below.

3.3. Garden Physics

Before getting into the attempt to specify the appropriate way to understand 'physics', though, it is worth noting that on even the most everyday account a very large portion, plausibly *all*, of the things we would recognise as entities, effects or outcomes are either physical or have a significant physical component. All material entities and processes on any scale are naturally thought of as at least partly physical, if only to the extent that aspects of them can be quantified in terms of physical magnitudes. It is partly for this reason that there is a great deal of *prima facie* plausibility to some form of physicalist supervenience thesis, as the following remarks by Kim suggest:

> To relieve a headache, you take aspirin: that is, you causally intervene in the brain process on which the headache supervenes. That's the only way we can do anything about our headaches. To make your painting more beautiful, more expressive, or more dramatic, you must do physical work on the painting and thereby alter the physical supervenience base of the aesthetic properties you want to improve. There is no direct way of making your painting more beautiful or less beautiful; you must change it physically if you want to change it aesthetically—there is no other way (1998: 42-3).

Given this, the claim that physics is, or could be, complete is both significant and interesting, if for no other reason than that the apparent consequences of such a claim would, if true, extend so far. If physics *is* complete there seems to be very little in the world for which no consequences would follow from that fact.

This rough and ready or intuitive sense that physics is 'all over the place' won't help much for the purposes of the arguments which follow. To make any progress we need to get down to business and try to fix 'physical' more precisely. This, though, brings us up against the very difficulties noted in the preceding sub-section. There is no point advancing a thesis which is either obviously false or trivially true, and a thesis taking either form would simply have failed to get at what is interesting and perplexing about the Completeness Thesis.

An answer to the question about the content of 'physics' will surely have something to do with the science we currently call physics. But that raises the first difficulty: If we tie the meaning of 'physics' to any specific details of present physical theory, it seems immediately unlikely that physics is complete. Present physics cannot yet provide physical explanations of all effects recognised as physical, and it is by no means obvious that physics forms a single science or even a network of compatible sciences.[13] In the terms developed above, either version of **ECT** is, on almost any account, *obviously* false where **X** is 'physical' as found in

---

[13] The specific question of the disunity of science, and the alleged heterogeneity of physics is discussed in later chapters, especially Chapters Three and Four.

current physical theories, and consequently any inference from the state of current physical knowledge to any version of **OCT** seems suspicious and implausible.[14]

While something along the lines of Smart's proposal, briefly discussed above, might seem like a lifeline, it won't work here. Tying 'physical' to some specific category of known physical properties might be a fair enough way of getting a mind-brain physicalism off the ground, but it's the Completeness Thesis, not any specific form of physicalism, which is at issue here. In a sense the perceived credibility of the Completeness Thesis can be seen as the source of the confidence physicalists like Smart have that the ongoing development of physics will not turn out to contain any nasty surprises for them. Looked at this way the present requirement is for an understanding of 'physical' which makes it possible significantly to ask whether this confidence is well founded.

It is as a result of the difficulties with associating 'physical' with current physics that, most of the time, when it is said that physics is complete, either the claim is meant ontologically, or, much the same thing, it is held that some future or ideal physics could, perhaps, or 'in principle', be suitably exhaustive in its specification of the antecedents of physical outcomes. But then in so abstracting away from current physical theory new dangers arise, including that of defining 'physics' as a science containing *whatever* is needed to establish the chances of physical effects (however they might be defined), and hence, especially given the noted scope of the physical *qua* effect, making what should be an empirical possibility look like it is being made true by definition.[15]

The challenge which arises here then, roughly put, is to specify the way 'physical' is to be understood so as to retain the capacity to discriminate significantly and appropriately, while still being sufficiently open-ended in the face of further scientific change and discovery. In what follows I consider a survey of existing approaches to the problem, discuss some critical positions, and offer an updated proposal specifically intended to facilitate critical discussion of the Completeness Thesis.

4. Physics II: Some Existing Approaches

The literature contains many attempts to specify how 'physical' should be understood, with respect to a variety of philosophical purposes. Poland (1994: 107-185) surveys many of them from the perspective of his attempt to develop an account of physicalism, and his detailed and fairly careful survey considers approaches and criticisms by, among others, Block (1980), Chomsky (1968), Hellman and Thompson (1975), Feigl

---

[14] Towards the end of Chapter Two I consider the question of the extent to which present incomplete knowledge might plausibly be taken to sustain belief in a causally closed physical domain. See especially section (7.1).

[15] As we saw briefly above, and will see in more detail below, proposals like Papineau's which make 'physics' trivially complete need not be unacceptable as long as there is some additional empirical question about what does and does not fall within the trivially complete physics.

(1969), Lewis (1983), Meehl and Sellars (1956), Post (1987), Putnam (1978) and Quine (1979). In the present section of this chapter I follow the main outlines of Poland's discussion.

4.1. Poland's Survey

Poland observes a distinction between what he calls *a priori* and *a posteriori* approaches to what he describes as 'isolating the physical bases', i.e. the task of saying what is to count as 'physical', where *a posteriori* approaches are those which rely upon 'appeals to the empirical study of nature … for grounds that distinguish the physical from the non-physical,' while *a priori* methods depend upon 'some form of conceptual or linguistic analysis or upon some other from of *a priori* argument purporting to provide definitive grounds for drawing the distinction between physical and non-physical…' (1994: 112).[16]

Poland begins by rejecting some views, which in his terms are *a priori*, (including Rorty 1979: 20) which characterise the physical by specifying that it is the *opposite* of the mental. He objects that these views fail to give any significant positive content to physics, and are also hostage to the need to specify what is meant by mental to get off the ground. While there is something in these reservations, it is also worth bearing in mind that physicalists themselves have been prepared to concede, with Feigl (1958: 420), that the notion of 'physical' appears to be *less* clear than that of 'mental', which would lend some support to views which attempt to use the mental as a starting point.[17]

It could also be objected, perhaps more seriously, that such approaches concede too much at the outset to a paradigmatically seventeenth century way of distinguishing the mental and physical by opposing them, exemplified in Descartes' account of unextended thinking substance, and unthinking extended substance. And to concede *that* is to concede too much to a broadly mechanist conception of matter as essentially inert,[18] at the very point where we are trying to ask what kind of thing physics is and what the physical is capable of.

Despite this, though, it is worth noting that from the perspective of the Completeness Thesis a negative characterisation may well turn out to play a useful role for the purposes of formulation and establishing what would count as evidence. As noted in the introduction to this thesis (section 1.1) Spurrett and Papineau (1999) argue that for the purposes of the standard type of argument for identity based on the rejection of

---

16 Poland's choice of the terms *a priori* and *a posteriori* to divide the approaches he considers is somewhat idiosyncratic, and it is far from clear that the approaches which he calls *a priori* are best thought as *a priori* in the typical contemporary senses of the term, even though they admittedly do not rely on features of current physical theories. Nothing important for my purposes hangs on the question of what to call the approaches, though, so for simplicity I will present them broadly in Poland's terms.

17 Note that starting off by *opposing* the mental and the physical is a considerably stronger move than the typical physicalist admission that 'mental' can be picked out and distinguished in virtue of attributes which are at least not *prima facie* physical.

18 The view persisted well beyond the seventeenth century. Thus Kant: 'And yet we cannot even think of living matter as possible. (The concept of it involves a contradiction, since the essential character of matter is lifelessness, *inertia*.)' (1987: 394).

overdetermination, what is needed to get the argument going is *any* plausible completeness premise. They then consider two versions of the overdetermination argument, one involving a negatively characterised completeness premise concerning the completeness of the non-mental.[19] I return to this point in section (6) below.

Turning for a while to *a posteriori* strategies, Poland (1994: 113-4) considers a well known proposal by Meehl and Sellars (1956) who distinguish between 'physical$_1$' and 'physical$_2$' as follows:[20]

> physical$_1$: terms employed in a coherent and adequate descriptive, explanatory account of the spatio-temporal order.
>
> physical$_2$: terms used in the formulation of principles which suffice in principle for the explanation and prediction of inorganic processes (1956: 252).

Although it might seem as though *physical$_1$* begs the question against, *inter alia,* the dualistic interactionist and emergentist, this is not actually the case. The purpose of *physical$_1$* is not to pre-empt such metaphysical considerations, but rather to capture the idea of a comprehensive and broadly scientific system of knowledge about the world. Both sides of Descartes' substantial dualism would comfortably find places in *physical$_1$*, as long as the purely temporal could be regarded as part of the 'spatio-temporal order'. Poland has it that *physical$_1$* 'was intended to capture the full vocabulary of natural science' and objects against it that the form of physicalism it leads to is 'much too weak' (1994: 114). Both claims seem questionable, though. There is nothing about *physical$_1$* that rules out the social sciences, economics, or psychology. Furthermore, it is far from clear that the purpose of *physical$_1$* was ever to characterise the physical *per se* (why else is there a separate account of *physical$_2$*?) but rather to set a general limit on what would be admissible as *science*. Certainly no interesting physicalist conclusions could be expected to follow from arguments concerning the relatively undiscriminating *physical$_1$*, for the reason noted more than once already, that physicalism is most naturally seen as a thesis about at least *two* classes of properties. Although Poland seems to expect *physical$_1$* to do more work than Meehl and Sellars did, my purposes also require a more narrow or restrictive characterisation of physical.

Indeed, Poland is, surely, correct to contend that '[t]he strength and significance of physicalist doctrine is, *ceteris paribus*, inversely proportional to the extent of the bases' (1994: 114). And it is precisely because *physical$_2$* is considerably narrower that it seems more significant and perhaps also promising for my purposes. Poland has a number of objections to it too. He points out, for example, that it lacks an 'independent

---

19 One of these is based on the proposal that the 'non-mental' might be complete, and the other that the scientifically 'quantifiable' might be. In both cases given the completeness premise an identity argument can be constructed.

20 Feigl (1958) adopts the Meehl and Sellars (1956) approach with some minor modifications, and continues to use the concept of the 'inorganic' to characterise his version of *physical$_2$*. He renders *physical$_2$* as '... the kind of theoretical concepts (and statements) which are sufficient for the *explanation*, i.e. the deductive or probabilistic derivation, of the observation statements regarding the inorganic (lifeless) domain of nature' (1958: 424).

characterisation of what 'inorganic' means' and that by implication it 'lumps together physics and chemistry' (1994: 115). Neither objection need be particularly damning, though, since while the exact boundary between organic and inorganic may well be vague, there manifestly is a great deal of unforced and untutored agreement about which things to count as living and which to exclude. Given this, 'inorganic' is a pretty useful way of discriminating a domain of phenomena, and could quite plausibly be regarded as a rather general natural kind. Not only that, there does not seem to be anything especially disastrous about combining physics and chemistry, especially not for Poland's purposes as a physicalist.[21] The exact relation between physics and chemistry is surely an empirical question, and the notion that chemistry is properly seen as a part of physics commands significant support.[22]

Poland also objects, following Chomsky (1968: 83) and Block (1980) that an account of physics based on *physical*$_2$ cannot handle the possibility of some kinds of emergent physical laws, specifically ones which come into play only in certain organic contexts, such as brains of certain sizes or degrees of complexity. Block puts the objection as follows:

> Briefly, it is conceivable that there are physical laws that 'come into play' in brains of a certain size and complexity, but that nonetheless these laws are translatable into physical language, and that, so translated, they are clearly physical laws (although irreducible to other physical laws). Arguably, in this situation, physicalism could be true—though not according to the account just mentioned (i.e. *physical*$_2$) of physical property (Block 1980: 296, quoted in Poland 1994: 116).[23]

This suggestion should not be mistaken with the more standard notion of an 'emergent' law where the law which emerges from a physical configuration is generally not *itself* regarded as physical.[24] Rather the key idea is of genuinely *physical* laws only manifest in some organic, or perhaps otherwise specialised,[25] contexts, a type of emergence which Poland describes as 'innocuous' as far as physicalism goes (1994: 117).

There are at least two issues here, and Poland's account suggests that he may be confounding them to some extent. On the one hand there is the issue of emergent physical laws, and on the other the specific notion of emergent laws associated with, say, organic systems. Obviously an emergent law which still fell properly

---

21 Chemistry has been the site of a recent emergentist revival, but since Poland makes no reference to this phenomenon his protest seems somewhat forced, especially since nothing in his wider argument indicates that he is especially committed to physicalism about *chemistry*. The question of emergence in chemistry is discussed below in Chapter Two, especially sections (2.1) and (7.1).

22 See Oppenheim and Putnam (1958: 22, 27), Smart (1989: 82), McLaughlin (1992) and Chapter Two below.

23 Block's argument is offered in an endnote where he is considering the view that a physical property is a 'property expressed by a predicate in some true theory adequate for the explanation and prediction of the phenomena of nonliving matter'. Compare McLaughlin on the scope of physics (1992: 53).

24 Emergence is discussed in more detail in Chapter Two below, and Block's proposal is considered specifically in section (3.2) of that chapter.

25 The question of the status of physical relationships, including laws, which may be limited to specific specialised conditions is discussed in more detail in Chapter Four below.

within the domain of the inorganic would present no difficulty for *physical*$_2$. Such a law could reasonably be expected to manifest itself in suitable inorganic contexts, and could also reasonably be expected to be the kind of thing physicists could find out about. Furthermore, it is at least unclear why we would want to maintain that a law obtaining *only* in organic systems, or some sub-set of them such as mammalian brains or the frontal lobes of transformational grammarians, would be properly thought of as physical. Block's stipulation that the law in question be translatable into the language of physics is insufficiently specific. Two ways in which it could be on the wrong track are, firstly, if he means either that the law be somehow *added* to physics so that the 'translation' would be a trivial achievement, or, secondly, if he is begging the question against the view that such a law would not be physical at all. Indeed our putative nomological oddity would seem to be a prime candidate for being recognised as a psychological or biological law, and that is exactly how most of the traditional emergentists would have seen it too.[26] For the time being I will treat Block's objection as not being damaging to *physical*$_2$ as an account of physical, and return to the problem in the course of the discussion of emergence in Chapter Two below.[27]

Poland (1994: 117) is surely correct, then, to shift some of the burden here onto the Block-style objectors, and demand that they offer some characterisation of 'physical' which makes sense of the objection, and also that they give some reason to take the speculation seriously. Poland also, though, decides that *physical*$_2$ has been found unacceptable, largely as a consequence of Block's worry, and continues with his survey, turning to the proposal that the meaning of 'physics' can be left to the physicists, or in some way read off the practices and conclusions of physicists (Poland 1994: 118). Nevertheless, he correctly points out that this fails to solve the problem for a range of reasons. Of the reasons he cites, the two most important for my purposes are, firstly, that such a proposal is of no direct use, since we still need to know how significantly to tell who the 'physicists' are, and secondly that unless we mean current physics we have not made any advance either. As noted above, if we *do* mean current physics, then the claim that physics is complete is almost obviously false, a point Poland also recognises (1994: 119).[28]

Despite this, there are those (including Feigl 1969, Lewis 1983, Putnam 1983) who hold that 'the physical is defined by modest extensions of current physics that are similar to it' (Poland 1994: 119). Poland contends that these approaches do not really deal with the problem here, for a range of reasons including their failure to specify the ways in which modestly extended physics will be permitted to vary. Poland concludes

---

[26] See Chapter Two below, especially sections (2) and (3). Note that Feigl (1958: 377) accepts physical$_2$ if 'emergentism is *not* required for the phenomena of organic life', and inclines to the view that physical$_2$ will hold the field against emergence.

[27] In section (6) below I adopt an account of physical which is very close to *physical*$_2$.

[28] It is worth noting, though, that if we take 'physical' as something like a natural kind, then 'physicists' are the people who try to figure out what the extension of the kind is. See the discussion of Snowdon in section (5.2) below.

that approaches tied to current physics even in this qualified way are unacceptable, and continues to seek a more general, and in his terms *a priori,* account of 'physical'. I would argue, though, that Poland's rather swift and impatient rejection of the views in question obscures some important points that are being made. Putnam puts it as follows:

> I shall assume that the fundamental magnitudes are basically the usual ones: if no restraint at all is placed on what counts as a possible 'fundamental magnitude' in future physics, then *reference* or *soul* or *Good* could even be 'fundamental magnitudes' in future physics! I shall not allow that naturalist the escape hatch of letting 'future physics' mean we-know-not-what. Physicalism is only intelligible if 'future physics' is supposed to resemble what we call 'physics' (Putnam 1983: 212).

Putnam is surely correct to maintain that the handing out of blank cheques to physicists by proposing unconditionally to be physicalist about *whatever* they eventually decide is philosophically poor form. His point has much in common with Chomsky's (1968) warning, briefly discussed in section (3.2) above, that without some kind of qualification or restriction the category of 'physical explanation' is unlikely to be able to do any discriminating work. That point has already been noted and discussed, though. What I want to emphasise now is that some kind of attention to present, and even previous, physics is unavoidable when we attempt to consider what it would be reasonable to expect from future or ideal physics. Whatever we can reasonably expect from a future or ideal physics has to have *something* to do with the state of present physics. In Chapter Four (section 6) below I return to the question of how best to make use of the limited evidence available from physics in its current state.

Not only that, Poland simply seems to have misunderstood Lewis. In the text to which Poland refers, Lewis characterises what he calls Materialism as 'the thesis that physics – something not too different from present-day physics, though presumably somewhat improved – is a comprehensive theory of the world, complete as well as correct' (1983: 361). As always it is important to keep one's modal ducks in a row when reading Lewis, whose point here is to assert something about the *actual* world, rather than to offer a general definition of physics which would pick out some specific set of properties at any world. Lewis's view is that at any world there will be some minimal set of properties, which he calls the natural properties, underlying the various processes and changes which take place at that world. He also maintains fallibly and empirically that at the actual world the set of natural properties is pretty close to the basic properties of current physics.

Poland concludes his survey of *a posteriori* approaches by claiming that the need for a broadly *a priori* distinguishing criteria for distinguishing 'physical' seems more urgent than ever, and stating that '[s]trategically, the problem for the physicalist is that, unless there is some antecedently specifiable principle for identifying the physical bases, physicalism cannot be formulated in significant ways' (Poland 1994: 119).

Poland's first candidates for *a priori* strategies are similar proposals by Lewis and Quine, to the effect that physics can be regarded as 'the branch of science whose goal it is to provide a comprehensive supervenience base' (1994: 120). Thus:

> One motivation of physics down through the centuries might be said to have been [...]: to say what counts as a physical difference, a physical trait, a physical state. The question can be put more explicitly thus: what minimum catalogue of states would be sufficient to justify us in saying that there is no change without a change in positions or states? (Quine 1979: 163-4).

> Physics (ignoring latter-day failures of nerve) is the science that aspires to comprehensiveness, and particular physical theories may be put forward as fulfilling that aspiration. If so, we must again ask what it means to claim comprehensiveness. And again, the answer may be given by a supervenience formulation: no difference without physical difference as conceived by such and such grand theory (Lewis 1983: 356-7).

Prima facie this seems as though it can't be quite right. Poland (1994: 120) has it, getting at part of what wrong here, that 'Quine and Lewis appear to have confused the aims of physics with the aims of physicalism', in other words that it seems mistaken to impute to *physicists* any particular interest in philosophical positions such as supervenience. Physicists might plausibly want to be eliminativists, for example. If we consider the worlds at which some physics is in fact complete, then a consequence of the completion of the work of physics would be that a supervenience base for other phenomena had been developed only at *some* of those worlds. But the supervenience base at some world (if there is one) need not be purely physical, so being part of a supervenience base cannot be a criterion for being physical. Not only that, the Completeness Thesis on its own does not entail what metaphysical dispensation will obtain at any given world at which it is true, since completeness is compatible with, *inter alia*, zombie worlds, qualia rich worlds, preëstablished harmony. If physicalists are to be correct then science must in principle be able to do what Quine and Lewis suggest, but it cannot be a defining characteristic of physics that it seek to do this, or that it can succeed.

For all that, it is worth noting that Poland seems, again, to be taking Lewis to be saying something different from what he is actually saying. Again, Lewis's point is one about the actual world. In the event that something like his Materialism is true, and given his further arguments for the identity thesis (Lewis 1966), the passage quoted above would be a reasonable description of the activity of physics. Not only that, the key notion in Lewis's account is that of 'comprehensiveness', and the talk of supervenience is a way of formulating that notion.

It is not that surprising that Poland has complaints regarding the Quine and Lewis strategies, though, since he has his own proposal regarding how 'physical' should be understood. It is to that which I now turn.

4.2. Poland's 'Physics'

Poland takes the failings and pitfalls identified in the course of his survey as an opportunity to develop his own set of criteria for a characterisation of physics, one largely connected with his desire for a formulation of physicalism. He sets down ten criteria, of which the most notable for my purposes are his requirement that the account employ 'no question-begging terms (for example 'inorganic', 'physicist')', that no 'inappropriate *a priori* constraints' be imposed on the form or content of physical theory, that the formulation be compatible with emergent physical laws, and that the characterisation not be so abstract that it makes it impossible to distinguish physics from other disciplines (1994: 123). All well and good. Poland summarises his account of physics as follows:

> ...physics is the branch of science concerned with identifying a basic class of objects and attributes and a class of principles that are sufficient for an account of space-time and of the composition, dynamics and interactions of all occupants of space-time (1994: 124).

This could use some explanation, since this version of Poland's position is, *prima facie*, not obviously distinguishable from the Meehl and Sellars (1956) account of *physical*$_1$ which Poland himself argued was inadequate. The difference here has to do with how Poland wants us to think of the *classes* in terms of which physics is to be understood:

> The crucial features of these classes are that they are minimal with respect to the descriptive and explanatory purposes they serve, that the magnitudes are defined for all regions of space-time, and that each occupant of space-time satisfies the principles governing those magnitudes. It is both the types of phenomena they are introduced to explain (i.e. composition, dynamics, interactions) and their complete generality that distinguishes these magnitudes and principles from others, and hence that distinguishes physics from other branches of enquiry (1994: 124).

Poland argues that these features of his account prevent the identification of physics with science in general, which would indeed distinguish his position from *physical*$_1$. Even though any science will typically refer to classes of objects and attributes which exhibit various types of composition, dynamics and interactions, such classes will not be even partly physical by Poland's lights unless they overlap with the minimal classes which he has in mind. So physicists are not interested in social phenomena '*as* social phenomena, although they are [interested in them] as occupants of space-time (i.e. in so far as they involve causal processes or entities which 'take up space')' (1994: 125).[29]

---

[29] Poland's remarks here echo those of Oppenheim and Putnam (1958: 10) who say that 'a physicist, when he speaks of "all physical objects," is also speaking about living things—but not qua living things.' Oppenheim and Putnam's proposals regarding the Unity of Science are discussed in Chapter Three, section 2.1 below, followed by an account of Dupré's (1993) criticisms of unity and completeness and responses to them.

This means that Poland does think that physics is a universal science, and he proposes that this universality needs to be understood by reference to the kinds of questions physicists ask, or which physical theories answer:

- What are the fundamental constituents of all occupants of space-time?
- What are the fundamental processes that underlie all causation and all interactions between such occupants?
- What parameters are relevant to the dynamic unfolding of all systems in space-time and hence to all change?
- What is the nature of space-time itself, its origin (if it has one), and its destiny? (1994: 125).

Poland thinks that his proposal is immune to the difficulties he perceived with the Quine and Lewis supervenience-based accounts, consistent with the possibility of emergent physical laws along the lines suggested by Block, and that it guarantees 'a relatively narrow conception of the physical bases' (1994: 126) while being consistent with the growth and development of physical theory, since it is not tied to any particular physical theories.

Certainly from my point of view it can be admitted that if physics is ontologically complete then an ideal physics along Poland's lines would be epistemologically complete in the required sense. But this admission is a truism, which signally fails to tell us what 'physics' actually is. Put another way, it is a feature of Poland's proposal (one which he simply fails to notice) that it is open to a form of the Chomsky (1968) objection noted above. As we saw, Chomsky pointed out that the notion of a 'physical explanation' could be extended to embrace whatever we decided was in need of explanation, and hence that the claim that something could be explained physically lost much of its force. Poland's stipulations that his classes be minimal prevents any shameless *ad hoc* extensions of 'physical' and hence cuts off some of the force of Chomsky's criticism. But not all. In the event that psychological, or biological, categories, say, were part of the minimal classes in Poland's sense then they would be part of what he calls physics. So, if Elsasser was correct to view the laws of physics roughly as special cases of the laws of biology,[30] Poland would make biology part of physics. Alternatively, both of the substances in Descartes' dualism would be part of the physical on Poland's account.[31] Whatever the merits of Poland's proposal, then, it manifestly cannot do the work that he himself demanded of it, which is to play a role in a general physicalist programme, including physicalism about the mental.

---

30 Elsasser (1958) more strictly makes both biological and physical laws special cases of some more general nomological framework. See Chapter Two section (2).

31 It is also likely that difficulties would arise for Poland in the event that there were fundamental attributes which were not shared by all the basic constituents of things, for example if not all fundamental particles had nonzero masses. Since it is not my project to *repair* his proposal, though, I will not get into this.

Rather than attempt to repair the defect, though, I now turn to the consideration of some arguments to the effect that 'physical' *cannot* be given a satisfactory characterisation. I argue that these criticisms can be answered, and that the weaknesses of the criticisms can help point the way to a solution sufficient for my purposes.

5. Daly's Criticisms

Daly (1998, see also 1995) has recently argued that a number of existing attempts to specify how 'physical' should be understood are defective in various ways, and come to the conclusions that 'there is no principled and well-defined distinction between physical properties and all other properties',[32] and, further, that programmes in metaphysics assuming such a distinction should be 'abandoned' (1998: 196). If Daly is correct, a number of apparently interesting and significant debates are in serious trouble.

In what follows I consider a number of the proposals discussed by Daly, defending them from his criticisms and arguing that they do permit significant theses to be advanced. Following this I offer some general suggestions concerning ways of delimiting sets of properties for the purposes of the kinds of debate Daly says should be abandoned. The one line of thinking aimed at giving criteria for membership of ***PP*** which Daly attacks but which I will not defend is that based on family resemblances (Daly 1998: 200-2).[33] As Daly correctly points out such an approach is likely to fare badly in the face of both physical properties not connected by suitable resemblances with other physical properties, and non-physical properties which are so connected.

Recall the discussion of physical properties in section (3.1) above, where I suggested that depending on the particular argument which was at stake, any of a number of sub-sets of ***PP*** would form a suitable account of 'physical.' Although Daly (1998: 197-99) maintains that discussion relating to such sub-sets of ***PP*** are hostage to the need for general criteria for being 'physical' it is by no means clear that this is the case. *Without* knowing what would decide whether *any* property was physical or not, it is clearly possible to make significant statements about, say, the difference between worlds where one large scale structure of space-time prevails, and worlds where another does. In general I argue below that as long as we keep clear about what it is that any given argument about 'physical' is intended to establish then we can see whether or not the account of 'physical' in question will do the required job or not. Exactly what and how much is needed will vary from case to case, but it is the argument which should determine this, not some blanket policy concerning what it is that we need to do or know in order to call something 'physical'.

---

[32] Recall the notion of the class of all contingently existing physical properties, or ***PP***, as introduced in section 3.1 above.

[33] Daly attributes this view to Frank Jackson in his John Locke lectures at Oxford University in 1995, noting in turn that Jackson acknowledged the suggestion to be Papineau's.

5.1. Natural Kinds

One strategy for characterising the physical might be to say that 'physical' is a natural kind, i.e. that there is some real difference (involving an essential feature or set of such features) between physical things, properties and relations, and non-physical ones, which we can find out about in the course of practical and theoretical scientific investigation into the natures of things satisfying the constraints on the kind. Snowdon (1989) makes a proposal along these lines, which is criticised by Daly.

Snowdon does not adopt a particular theory of natural kind terms, but Daly seems correct when he reads Snowdon's suggestion along the lines of a well known account of such terms, due to Kripke (1980) and Putnam (1973, 1975).[34] On this view natural kind terms pick out a real extension, where the extension in question typically takes the form of a physical microstructure, or at least some shared theoretical feature, and where suitable causal contact with these microstructures enables successful reference. More specifically, Putnam proposes that meaning can be defined on the analogy of a vector with four components: syntactic marker, semantic marker, stereotype and extension. Ordinary linguistic competence does not presuppose knowledge of the extension, for example H_2O, of the word 'water', but rather:

> It means that (*we* say) the extension of the term 'water' as *they* (the speakers in question) use it is *in fact* H_2O (Putnam 1975: 269).

Competent speakers do, though, according to Putnam, need to be masters of the markers (semantic and syntactic) and the stereotype, which in the case of 'water' might include 'colourless; transparent; tasteless; thirst-quenching; etc.' (1975: 269). On this view, then, there will be a real fact of the matter about the extension of a natural kind term which it may or may not be possible to discover, and which may or may not be known by any member of the community using the term.

I note in passing at this point that the Kripke-Putnam approach is not without opponents, and particularly that it has been vigorously criticised from the point of view of *biological* natural kinds by, for example, Dupré (1993). As far as biological kinds go I am in broad agreement with his objections, but maintain that in cases where the extension of the kind is determined by *composition* or *constitution* (rather than genealogy, function and so on) the Kripke-Putnam view is at its most plausible. It is also not required that we regard it as a generally acceptable theory of meaning for the purposes of the present argument. All that is needed is that we see it as a plausible model of how a scientifically motivated classification of a kind which really is compositional might work. Let us turn to Snowdon's proposal.

---

[34] The Kripke-Putnam account of natural kind terms is discussed at greater length in Chapter Three below (section 2.2) in connection with Dupré's criticisms of it.

Snowdon's general project in the paper to which Daly refers is an attempt to formulate materialism (or physicalism — he uses the terms interchangeably) and dualism. Snowdon contends, on this point agreeing with Stroud, that *a priori* approaches to explaining the predicate 'physical' are unlikely to succeed (Snowdon 1989: 152, Stroud 1987: 277). I return to Snowdon's reasons for this shortly, but for the time being note only that he takes the limitations of *a priori* approaches to indicate the need for an elucidation of 'physical' which is 'sensitive to the empirical exploration of the physical world' (1989: 152). Rejecting attempts to tie 'physical' to current physics, for reasons along the lines of those noted above in section (2), and also eschewing entirely undiscriminating gestures at physics '*as it ultimately evolves'* Snowdon instead proposes that:

> We apply the term 'physical' to a range of objects, for example, tables, chairs and stones, which we think we encounter in perception. The term is intended to mark out their most basic and shared essential features. That is, 'physical' is an extremely basic natural kind term, a term for the most all-embracing (natural) kind with which we are acquainted. The discovery (if such is possible) of the essence of the kind is a posteriori. So, on this suggestion, the term's restrictions flow in the same way that those of natural kind terms do in other cases (1989: 154).

This suggestion is rather compressed and short of detail, but the idea would seem to be that we can be plausibly be taken already to have a developed sense of what we can call the *stereotype* of 'physical',[35] which, along the lines of the Kripke-Putnam view, would fix the extension even if we do not know, and may indeed never know, the essence particular to the extension of the word. There is something fundamentally correct about this too, since we clearly can, and do, have a cluster of activities and programmes under the broad heading of 'physics' *without* knowing what or how much any of these programmes will eventually discover.[36] It is Daly's demand that we pin down ***PP*** before we even get started which at fault here.

For Snowdon's purposes this seems fair enough, since the two forms of dualism he goes on to identify are united in denying that the physical natural kind will turn out to form the foundation of all phenomena, either because the mental will be found to have its ground in something immaterial, or because mental phenomena 'are themselves amongst the fundamental features of the world' (1989: 156-58). In the event that physics was not complete in the sense explained above, because one of Snowdon's forms of dualism was correct, it would be the case that physics was unable, even in principle, to provide a causal ancestry for at least some physical occurrences. Whatever its other merits, Snowdon's approach does enable him to advance the theses which he is attempting to advance.

---

35 For more on stereotypes see Chapter Three section (2.2)

36 Note that on the Kripke-Putnam view we do not have to be especially perspicacious in our identification of the stereotype with which we begin our enquiries, since it is the ongoing investigation into the *extension* which carries the greater part of the burden here. So, for example, no-one had to *begin* by thinking that diamonds and coal were basically the same in some important sense to find out about carbon.

What of Daly's criticisms of Snowdon? For some reason Daly identifies as Snowdon's 'conclusion' (Daly 1998: 203) a remark to the effect that 'we can explain being physical as being the way (whatever it is) that all these objects (for example tables and chairs) fundamentally are, given that they are objects with the capacity to exist independent of perception and are in space (and so on)' (Snowdon 1989: 153). As it happens that particular remark is an attempt by Snowdon to forestall an imagined idealist objection by tactically modifying the formulation which I quote above. Nonetheless Daly objects to this 'conclusion' that many paradigmatically non-physical entities have been supposed to exist in the absence of perception, including 'Platonic Forms, Fregean Thoughts, God, and Cartesian souls' (1998: 204). The point is well taken, and bears comparison with a remark made by Snowdon one page before the sentence Daly lights upon. As noted above Snowdon is in agreement with Stroud regarding the implausibility of *a priori* approaches to the problem, and notes in this regard that:

> ...it is hard to be confident of even the most general notions, such as occupation of space, that, as understood, they *must* apply to all physical objects, given the surprising entities physicists are prepared to endorse. On the other hand, even if we could articulate such a highly general, and a priori defensible, property of all physical entities, it is unlikely that it could yield a sufficient condition for being physical. After all, Locke held that spiritual substances had a location, but were not physical (Snowdon 1989: 152).

It is abundantly clear that Snowdon's point regarding objects being in space, and independent of perception, is one about tables and chairs, directed at his hypothetical idealist opponent, rather than an attempt at laying down some strict *a priori* criterion for all and only physical objects, an attempt which, as we have seen, he gives good reasons for rejecting. In any event it is surely worth noting that Snowdon's point against the imagined idealist is, clearly, that the *stereotype* (or nominal essence) of 'physical' objects involves being in space, not that spatiality is a general condition for all physical properties, or that being in space is an *a priori* requirement for being part of the extension of 'physical'[37]. Since Snowdon does not actually make the claim Daly accuses him of, it would be possible to set aside Daly's attempt to show that the requirement that physical objects be in space is doomed to circularity.[38]

---

[37] That said, it is worth noting that Schlick (1953: 530f) contends that '*spatial* characteristics are the indispensable conditions for a concept of science' and notes that this view seemed to be shared by the neo-vitalist Driesch, whose entelechies, Schlick argues, were supposed to be non-spatial, and especially non-localised, even though having spatial effects.

[38] Daly's objection hinges on the claim that the attempt to require that physical properties be in space is likely to collapse into circularity in one of two ways. On the one hand, if space is defined as a system of relations between physical objects, then the natural kind *physical object* would be partly explicated in terms of the notion of physical space. But that notion of space would in turn be accounted for in terms of the fact of space being a system of relations between physical objects, which would be circular. If, on the other hand, space is taken to be a physical object itself then the account of the natural kind *physical object* would be in terms of a particular physical object, which would also be circular. But there need be no circularity in either case. A system of relations could be a set of physical properties dependent upon some *other* set of properties.

Following his general suggestion concerning 'physical' as a natural kind, Snowdon goes on to propose that a 'property is a physical property if it can be instantiated in a domain consisting only of physical objects' and that 'a state of affairs is a physical one if it involves or consists only of physical objects instantiating physical properties' (1989: 154). Daly objects that this would seem to require properties which are not readily thought of as by definition physical ones (such as biological and psychological properties) are to be made physical by definition in the event that it turns out that they can be instantiated in a domain consisting only of physical objects (1998: 205). He also complains that even if this is the kind of thing which might indeed turn out to be true, it is not the sort of thing which should be trivially true.

To the extent that it is reasonable to take Snowdon's proposal as an attempt to lay down general and pre-emptive criteria for being physical there might be something in this. Since, however, Snowdon's project was merely to state clearly what was at issue between dualism and materialism it is difficult to see where the problem is. Snowdon is surely correct to say that a dualist (as he describes the two main variants of that doctrine) would not accept that mental properties could be instantiated in a domain consisting only of physical objects. And if they *could* be, that would indeed make them physical by definition for the purposes of his discussion.

For all its brevity and minor flaws Snowdon's proposal is onto something important, which is that we can at least try to characterise 'physical' by linking the extension of the term to an ongoing project. He marks this point by saying, as noted, that we need a characterisation of 'physical' which 'allows its elucidation to be sensitive to the empirical exploration of the physical world' (1989: 152). What I want to do now is look at an approach which suggests that our ongoing investigation of the 'physical' can be driven by a focus on paradigmatic physical *effects* rather than physical *objects*.

5.2. Paradigm Effects

Daly also considers a proposal, mentioned above in section (3.2) and due to Papineau, to the effect that some 'pre-theoretically given class of paradigmatic physical effects, such as stones falling, the matter in our arms moving, and so on' can be postulated as independently given, and physics characterised in terms of 'all the categories that need to be brought in to explain those paradigmatic physical effects' (Papineau 1993: 30).[39]

Alternatively a collection of physical objects could have the property of being 'in' some other physical object which could itself be independently characterised.

[39] Daly (1998: 206) notes that Papineau's proposal is similar to Feigl's characterisation of 'physical$_1$' and 'physical$_2$' in Feigl (1958) p. 377. Feigl's final formulation of physical$_2$ has it as embracing 'the kind of theoretical concepts (and statements) which are sufficient for the *explanation*, i.e., the deductive or probabilistic derivation, of the observation statements regarding the inorganic (lifeless) domain of nature.' (1958) p. 424. Daly does not note that Feigl's version is itself a modification of a proposal by Meehl and Sellars (1956). See sections (4.1) above and (6) below.

Daly has three major objections to Papineau's approach: that it is too narrow, too broad and, for good measure, that it is circular.

The first criticism is that Papineau's criteria may suffice at best to 'pick out only the class of actual physical properties,' (1998: 206) hence being too narrow. There are two ways of reading this criticism. On the one hand it might be taken to suggest that Papineau's approach would only pick out the general physical properties of the actual world, so that the properties instantiable under some or other alien physics would not be included. On the other hand, Daly gives an example of '*being-a-solid-sphere-of-gold-with-a-diameter-of-one-mile*' (1998: 207) which indicates that he may mean that the set would be limited to the actually *instantiated* physical properties. In either case it is difficult to see that there is a significant problem. For a start there is good reason to think that delimiting the actual physical properties is all Papineau was attempting to do with his proposal, since his broad objective was to defend physicalism with respect to the *actual* world by means of an argument using the premise that physics is complete (Papineau 1993: 29-32). For such a project a procedure which specified the actual physical properties would be exactly what was needed. Indeed it is difficult to see how starting from actual physical effects and working backwards to the factors fixing their likelihoods one could go anywhere except to other actual properties, which, with Papineau, we could agree for the purposes of argument to call physical. The gold sphere example, on the other hand is clearly a red herring. Physics does not generally deal in particular properties of that sort, but can more plausibly be seen as dealing, or attempting to deal, with what Lewis (1983: 364) would call natural properties: that minimal catalogue of attributes in terms of which the descriptions of actually existing things can be analysed. On this view were a gigantic gold sphere to be discovered it could be perfectly well understood in terms of natural properties, but it is surely perverse to demand that it actually be *anticipated*. To demand that is akin to demanding of a theory of grammaticality that it come with a list of all the possible grammatical sentences. Really, it is Daly's claim that Papineau's approach is too narrow that is itself in need of justification.

Nevertheless Daly imagines that in order to meet this difficulty, if difficulty it is, Papineau might want to say that some property is a physical property if and only if there is at least one world in which that property is needed to explain an effect which is of the same type as one of Papineau's class of paradigmatic physical effects (1998: 207). This seems like a rather poor lifeline, though, since Daly does not explain why we should think that physical effects at other worlds will be suitably similar to those at ours for his modified procedure to capture what he or Papineau are after. We can agree, though, that even if Daly's proposal would not necessarily give us criteria for membership of ***PP***, it would point to some set of properties more extensive than $\boldsymbol{PP}_{@}$. But Daly then argues that his proposed modification of Papineau's position has the defect that it would make *any* property needed for explaining physical effects into a physical property, including for example mental properties at worlds where dualism was true, and hence where some paradigmatic physical effects had mental causes. Quite so. It is unclear why this is supposed to follow from modifying Papineau's approach, though, since, as

noted, that approach precisely *sets out* to make physics trivially complete, by characterising as 'physical' *anything* needed to fix the likelihoods of the paradigmatic physical effects. Physics can be trivially complete in this way without some or other form of physicalism being made trivially true only if further considerations are added. And, as already noted, Papineau further hypothesises, with precisely this issue in mind, that physics not make reference to psychological categories (1993: 30). This hypothesis is empirical, and it is important to note that it gives a contingent edge to the question of the completeness of physics. If physics is complete by definition, then it cannot *also* exclude psychological categories by definition, and *vice versa*. This is so, since at some worlds physical effects will indeed follow psychological antecedents in just the ways Daly suggests. A somewhat extended version of Papineau's approach is discussed in section 6 below, for now let us stay with Daly's criticisms.

Daly objects that Papineau's account of the physical properties of the actual world is too narrow in a second way. This apparently follows because the stipulation that physics contain only the categories that *need* to be brought in to explain the paradigmatic physical effects, has the effect that some physical properties are excluded:

> Every paradigmatic physical effect which can be explained by reference to *density* can be explained by reference to *mass* and *volume* instead. It follows from Papineau's definition that *density* is not a physical property (1998: 207).

There are at least two ways of responding to this criticism. On the one hand it might be possible to demand an explanation of how *mass* and *volume* could possibly be referred 'instead' of density. Density, after all, just is a function of mass and volume, so it is just not clear at all what Daly means by the suggestion that an explanation which simply referred to mass per unit volume would not involve reference to what we call density. At most it seems that he can say we could use substantially the same *description* or *explanation*, but not bother to assign a *name* to the quantity resulting from measuring mass per unit volume. Perhaps, though, the problem here is with Daly's example.

An alternative response, then, would be to accept the criticism and then ask what damage it actually does. For this purpose let us set aside the mass, volume, density example and imagine *any* property or set of properties which, for whatever reason, are not needed, in Papineau's sense of needed, for the explanation of paradigm physical effects, or the effects of which the paradigms are paradigmatic. We can also imagine for good measure that these properties are ones that we would be strongly inclined, for whatever reason, to call physical. Recall also that Papineau's account of 'physical' is intended to contribute to establishing a completeness premise which will play a role in further arguments for a particular form of physicalism. Daly is worried that the specific focus on only those properties which are *needed* to explain the class of physical effects will result in picking out some set of physical properties itself forming a sub-set of actual physical properties. So what? For

physics to be complete, according to Papineau, it must be the case that 'all physical events are determined, or have their chances determined, by prior *physical* events according to *physical* laws. In other words we need never look beyond the realm of the physical in order to identify a set of antecedents which fixes the chances of every physical occurrence' (1993: 16). And if there is some set of physical properties which contains all the physical properties needed to set the likelihood of all physical outcomes, then if nothing else, physics is complete in the required sense. Put another way, if we never have to 'look beyond' some sub-set of the actual physical properties, then, trivially, we will never have to look beyond either $\boldsymbol{PP}_{@}$ or ***PP*** either. If we bear in mind the particular point of Papineau's argument, it is clear that there just is no requirement for an exhaustive specification of criteria for membership of either ***PP*** or $\boldsymbol{PP}_{@}$ here. Again, it is the claim that one *is* needed which stands in need of justification.

Daly's third and final objection is that we cannot decide *which* effects to class with the paradigmatically physical effects needed to get Papineau's project off the ground unless we already have some criterion of being physical:

> And since Papineau's account was meant to provide a characterisation of what a physical property is, his account proves circular (1998: 208).

This misses the point of Papineau's approach entirely. Simply to characterise physical properties *as* physical properties would indeed be as useless and circular as saying all members of ***PP*** are defined *as* members of ***PP***. It is not the same thing, though, to accept that some properties are already known or commonly taken to be physical[40] and to characterise further properties of the same kind by specifying in a discriminating way what role those properties will play in a certain type of explanation would not. What we call 'physics' is the science which, *inter alia*, attempts to provide that kind of explanation, and at any given world it will be a matter of fact how much can be so explained, and hence a matter of fact whether or not 'physics' is complete (i.e. whether some 'physical' effects have non-'physical' causes), whether 'physical' so defined does or does or does not include psychological categories, and so on. In section (6) below I defend a modified version of Papineau's approach which I contend is immune to all objections of the type Daly advances. First, though, I want to look at a one more proposal for delimiting the 'physical', which begins by characterising the discipline and activity of physics.

---

[40] As Daly himself does when giving '*being-5kg-in-mass, being-5-light-years-from-Neptune,* and *being-a-magnetic-field*' as examples of physical properties, and then also giving some examples of non-physical properties, while attempting to explain what is at stake in delimiting *PP*, Daly (1998: 198). Strategies such as Snowdon's and Papineau's demand no more from our everyday sense of 'physical' than Daly is prepared to allow for the purposes of formulating many of his own arguments.

5.3. Physics and the Physical

As we have seen Poland (1994) has recently attempted to address the problem of characterising *physical*. He does so, after canvassing and rejecting a range of options, by developing an account of physics as a science, including an explanation of what he takes to be the crucial universality of physics. While Poland rejects suggestions along the lines of Quine and Lewis to the effect that the aim of physics is to provide a supervenience base for other forms of description and explanation, his own account is best understood with supervenience in mind. Poland's complaint against Quine and Lewis is that they confuse the aims of physicists with those of physicalists, and as we saw above he himself eventually characterises physics as follows:

> ...physics is the branch of science concerned with identifying a basic class of objects and attributes and a class of principles that are sufficient for an account of space-time and of the composition, dynamics and interactions of all occupants of space-time (1994: 124).

Daly objects that Poland's approach is insufficiently discriminating, and contends (in a manner reminiscent of his criticism of Snowdon) that Poland has no criteria which can justify the exclusion of psychology and economics from being parts of physics, since both make reference to the composition and interaction of occupants of space time (such as human beings and banks) and construct explanations of their dynamics (Daly 1998: 210-11). He notes Poland's stipulation that 'all dynamics and all interactions among phenomena are constrained by fundamental physical principles' (Poland 1994: 125, quoted in Daly 1998: 211) but maintains that this does not stand in the way of regarding the laws of biology, chemistry, psychology, economics, etc. as part of physics, since Poland's criterion of 'generality' is not properly explained. The two sentences in Poland immediately following the one Daly quotes are worth spending a little time on:[41]

> But this does not mean that physics is concerned with complex phenomena with respect to *all* of their properties. For example, social phenomena are not of interest to the physicist *as* social phenomena, although they are of interest as occupants of space-time (i.e. in so far as they involve causal processes or entities which 'take up space'). (1994: 125).

With a little unpacking this both enables Daly's criticism to be answered, and shows some significant connections between Poland's approach and Papineau's. A little more detail from Poland will be helpful here. Poland contends that it is 'not incorrect to say that physics is a universal science', (1994: 125) a remark reminiscent of Snowdon's characterisation of 'physical' as 'a term for the most all-embracing (natural) kind with which we are acquainted' (1989: 153). As we have seen, Poland suggests that this universality can be understood by reference to general features of the questions physical theories attempt to answer.

---

[41] It is not clear why the sentences in question are neither quoted nor referred to by Daly.

So would Poland really be required to grant psychology, say, a position within physics? The answer is that he would not in the sense that Daly suggests, but that there are worlds in which he would, just as there are worlds in which Papineau's trivially complete physics would include psychological categories. Daly is surely correct to point out that psychology includes dynamics and interactions, and that it involves consideration of the composition of the entities it studies, which are themselves occupants of space-time. But what Poland says on the matter gives a principled and non-ad hoc way of saying that psychology (or any other special science) is not *automatically* a part of physics. The two crucial points here are Poland's account of the generality of physics, and his stipulation that physics concern a class of attributes and principles which are *sufficient* for the purposes he details.

Some of the occupants of space-time are indeed social phenomena. But what does Poland mean by saying that physics is not interested in them *as* social phenomena? Let us first consider a case along the lines most clearly suggested by Daly's criticism, where it is *not* assumed that there are *sui generis* psychological entities and causes, but that psychological properties at least supervene on physical ones.[42] Then the criticism of Poland would be an internal one, to the effect that even where there is in principle some physical description of a process, Poland's approach cannot give a principled criterion for not calling any other possible description a part of physics. But in this case the interest of physics according to Poland would be, as noted, in the fundamental constituents of *all* occupants of space-time, and the forms of interaction and causation typical of them. Poland specifies that the classes of objects, attributes and principles which form the object of physics are 'minimal with respect to the descriptive and explanatory purposes they serve,' (1994: 124) which is to say that forms of description which are possible but not essential are just not part of physics. Although we have seen that Poland denies the view that the aim of physics is to provide a supervenience base, we can see that for him the physicist would be interested in those things which might *turn out* to form the supervenience base of the phenomena, in this case psychological ones, to which Daly refers.

This answers Daly's objection, but it is worth considering an alternative case where there are in fact irreducible *sui generis* causes of some type we would not intuitively want to call physical, for example mental causes, which had effects in space-time. Then it would follow that such things would indeed be part of the object of 'physics' as Poland describes its interests. Then, in a similar way to what we saw with Papineau, we would still have a complete science which we had decided to call 'physics' but which would not permit inferences to any kind of supervenience thesis about the mental, which is exactly what we would expect in a case where there were irreducibly mental causes. Poland himself does not explicitly consider such cases, since he seems confident that there are no such causes to consider. For all that, he is careful to make clear, though, that physicalism is significant in inverse proportion to the extent of what he calls the physical bases. More

---

[42] Difficulties with supervenience are discussed briefly in the conclusion to this thesis.

plainly, the less that is physical the more interesting it would be to say that other things supervene on the physical, or stand in some other relation of being determined by the physical.

Three accounts of 'physical' have, so far, been defended here: Snowdon's 'most general natural kind', Poland's 'fundamental attributes of all occupants of space-time', and Papineau's 'set of categories needed to explain paradigm physical effects'. I believe that Daly's most important objections to each of them has been answered. Following his criticisms Daly opts for a roughly inductive, admittedly provisional, conclusion to the effect that it is 'doubtful whether any satisfactory account will be forthcoming' (1998: 213). I suppose I might opt for a similar induction here, but it takes no effort to imagine an indefinite exchange of criticisms and defences of specific proposals concerning the 'physical' with their own inductive conclusions concerning the prospects of future efforts. Rather let me try to draw a general moral from what has been said so far, and also go on to say something more direct about the reasons for thinking that a relatively restricted physics might be complete.

6. Trivial Completeness for Pleasure and Profit

In one way and another I have, in the preceding section, been arguing that it is possible to sustain philosophical debates over the physical without having to specify criteria for membership of ***PP***. (That is to say without saying in any exhaustive or preëmptive way what is to count as physical.) All of the proposals which have been defended here, though, might seem to be poor alternatives to some kind of *a priori* characterisation of the 'real' boundaries of ***PP*** along the lines demanded by Daly. In each of the cases discussed here, though, the criteria offered have been related to and driven by the needs of the particular argument being made, whether by Snowdon, Papineau, Poland or for that matter Smart. And in all cases there has been a significant element of open-endedness to the criterion, allowing room for further enquiry on the part of the physical sciences.

The task of providing a principled *a priori* criterion might have seemed simpler in the seventeenth century.[43] Mechanists thought they had good reason to limit the set of physical properties very severely indeed, in Descartes' case to modes of extension, including size, shape and motion, but more usually to a slightly larger set at least including impenetrability. But it would be a mistake to be nostalgic about the good old days when we thought we knew *a priori* what all material or physical properties actually were. The mechanists in question were not only wrong about the specific properties of matter, but wrong to think that those properties could be known *a priori* at all.

Another way of making the point which I have been making in various ways in the preceding discussion of Daly's criticisms is that we just don't *need* to know this. What I would like to do now is return to

---

[43] A point noted in, for example, Stroud (1987: 265,267).

Papineau's proposal, and develop it somewhat with a view to settling the formulation question, and then moving on to see just how restricted a version of 'physics' might plausibly be seen as complete.

Before doing so it will be worth returning to an account of 'physical' discussed above, that proposed by Meehl and Sellars (1956). As we saw, they distinguish between two senses of physical for their purposes, one simply intended to capture something like 'scientific', or to express a general commitment to naturalism, and the second more specifically aimed at physical science. Their two characterisations are as follows:

> physical1: terms employed in a coherent and adequate descriptive, explanatory account of the spatio-temporal order.
>
> physical2: terms used in the formulation of principles which suffice in principle for the explanation and prediction of inorganic processes (Meehl and Sellars 1956: 252).

It seems reasonable to say of *physical1* that it is complete by definition. If there is some account of the spatio-temporal order which is coherent and adequate, then it would be complete in the sense at issue here.[44] Matters are somewhat different with *physical2*. Meehl and Sellars discussion is a defence of emergentism from a particular line of criticism due to Peper (1926) to the effect that the doctrine of emergentism is faced with an unresolvable dilemma.[45] Meehl and Sellars, while not specifically arguing that emergentism is true, argue that it is immune from Peper's criticisms, and characterise the emergentist as one who denies that some phenomenon or other[46] can be considered a part of *physical2* even though accepting it could be a part of *physical1*. This makes it seem as though *physical2* could not possibly be complete by definition, although it is possible that it could be complete by definition, and also incapable of describing the entire spatio-temporal order by definition. Whether or not *physical2* is complete by definition depends on how Meehl and Sellars want the 'in principle' clause to be understood. Fortunately I don't need to work out how to do this, and I will assume a weak reading which does not make *physical2* complete by definition, so that the most significant feature of their account is the limitation of 'physical' to inorganic phenomena.

As we have also seen, on Papineau's account 'physics' *is* trivially complete: anything which bears on the likelihood of any physical outcome is part of physics. For the purposes of the argument of the present section of this chapter I want to take the notion of a trivially complete physics as a premise. Doing this does not make epistemological completeness trivially true: we cannot make our knowledge of the causal processes behind certain types of outcome complete with a wave of the hand, as a glance at **ECT2** will make clear.

---

44 For the purposes of the present point I am fudging the fact that Meehl and Sellars' *physical1* is not framed in terms of causal relations, but rather explanation. Since they incline to a deductive nomological account of explanation, this is not an especially significant issue.

45 Peper's criticism is discussed briefly in Chapter Two section (3.3)

46 Their particular examples are 'rawfeels' and the like, i.e. qualia (Meehl and Sellars 1956: 252).

ECT2: A science **X** is complete iff **X** is capable of fixing the likelihood of **X**-type outcomes solely by reference to **X**-type entities, powers, states, laws, etc.

So physics would be epistemologically complete iff we could fix the likelihood of physical outcomes by reference to only physical factors. But we cannot simply define our epistemological resources into completeness, since whether or not they fix the likelihood of the outcomes we designate as physical is a matter of fact. So the correct way to understand Papineau's proposal is ontologically:

OCT2: A domain of reality **X** is complete iff **X** occurrences always have sufficient **X**-type causal ancestries.

The paradigmatic class of pre-theoretically given physical outcomes are the **X** occurrences, and making physics trivially complete means making anything which is needed (i.e. *whatever* is part of the causal ancestry of such outcomes) to fix the likelihood of those outcomes be part of 'physics'. Whatever its relations to what we might otherwise call physics, the science picked out by this proposal ignores our intuitions in the interest of causal closure. Perhaps for the time being we can call it 'schmysics' to forestall objectors with worries along the lines of Daly or even Poland.[47] Until further notice the present thesis is about the question of the completeness of schmysics. *Prima facie* the range of things included in schmysics could turn out to be close to the range picked out by Meehl and Sellars's *physical*$_1$. Schmysics is trivially complete, and *physical*$_1$ seems likely to be complete by definition.

I have already admitted that I propose to argue for the completeness of physics here. Given that, this last move concerning the completeness of schmysics may well seem to have something of the flavour of a philosophical party trick. Certainly 'schmysical', as I have just suggested it might be conceived, is the name of a pretty uninteresting, if not entirely pointless, set of properties, and that for the simple reason that the criteria for membership of the set are remarkably indiscriminate. It's worth pointing out that they're not *completely* indiscriminate, since truly epiphenomenal properties, for example, which are not needed to fix the likelihood of schmysical effects, would obviously *not* be included. And the schmysical properties are not pointless, because by definition they are the ones that do causal work with reference to physical outcomes. But there's more. Taking another cue from Papineau we can supplement the completeness by definition with some or other kind of exclusion hypothesis as well. In Papineau's case, as we saw, the conjecture was that 'physics' not need to make reference to psychological categories. Let's do the same with schmysics, so that we hypothesise that the schmysical, i.e. that which does causal work with reference to physical outcomes, will not include any psychological categories.

---

47 I stipulate that there are no schmysical effects which can be independently characterised: the 'schmysical' is that which is needed to account for *physical* effects, taken as pre-theoretically given in the sense of Papineau. Physical effects are schmysical by fiat.

What needs emphasising here is that it is not particularly important *what* exactly is excluded by the additional hypothesis. The first and most important point to note is that the proposed restriction that 'schmysical' not include some type of independently distinguishable properties *cannot* be true by definition in cases where schmysics *is* complete by definition. You can't have both. This means that for any given exclusion, it is an *empirical* question whether schmysics is both complete and devoid of the kind of property in question. (Since physicalism is usually a thesis about two classes of attributes, only one of which is *prima facie* physical, part of what makes the completeness premise interesting in *any* given case is exactly what it is that is excluded.) That is why the trivial completeness of schmysics is not a party trick - there's an empirical and contingent aspect to the proposal too. Returning to 'physics' for a moment, it's worth noting that in a sense all that is done here is the conceptual inversion of some details of fairly traditional ways of thinking. As noted above there are those who take it to be a matter of definition that 'physical' and 'mental' are opposed. For such thinkers the failure of the mental to fall within the domain of the physical is the trivial bit, and the question of the completeness of physics the difficult empirical question. With schmysics things are the other way around.

Having both the trivial completeness of schmysics and the contingent exclusion hypothesis helps specify the kinds of thing which might be expected to count as *evidence* for or against the view that the schmysical is both complete and free of whatever is excluded. So, for example, if Eccles (e.g. 1980)[48] was right and particles in the brain accelerated from time to time for no discoverable *physical* reason, but these events also turned out to be correlated with independently identifiable psychological states, then that would tend to support the view that these outcomes *needed* reference to psychological categories in order to fix their likelihood. That would count against the combined thesis that schmysics was complete and free from psychology. Either the complete science would have to include psychology hence sustaining no interesting schmysicalism about the mental, or we would have to admit that it was not complete.

If excluding something made schmysics more interesting (so that if it *was* true that the mental was not needed for a complete schmysics, then it would be possible to be a schmysicalist about the mental) then further exclusions might reasonably be expected to make it more interesting, and perhaps controversial, still. Recall Poland's (1994: 114) observation that the significance of physicalism is inverse proportion to the extent of 'the physical bases'. The reason is clear: the less there is that is 'physical' the more interesting it is to say, on the basis of further argument from the completeness of the physical, that *everything* is, in some sense, dependent on the physical. As with physics so with schmysics.

---

[48] See Chapter Two section (2.2).

The same basic approach will work just as well for debates other than those concerning the relationship between the psychological and schmysical. Depending on the thesis at stake we can impose variously strong constraints on what we allow to be part of the schmysical. Staying with the current exclusion principle for a moment we can see how this would work. In the event that there were *biological* properties (i.e. of living entities) which needed to be referred to in explanations of the paradigmatic physical effects in question, but not psychological ones, then what Papineau (1993) might call 'physics' and I am calling schmysics would be complete in the required sense, even though including parts of what we would usually call biology. From Papineau's point of view this is only an objection to the extent that we are worried about the word 'physical' though, since a complete science of such a sort would still be philosophically significant as far as the place of the mental went.

If, though, we wanted to make some kind of claim about the relationship *between* biological properties and schmysical ones along the same lines as we did with psychological properties then we could simply make the exclusion more restrictive, and generate a new empirical thesis to the effect that schmysics is complete and need never refer to biological properties. And as before there would be various kinds of evidence which might bear on the question. Nineteenth century anatomists, for example, searched for evidence in the energy intake and consumption of living things that might show their access to additional forms of energy of some kind, and failed to find it.[49]

Where has this taken us? Well, as it happens schmysics is now the trivially complete science of the causal antecedents of physical outcomes, which we hypothesise need make no reference to either psychological or biological categories. In other words, schmysics is about the inorganic, just like Meehl and Sellars's $physical_2$ which means that I can drop the charade, and start calling it physics again. We need not stop here, though. Recall that Poland (section 4.1 above) was worried about the lumping together of physics and chemistry implied by Meehl and Sellars.

The most restrictive version of the exclusion hypothesis that seems likely to have much application, then, would be one excluding *chemical* factors. There is a historical background for exactly this possibility. For a time it was thought that chemical phenomena required reference to forces of affinity between elements and substances of various kinds. If true, this would have meant either that a complete physics would have had to include chemical forces, or that a physics with no chemical categories could not be complete. But the explanation of chemical bonding by quantum mechanics seemed to show that no such forces were required.[50] Similarly various experimental programmes in the nineteenth century appeared to show that it was not

---

[49] Emergentism is discussed in more detail in Chapter Two below, and these short remarks on evidence are considerably amplified in section (7.1) of that chapter.

[50] For example Schlick (1953: 524) states emphatically that the laboratory production of Urea in 1828 by Wöhler 'refuted once and for all the doctrine that the synthesis of organic compounds requires a special force.'

necessary to suppose that additional 'vital' forces were required to account for biological phenomena from digestion to perception.

For the purposes of the arguments which follow two versions of the Completeness Thesis present themselves as worth of discussion, then, one which, in Kim's idiom, 'draws the line'[51] at biological categories, and one which does so at chemical categories. This gives us the following:

> **OCT3**: The physical domain of reality is complete iff physical occurrences always have sufficient causal ancestries including no biological, or psychological aspects.
>
> **OCT4**: The physical domain of reality is complete iff physical occurrences always have sufficient causal ancestries including no biological, psychological, or chemical aspects.[52]

So much for formulation, and the most preliminary considerations of plausibility. Rather than attempt a general survey of the possible evidence *for* the completeness of physics,[53] I will approach matters as follows: In the remainder of this thesis I consider a range of arguments which, if correct, indicate that one or both of these forms of the Completeness Thesis are *false*. Evidence for the Thesis, and further arguments for its truth or strength are developed and defended in those contexts, rather than pre-emptively. Before concluding the present chapter, then, it would help to spend some time considering ways in which evidence of whatever sort might bear on the question of the truth or falsehood of the Completeness Thesis.

7. Ways of the Completeness Thesis Being False

As is to be expected it is the epistemological form of the Thesis which is most directly amenable to evaluation. Not only that, as matters stand it is, as noted, unarguably *false* for any reasonable interpretation of physics. Recall:

> **ECT2**: A science **X** is complete iff **X** is capable of fixing the likelihood of **X**-type outcomes solely by reference to **X**-type entities, powers, states, laws, etc.

Physics as it stands is manifestly incapable of delivering the required goods. It does not form an integrated body of knowledge, and as well as not being integrated contains, if taken as a whole, various areas of

---

51 In a discussion of causal closure principles Kim (1998: 40) describes the key consideration as whether in the course of tracing the causal ancestry of a physical event, it will be necessary to 'crossing the boundary between the physical and nonphysical'. My concern here is with how far down the line can be pushed while still preserving completeness.

52 Feigl (1953: 384) admits the difficulty of giving a non-arbitrary criterion for distinguishing physical from non-physical laws, and proposes that a 'more definite meaning may be suggested by formulating the thesis of physicalism in the following way: The set of physical laws which enables us to deduce the facts of chemistry will be sufficient for biology and psychology.'

53 Although I make reference to such evidence in a number of places in what follows, including Chapter Two (especially section 7).

conflict and inconsistency.[54] Even if we were to assume with heroic optimism that all of these problems could be dealt with by means of the proverbial 'large enough sheet of paper and sharp enough pencil' approach, the thus integrated body of theory would not be complete in the required sense, for the simple reason that there are all manner of unresolved problems and empirical questions in physics at all scales and with respect to all phenomena.[55] To assume that these can all be answered by physics in the end is partly to beg the question, and also, perhaps inevitably, to create a need for an even larger piece of paper and a vastly sharper pencil to deal once again with the consistency problem.

The epistemological version of the Thesis would be on its ideal ground in the event that the world was best described by some deterministic theory. Persistent scepticism aside, being in a position to *predict* actual outcomes with the accuracy made possible by a true deterministic theory would be just about the best available kind of confirmation. Determinism is less straightforward and simple than classical modern thinkers, and the architects of rational mechanics, such as Laplace, tended to think, though. John Earman (1986) has discussed the ways in which even classical mechanics is not deterministic in the ways traditionally expected, and various other considerations indicate that even the behaviour of some deterministic systems may well be impossible to predict (e.g. Lorenz 1963).

Failures of epistemological completeness do not in general indicate any automatic difficulty for ontological completeness though. This is, in fact, one of the principal theses defended in what follows. As the case of non-linear dynamics shows, even deterministic systems may be genuinely unpredictable. More generally, any direct inference from failures of prediction to failures of ontological causal closure would be a *non sequitur*. Alleged counterexamples are always subject to reinterpretation.[56] And even without a successful predictive theory we might well have good grounds for believing in physical causal closure, in the event that we could defend the claim that the only kinds of causal process there was any reason to believe operated in some given system were physical ones. Let us look, then, at the ontological form of the Thesis:

> **OCT2**: A domain of reality **X** is complete iff **X** occurrences always have sufficient **X**-type causal ancestries.

In this case it is clear that where the domain **X** in question is the physical, it would suffice for the falsehood of the ontological Completeness Thesis if there was any well substantiated evidence that there was some class of physical occurrences which either did *not* have sufficient physical causal ancestries, or the

---

54 These theoretical difficulties vary from the well-known issue of the apparent conflict between general relativity and quantum mechanics, to wider problems about the relationships between any fundamental law and the phenomenological laws relevant to particular phenomena. Cartwright's criticism of fundamental laws (1980, 1983) is discussed in Chapter Four below.

55 The difficulties here vary considerably too, from cases of hypothesised but either undiscovered or contested particles to questions about the large-scale structure of space-time.

56 The classic discussion of this point is in Duhem (1954).

sufficient causal ancestries of which *did* have a non-physical component. What might lead us to think that such a state of affairs obtained?

For a start it is important to note that the mere availability of some causal claim which relates a physical outcome to a non-physical antecedent, for example 'my wanting caused my walking' will *not* by itself be decisive. More generally any claim so relating a physical effect with a non-physical cause will be unproblematic from the point of view of the Completeness Thesis just as long as it is still possible to defend the claim that there is a sufficient physical cause as well. Given two such claims there are a range of possible responses including, *inter alia*, accepting the overdetermination or rejecting it and arguing for some kind of identity thesis.[57] It is not my intention here to get into those questions, though, simply to focus on the question of completeness. And what is needed to create difficulties for the Completeness Thesis is reason to believe that there is not a sufficient physical ancestry for some physical outcome.

There are various ways in which this might occur. Dualistic interactionism is one, any of a number of forms of causal emergence involving 'downwards causation' another. The matter of fact in all cases is the same, though, and that is this: if any of these views is true, there will be some physical outcomes which just do not have sufficient physical causal ancestries, however those ancestries are to be understood.

The ontological form of the Completeness Thesis could turn out to be not so much false as utterly unacceptable. What I mean by this is that rather than evidence *for* positive failures of completeness, it might turn out that there were reasons to think that the very notion of a well defined domain of the physical was incoherent, or at least that our best evidence indicated that there was no such domain to be found at the actual world. This is one way of thinking of the force of Cartwright's assault on traditional confidence in high level theoretical laws of physics, a view she dubs fundamentalism. I postpone treatment of that view to the fourth chapter of this work, and in the interim focus more directly on views which suggest that the Completeness Thesis is false.

8. Conclusion

So, contrary to the arguments of some, the Completeness Thesis can be satisfactorily formulated. Not only that, it has sufficient *prima facie* grounds of plausibility for it to be worth investigating seriously. As already noted, the following three chapters are dedicated to considering a selection of critical positions with a view to seeing how far the Thesis can be defended, and to work out the ways in which it may need to be modified or revised.

---

[57] For example Lewis (1966), Papineau (1993).

Given the noted obstacles in the face of any chance of direct demonstration of either the truth or falsehood of the Thesis, what will be needed is a careful attempt to weigh up the available evidence and also to establish how best to interpret that evidence. The two biggest errors to be avoided throughout are forms of dogmatism: on the one hand the physicalist evangelist may attempt to maintain the 'in principle' completeness of physics in the face of all apparently recalcitrant counter evidence, by heroically attempting to redescribe the evidence in acceptable terms.[58] On the other hand those anxious for whatever reason to see the Thesis defeated might persistently attempt to run the noted *non sequitur* from failures of epistemological completeness to the falsehood of the Thesis in its ontological form. Both forms of dogmatism need to be resisted.

[58] The problem here is one I have discussed in connection with the question of evidence for physical determinism, see Spurrett (1997, section 2).

Chapter Two: Bhaskar and Emergence

'...the manner of go in the enriched system is different.' – Conway Lloyd Morgan (1923: 20).

1. Introduction

As noted in the preceding chapter one of the ways in which the Completeness Thesis could turn out to be false is in the event that a suitable kind of emergence turned out to be true. Specifically what would be needed would be grounds to believe that there were causal laws or causal powers with physical consequences, which were not themselves physical, or were at least not completely physical in their causally relevant respects. Emergentism, which is typically committed to just such powers or laws, was popular for much of the later nineteenth and early twentieth century, declining in popularity from the 1920s, and then more recently enjoying, in name at least, a striking revival.

Most arguments for emergence are empirical: either the inability of physics to account for certain phenomena is urged, or the credentials of some causal relation to be recognised as an emergent law defended, and in some cases both. In these cases the main burden of philosophical argument is to show that emergentism is coherent, and to offer a schematic account of the differences between systems in which emergence in manifest and those in which it is not. A somewhat unusual feature of Bhaskar's (1978) defence of a kind of emergence is that it claims to arrive at an emergentist conclusion from a primarily philosophical argument. Bhaskar's approach is discussed in particular detail here both because of this peculiarity and also because features of his treatment of experiment will be useful in the discussion of the work of Cartwright which appears in Chapter Four.[1] There are a small number of attempts to show that the very notion of emergence is incoherent. These arguments are discussed in what follows, but it is striking that even anti-emergentists tend to reject them, and assert instead that the failings of emergentism are empirical rather than philosophical.

Although for most of my purposes emergentism and dualism can be treated together (see section 3 below) emergentism is typically more closely related to specific empirical claims and experimental data than dualism. Part of the reason for this is that emergentism is commonly (although not essentially) a view which includes reference to biological processes, often involving interpretations of specific features of the behaviour and development of living systems.[2]

---

1 A further motivation for my criticism of Bhaskar on this point is that I used to agree with him. See Spurrett (1997)

2 Recall the reasons given for my attempt to pursue the question of the truth of the Completeness Thesis with as little direct contact with the philosophy of mind as possible (Introduction section 1.3). Since emergentism involves the same type of clash with completeness as dualism, no significant arguments are excluded by focussing on science and philosophy of science in preference to philosophy of mind here.

Before examining Bhaskar's argument for emergence, I spend some time setting out the background to emergentist thought, especially its roots in the work of Mill, and also set out a few important distinctions between different emergentist theses, not all of which are directly relevant to the question of the defensibility of the Completeness Thesis.

2. Background to Emergentism

There is a wide variety of forms of emergentism, and a number of more or less independent emergentist traditions. Not only that, emergentism exerts a powerful attraction on the imagination and attention of eccentrics and crackpots of various kinds. This is not the place to attempt an explanation of why this might be so, or what it is about the particular mixture of apparent scientific respectability with seemingly unlimited speculative license that makes emergentism seem like a life raft for a range of under-motivated and indifferently argued positions.[3] Rather I mention it because, despite all that, there is a responsible and coherent core emergentist position, which may turn out to be empirically defeasible, but cannot be rejected on the grounds of philosophical incompetence on the part of its adherents.

There are several emergentist traditions, although in the twentieth century they are far more difficult to separate into more or less independent strands. There is, though, a fairly coherent British emergentist tradition including figures such as Mill, Alexander and Broad. Other major emergentist figures include Driesch,[4] Kapp,[5] Elsasser,[6] and others.[7]

---

3 The most striking examples I ran into in the course of innocently following the bibliographic trail created by the term 'emergence' are Mott and de Selby. There are surprisingly many works by F. K. Mott (e.g. 1948) whose brand of emergentism, with 'configurational forces' and all, is apparently substantiated by reference to 'evidence' (his word) gathered in the course of 'dream analysis' (1948: vi). On the other hand the turn of the century experimentalist de Selby advances spectacular claims in favour of emergentism with reference to the phenomena of chemical bonding on the basis of several irreproducible researches involving a device which he called the 'water box', aspects of which are described in O'Brien (1964).

4 See, e.g., Driesch (1908) which is in many ways similar to the emergentist writings I discuss below, at least insofar as the bulk of Driesch's neo-vitalist approach is concerned with biological matters, but which differs from British emergentism in his concern with, for example, the 'Logic of History' (1908 (vol. 1): 297-324). Broad's derision about 'entelechies' (1925: 58, 69) notwithstanding, Driesch is careful to explain how his view of entelechy is, supposedly, compatible with the conservation of energy (1908 (vol. 2): 218-225) and it is clear that his claims are intended to be understood as empirically substantiated, and offered in a generally naturalistic spirit.

5 Kapp (e.g. 1940, 1951) is cheerfully, and generally unfairly, abusive of the emergentist tradition (1940: 106-7, 214), the 'roots' of which he sees, making no mention of Mill, in Alexander (1940: 211). Although rejecting the label emergentist himself Kapp's major contribution is the idea of the 'diatheme' (1951: 77-125, 1940: 256-276) which, in the manner of all even moderately thoughtful emergentists he takes care to reconcile with the conservation of energy (1951: 130-132), and which he regards as being a 'selecting, guiding, controlling activity' (1940: 256) not explicable in terms of pure mechanism.

6 Elsasser (1958) mixes elements of standard emergentism with mid-century scientific sophistication, and considerable caution. So, he offers the standard rejection of Cartesian dualism (1958: 1), and also Driesch-style entelechy (1958: 2-3), while also being up to date with quantum mechanics, cybernetics and information theory. On the basis of the supposed increase in information represented by the development of a living system, Elsasser argues for what he calls 'biotonic causation' (1958: 19ff) which is significantly different

My task here is not the construction of a general history of emergentist thinking, and for the purposes of exposition I will concentrate on the work of the British emergentists. I have seen no evidence that there are any significantly different arguments to be found in other traditions, or that they have any independently interesting analyses of any of the experimental data typically taken to support emergence. As we will see the crucial issue, that of downwards causation, is ultimately empirical in any event.

The main outlines of the history of British emergentism (following McLaughlin 1992) are as follows: It begins with Mill's *System of Logic* (1843)[8] which sets out the core idea of there being different ways in which causal influences can combine, one (Mill does not use the term 'emergent') of which is 'heteropathic' or non-mechanical. In different ways this line of thinking is visible in Alexander Bain's *Logic* (1870), George Henry Lewes' *Problems of Life and Mind* (1875), Samuel Alexander's *Space, Time, and Deity* (1920), and Lloyd Morgan's *Emergent Evolution* (1923). It is Lewes who coined the term 'emergent' to describe the causal combinations which Mill had called heteropathic. The last great work in this tradition, following McLaughlin's reading of the history, is C. D. Broad's *The Mind and Its Place in Nature* (1925), after which the line all but died out, although seeing intermittent support from such figures as Sperry, Eccles and Popper, discussed further below.[9]

British emergentism (or sometimes for the remainder of the present section simply 'emergentism') is naturalistic – the emergent causal powers to which it is committed are not supposed to be thought of as souls, entelechies or elan[10] (McLaughlin 1992: 49). Indeed emergentism is typically a form of *materialism*, which embraces anti-dualism, or at least opposition to any form of substantial dualism, while also rejecting a certain form of materialism which some emergentists called mechanism (e.g. Broad 1925: 45-6), and which is one

---

from mechanical causation in typical emergentist ways. Although Feigl (1958: 376) attributes to Elsasser the view that the laws of physics are a special case of the laws of biology, Elsasser himself is careful to distance himself from what he takes to be Bergson's thesis that 'inorganic matter is a degenerate form of organic life' (1958: 166), seeing both mechanical and biotonic causation as limit cases with almost all actual causation in living systems somewhere in between (1958: 165). Ultimately Elsasser's view should be seen as Bohrian, in Elsasser's words involving 'generalized complementarity' (1958: 13).

7 More recent emergentists besides those directly covered in the text include Whitehead (1978), Lovejoy (1961). Feigl (1958: 376) attributes to Whitehead the view that the laws governing the behaviour of electrons would be different in living bodies.

8 In his *Autobiography* Mill notes the influence of a chemistry text by a certain Doctor Thomson, which he indicates played a formative role in his thinking on these issues (Mill 1989: 129). The text in question is likely to be the *System of Chemistry* by Thomas Thomson. I have not followed up this clue to the prehistory of emergentism.

9 Stephan (1992) argues for a three-stage history of emergentism largely compatible with the account I have given above, follows McLaughlin (1992), but regarding Sperry, Popper, Eccles and others as a distinct third stage of emergentist thinking. It is possible to regard McLaughlin's account of the demise of emergentism in the face of quantum mechanics as slightly overstated, in which case his account and Stephan's are in greater agreement. The question of quantum mechanics and emergence is taken up in section 7 below. Feigl (1958: 411-413) also makes some relevant observations concerning the role of QM in the downfall of emergentism.

10 Although Driesch uses the term entelechy, what he has in mind is closer to the general spirit than the typical critical remarks by other emergentists suggest. See footnotes 4-6 above.

possible version of the thesis that physics is complete.[11] Most emergentists believed in the existence of fundamental physical particles, but rejected the notion that any description couched only in terms of these particles could be exhaustive in all cases:

> According to British Emergentism, there is a hierarchy of levels of organizational complexity of material particles that includes, in ascending order, the strictly physical, the chemical, the biological, and the psychological level. There are certain kinds of material substances specific to each level. And the kinds of each level are wholly composed of kinds of lower levels, ultimately kinds of elementary material particles. Moreover, there are certain properties specific to the kinds of substances of a given level. These are the "special properties" of matter. Physics is at the base of the hierarchy of sciences in that it is concerned with the properties common to all or nearly all material substances at whatever level of organizational complexity: These include properties such a inertial and gravitational mass, and electrical charge. Physics studies the organizational relationships such properties are responsible for, e.g., gravitational attractions, electromagnetic attractions and repulsions, etc. But the material substances of a given level participate in certain organizational relationships in virtue of their special properties too. And it is part of the business of a special science to study the organizational relationships peculiar to a specific level and to formulate the laws governing those relationships (McLaughlin 1992: 50).

As can be seen from this overview, emergentism can properly be seen as a view about the status of what would now be called the special sciences, and their relationship to an assumed basic physics. For an emergentist at least some special sciences are distinct and autonomous from physics because they have as their object phenomena characterised by, and only explicable by reference to, causal powers arising only when the objects studied by physics are combined in certain particular ways.

Following McLaughlin (1992: 52f) I will present emergentism as a doctrine about forces, in which case the emergentist thesis can be expressed as one to the effect that there are what McLaughlin calls 'configurational forces', that is forces which are exerted 'only by certain types of configurations of particles' (1992: 52).[12] As forces an important feature of configurational forces is that they are capable of bringing about changes in velocity, i.e. accelerations. That these forces are exerted by configurations distinguishes them from standard physical forces which can typically be exerted by suitable pairs of particles, as is the case with gravitational forces. A crucial feature of configurational forces is that they are supposed to be *fundamental*, that is not

---

11 As we will see, in section (7.1) below, the form of mechanism these thinkers had in mind points to a restrictive notion of completeness, where all causal relations are mechanical and operate additively.

12 In some cases physical laws are typically formulated in terms of energy rather than forces. For the purposes of the account of emergentism here I express everything in terms of forces, a strategy in which I follow McLaughlin (1992). As McLaughlin notes it would be possible to describe all of this in terms of energy, but with no significant gain in clarity. Besides, the emergentists themselves tended to think in terms of forces, for example Bain (1870:21) on the subject of 'nerve force'.

reducible to or explicable in terms of forces which are not so reducible.[13] Particle pair forces need not *only* be exerted by pairs of particles: the individual forces can act in combination when more than two particles are present. The key point is that in contrast with supposed emergent forces, it does not take a special arrangement to bring about the action of particle pair forces.

Since these forces have effects at the level of fundamental particles, or more generally at the next level down in the emergentist hierarchy if that level is not fundamental particles, they are one way of thinking of the operation of 'downwards causation' (Campbell 1974).[14] It is this which raises the possibility of conflict with the Completeness Thesis, since in the event that there is non-redundant downwards causation from the non-physical to the physical there will be physical outcomes which do not have complete physical causal ancestries.

Given its central role there are few better places to start than with a work which played a crucial role in the development of various emergentist positions in the nineteenth century, Mill's *System of Logic*.

2.1. Mill and British Emergentism

Mill's account of emergence (he does not use the term himself) is the *locus classicus* of the British Emergentist tradition.[15] Given that Mill's treatment is utterly paradigmatic, and that my objective here is not a history of emergentism, I will concentrate on what he says.[16] The crucial discussion occurs in Book III, Chapter VI, entitled 'On the Composition of Causes'. Mill distinguishes between two models of the combined action of forces, the mechanical and the chemical.[17] On the mechanical mode he says:

> In this important class of cases of causation, one cause never, properly speaking, defeats or frustrates another, both have their full effect. If a physical body is propelled in two directions by two forces, one tending to drive it to the north and the other to the east, it is caused to move in a given time exactly as

---

13 So forces such as London forces not exerted by particle pairs are not configurational in the sense used here because they are not fundamental; they are explicable in terms of electromagnetic forces (McLaughlin 1992: 52).

14 Campbell's (1974) label 'downwards causation' is generally taken to suggest the operation of independent causal powers in the manner of what I call strong emergence. His paper is ambiguous on this point, since he that 'downwards causation' operates 'in complete compatibility with the physical model of causation' (1974: 180). On the other hand he also suggests that the higher level laws (which he sees as following from natural selection) 'determine in part the distribution of lower-level events and substances' (1974: 180). The term 'downwards causation' is also sometimes used in cases where strong emergence has been ruled out entirely: Auyang's (1998: 54f) endorsement of talk of downwards causation is immediately succeeded by the qualification that systems in which downwards causation operates have no 'independent funds of causal power'.

15 My discussion of Mill in the present sub-section owes a great deal to McLaughlin's elegant and careful (1992) essay.

16 References to Mill (1843) are to the text of the first edition of his *System of Logic*, while references to Mill (1973) are to the edition in Mill's *Collected Works* including all significant changes and variations between the eight editions of the work. Quotations from this edition follow the wording of the 8th (1872) edition text.

17 In the first edition of the *System of Logic* he notes that the distinction is so 'fundamental' as to require a chapter devoted to it. In all subsequent editions the distinction is rather described as 'radical' (1973: 370).

> far in both directions as the two forces would have separately carried it; and is left precisely where it would have arrived if it had been acted upon first by one of the two forces, and afterwards by the other. This law of nature is called, in dynamics, the principle of the Composition of Forces: and in imitation of that well-chosen expression, I shall give the name of Composition of Causes to the principle which is exemplified in all cases in which the joint effect of several causes is identical with the sum of their separate effects (1843: 428).

The model for what Mill is referring to here is Newton's account of the composition of forces in the *Principia Mathematica*. There (1934: 14)[18] Newton specifies that the net effect of two mechanical forces can be shown to be represented in direction and magnitude by the diagonal vector of a parallelogram having as two of its adjacent sides vectors representing the magnitude and direction of each individual force. The notion of separating a tendency to motion[19] into components is a crucial feature of Galileo's account of projectile motion and falling bodies generally (e.g. 1914: Fourth day), and is also worked out to a different purpose in Descartes' account of the reflection of light (1985: 156-164). For these and other thinkers the result of multiple simultaneous mechanical forces or tendencies to motion can be seen as the simple sum of the individual forces or tendencies.

Lewes (e.g. 1975: 414) called such additive forces *resultants* to distinguish them from the other type of force which emergentists acknowledge, which he called *emergents*. Mill's version of the distinction was, as noted, between 'mechanical' and 'chemical' modes of composition, and while maintaining that *most* causation followed the mechanical mode, he also claimed that the principle of composition of causes did not have 'universal validity', that is that not all effects could be regarded as caused by the simple sum of a number of individual influences:

> The chemical combination of two substances produces, as is well known, a third substance with properties different from those of either of the two substances separately, or both of them taken together. Not a trace of the properties of hydrogen or oxygen is observable in those of their compound, water. The taste of sugar of lead is not the sum of the tastes of its component elements, acetic acid and lead or its oxide; nor is the colour blue vitriol a mixture of the colours of sulphuric acid and copper (1973: 371).

Mill holds that the properties of at least some chemical compounds cannot be deduced from our knowledge of the properties of the constituents of the compound, although in a fairly typical emergentist gesture he adds that this is so 'at least in the present state of our knowledge' (1973: 371). I suggest that what Mill has in mind here, though, is not merely that the properties of compounds are not deducible from those of

---

[18] The composition of forces is discussed at length in Chapter Four section (5.4) below in connection with Cartwright.

[19] The term 'force' is either not used by these thinkers, or not used in its strict Newtonian sense, which is the reason I use the term 'tendency to motion' with respect to Galileo and Descartes.

their constituents, but that the laws by means of which the properties come to be instantiated are themselves not deducible either. This is to say that it is the heteropathic nature of the laws which lies behind the non-deducibility of the properties.[20] Mill sums up the differences between the mechanical and chemical modes as follows:

> There is, then, one mode of the mutual interference of laws of nature, in which, even when the concurrent causes annihilate each other's effects, each exerts its full efficacy according to its own law, its unique law as a separate agent. But in the other description of cases, the agencies which are brought together cease entirely, and a totally different set of phenomena arise: as in the experiment of two liquids which, when mixed in certain proportions, instantly become, not a larger liquid, but a solid mass (1973: 373).

Although urging the distinction between mechanical and heteropathic composition, Mill is also adamant that the mechanical mode is generally operative, and regards cases of heteropathic causation as 'exceptional' (1973: 373). Not only that, even in cases with a heteropathic aspect the hand of mechanical composition can still apparently be discerned:

> There are no objects which do not, as to some of their phenomena, obey the principle of the Composition of Causes; none that have not some laws which are rigidly fulfilled in every combination into which the objects enter. The weight of a body, for instance, is a property which it retains in all the combinations in which it is placed (1973: 373).

Heteropathic or emergent laws are to be thought of as having specific effects, and operating at definite times, or at least under definite conditions, so that the laws applicable to an entity in virtue of the laws governing the behaviour of the elements out of which it is composed when they are *not* configured in some way which occasions the production of a heteropathic effect continue to operate at other times. Vegetables, for example, are not to be thought of as *losing* their mechanical and other aspects, but rather as continuing 'to obey mechanical and chemical laws, in so far as the operation of those laws is not counteracted by the new ones which govern them as organised beings' (1973: 373). Not only that, it can, according to Mill, happen that the new laws may co-exist with some portion of the old, yet supersede another, and 'may even compound the effect of those previous laws with their own' (1973: 373), furthermore the new laws may themselves enter into simple additive compositional relationships. (1873:376). Through all these various possibilities, the hallmark of heteropathic causation in general, though, is described as follows:

> ... the principle of Composition of Causes [...] fails [...] where the concurrence of causes is such as to determine a change in the properties of the body generally, and render it subject to new laws, more or less dissimilar to those to which it conformed in its previous state (Mill 1973: 377).

---

[20] The importance to the early emergentists of what we would call emergent laws rather than emergent properties is easily underestimated, especially in many more recent discussions which often take the truth of the Completeness Thesis as a premise. See section (3.1) below.

The image Mill has in mind for the heteropathic laws is that they 'supersede' the lower level laws (1843: 433-4), rather than *contravening* or *violating* them (I return to this point in section 3.2 below). A final point which needs to be noted with reference to Mill is that he regards his commitment to heteropathic laws as entirely consistent with his thorough-going determinism, since the net effect of all the laws in play in any given situation is always unique, and hence if we knew all the relevant laws and facts, we would be able to make successful predictions:[21]

> If we could determine what causes are correctly assigned to what effects, and what effects to what causes, we should be virtually acquainted with the whole course of nature. All those uniformities which are mere results of causation might then be explained and accounted for; and every individual fact or event might be predicted, provided we had the requisite data, that is, the requisite knowledge of the circumstances which, in the particular instance, preceded it (1973: 377-78).

Mill's account of emergentism is not an argument for the truth of that doctrine, but rather a case for the coherence of a certain view of the laws of nature. The question of whether there are heteropathic laws, and if so which ones, are entirely empirical as far as Mill is concerned. In what follows I will take the main outlines of Mill's position as paradigmatic of emergentism, and for the most part (except in the discussions of Bhaskar) focus on the problem of what evidence can be brought to bear on the question of the defensibility of emergentism thus conceived. A few brief observations about contemporary emergentism may help make clear why little of significance is lost by thus focussing on Mill.

2.2. Contemporary Emergentists

Subsequent to the heyday of traditional emergentism, there have been a number of thinkers who have endorsed some kind of emergentist thesis for a variety of reasons. Of these the best known are probably Sperry, Eccles and Popper. There has also been a recent revival of interest in emergentism, largely although not exclusively fuelled by inter-related research into complex systems, cellular automata and chaos. It is important to keep these two streams separate, since the motivations for the use of the term 'emergent' and the factors taken to count as evidence vary in important ways in the two cases.

The work of Popper, Eccles (e.g. 1977) and Sperry (e.g. 1966, 1976, 1978, 1983, 1986) is for the most part standard emergentist fare along the lines we have seen with Mill. Popper, for example, wants to reject the 'materialist doctrine that men are machines' (1977: 14), endorses Campbell's notion of downwards causation (Popper 1977: 19, Campbell 1974) and states in his vocabulary that 'World 1 is not causally closed' (1977: 57). Eccles speaks in terms of a causal interaction between mind and brain which he calls 'liaison' (e.g. 1977: 361) and which bears comparison with Kapp's 'diatheme' (Kapp 1951: 77-125). Sperry makes a case for a kind

---

[21] This point is discussed again in connection with Oppenheim and Putnam (1958) in Chapter Three below (section 2.1) and also in the criticism of Bhaskar in the present chapter (section 6).

of macro-determination inconsistent with the completeness of physics, and largely based on interpretations of experimental evidence from his work on neurobiology and psychology. Unlike traditional emergentists almost all contemporary emergentists (certainly all of those considered here) seem to be evolutionary emergentists, which is to say that they couple an emergentist notion of heteropathic laws (in Mill's sense) with an evolution-inspired picture of increasing complexity over the course of time, in the sense of Lloyd Morgan.

What is important for my purposes is that none of the contemporary emergentists have significant independent arguments for the plausibility of emergentism, and hence that the plausibility of their views remains an empirical matter.

On the other hand much of the contemporary revival of interest in emergence has to do with a quite different set of considerations. Researchers in various fields, typically making using of computer models and simulations, have found in different ways that reductive descriptions or explanations of the phenomena which they study are inadequate or even impossible (see, for example, Forrest 1991, Soucek *et al* 1992). There is no question at all that these models and simulations are causally closed at at least *one* level of description, since they are all the products of deterministic algorithms and programs running on digital computers. The use of the term 'emergent' here relates essentially to the putative non-deducibility of larger scale properties of the simulations. In terms to be explained shortly this research relates to *weak* emergence, and has no particular interest from the point of view of the Completeness Thesis.[22]

3. Consequences for the Completeness Thesis

In the preceding discussion of emergentism I have, for the most part, focussed on the aspects of emergentist thought most directly relevant to the present enquiry into the Completeness Thesis. Before explaining in a more pointed fashion how emergentism and the Completeness Thesis are at odds, some brief comments on other emergentist theses, and the relationship between emergentism and dualism, are in order.

3.1. Strong and Weak Emergence

Before going any further it is important to clear up a potentially dangerous ambiguity in the use of the term 'emergence'. As can be seen from the preceding discussion of traditional emergentism there are at least two, sometimes related, categories of emergence the one, roughly, involving properties and the other causes. It is a feature of the contemporary revival of interest in emergence, and especially the current vogue for the term in the study of chaos, complexity and related fields, that the two senses of emergence are often systematically equivocated.

---

[22] That is to say that it poses no *challenge*. In Chapter Three below (section 5) an argument concerning weak emergence is deployed against Dupré's contention that failures of reduction are a threat to the Completeness Thesis.

It is easy enough to see, though, that a case in which it might be defensible to regard some property as emergent need not be one which carries any implication that some causal power is emergent also. Bedau (1997) gives the name 'weak emergence' to cases where as a result of causal processes at the micro-level of some system, structures, processes and relations arise at larger scales which could not typically have been predicted by reference to the attributes and possibilities for interaction of the fundamental constituents. In these cases it is also typically difficult or even impossible to describe the higher level phenomena in the terms suitable for micro-descriptions, even though all processes at the micro-level continue to take place in strict accordance with the standard and preëxisting micro-laws. I will follow Bedau's practice and refer to this type of emergence as 'weak' emergence in what follows, which naturally suggests using the term 'strong emergence' for cases where there is some emergent causal mechanism or influence, whether or not there is any suggestion of emergent properties.[23]

(Those with strong views about the metaphysics of properties, and the protocols of speaking about such things, may become jumpy at all this talk of properties 'emerging'. It would, I concede, be more correct to say that the conditions for some property being instantiated are emergent. Those wedded to minimalism about properties[24] might even prefer me to say, rather, that the conditions for some description being applicable are emergent. In what follows I will for the most part ride roughshod over these niceties, but at no point will I mean anything metaphysically loaded by the suggestion that a property emerges.)

Returning to strong emergence, though, we can say that the hallmark of this type of emergence is the action of new, or higher level, causal powers, which have an effect on the development of the system in which they are active, including an effect on the behaviour of the constituents of the system. A common label for this possible phenomenon is 'downwards causation' (Campbell 1974), or, in Lloyd Morgan's charming phrase we could say that 'the manner of go in the enriched system is different' (1923: 20).[25] Whatever their other commitments might be, the classical emergentists were united in their belief that the world was characterised by instances of strong emergence.

More exhaustive taxonomies of emergentist theses than my division into weak and strong varieties are possible, for example Stephan's (1992) categorisation in terms of non-additivity, novelty, non-predictability (in

---

23 It is a potentially interesting question whether strong emergence implies weak emergence, but not one into which I need to delve. (It does seem reasonable to suppose, though, that strong emergence could arise without any 'new' properties being instantiated at the same time, i.e. that strong emergence need not imply weak in all cases.)

24 The paradigmatic contemporary property minimalist is Armstrong (e.g. 1978), see also Swoyer (1996).

25 Lloyd Morgan uses this striking phrase in a number of places (including 1923: 88-9) but it is in his (1926: 68) that he explains that the idiom comes from the fact (Lloyd Morgan gives no reference) that 'Clerk Maxwell, when a boy, was wont to ask of everything: "What's the go o' that?".'

various forms) and non-deducibility.[26] Such taxonomies do not always co-incide my concerns here, so, for example, it is possible to read all of Stephan's options as forms of weak or property emergence. (Indeed it seems as though this is precisely what he has in mind.) In any event such finer distinctions tend to raise a host of questions and problems, such as distinguishing deducibility from predictability, which are not relevant to the question of the defensibility of the Completeness Thesis. For these reasons, in what follows, I will work with the strong/weak emergence distinction in the simple form in which I have just set it out.

Whatever might be involved in the emergence of properties, that is whether the key issue is non-deducibility, novelty or whatever, weak emergence just does not pose any difficulties for the Completeness Thesis. The reason that weak emergence need not trouble the defender of the Completeness Thesis is that the Thesis is not an inclusive view about what properties there are, but rather an exclusive one about the causal closure of the physical properties. Some property could have *all* the hallmarks which might move us to think of it as emergent, but as long as it has no impact on the 'manner of go' in the system in which it emerged, it remains uninteresting from the perspective of the present investigation.

3.2. Emergentism versus the Completeness Thesis

As we have seen, though, most of the classical British emergentists follow Mill in thinking that it is ultimately possible that there be some complete, deterministic and deductive theory embracing all occurrences. So they would, in the terms of Chapter One above (sections 4.1 and 5.4), have agreed that $physical_1$ could be complete. They also maintain that the behaviour of fundamental physical particles (in the existence of which they are unanimously believe)[27] can be affected by organisational features of the systems of which they are a part, in cases where this organisation is such as to result in the exercise of causes which are both fundamental and non-physical. To the extent that they do so, i.e. to the extent that they are strong emergentists who also hold that there are emergent laws over and above chemical laws, they would reject the completeness of anything like $physical_2$, and therefore hold the Completeness Thesis in the form I am considering it to be false.

It is worth noting an important qualification, which is that there is no *a priori* limit to the types of thing about which it is possible to be an emergentist, and that it is only those which involve commitment to non-physical parts of the sufficient causes of physical outcomes that present any risk of difficulty for the Completeness Thesis. Recall the following formulations of the Completeness Thesis from Chapter One above:

---

[26] Stephan's classification seems to view all of the types he identifies as variations on what I am calling weak emergence, since for the most part he takes the Completeness Thesis as a premise. The distinction between strong and weak emergence which I use is essentially the same as that of Bedau (1997). Another close precedent is Silberstein and McGeever's (1999) distinction between 'ontological' and 'epistemological' emergence, discussed in section (7.1) of the present chapter below.

[27] Alexander (1920), in keeping with the contemporary interest in relativity, seems in some ways to make space-time rather than fundamental particles primary. Nothing important for the purposes of the present discussion hangs on this question, since his space-time is a framework for a basically particulate physics.

ECT2: A science **X** is complete iff **X** is capable of fixing the likelihood of **X**-type outcomes solely by reference to **X**-type entities, powers, states, laws, etc.

OCT2: A domain of reality **X** is complete iff **X** occurrences always have sufficient **X**-type causal ancestries.

The strong emergentist threat more specifically, then, arises where directed at anything non-physical (chemical, biological, psychological, etc.), and entails the rejection of both versions of the Thesis where **X** is physical, since there will be physical occurrences which do not have sufficient physical causal ancestries, and hence with respect to which no science referring only to physical entities, powers, states, laws, etc., could possibly fix the likelihood.

Causes or laws which are emergent, but nonetheless physical, need not present any problem, though. Were one to hold, for example, that the processes by which crystals were formed were heteropathic in Mill's sense, but still to regard crystals as unproblematically physical then we would have an example of a strong emergentist thesis which was entirely compatible with the Completeness Thesis. Some examples of such theses are discussed in section (7) below. Before turning to the relationship between dualism and emergentism, I want to return briefly to the objection to *physical*$_2$ raised by Block (1980) supported by Poland (1994), and discussed briefly in Chapter One (section 4.1) above.

As was seen then, Block's concern was with the possible existence of laws which arose only in living brains of some significant but unspecified complexity, but which once discovered turned out to be translatable into physical laws. The conceivability of such laws was supposedly reason to be wary of using anything like 'inorganic' to characterise the physical.

It is not entirely clear what Block thinks he is worrying about here. Part of the problem is that Block's description fails to make precise what he means by the translation of his putative laws into physical laws, which suggests a range of procedures from the trivial to the major scientific achievement.[28] Nonetheless the preceding discussions of emergentism allow a number of significant possibilities to be identified.

One is that the translation would involve expressing the law as a physical consequence (whether an event regularity or a description of a capacity or tendency) of some set of physical conditions. In the event that this was possible the law would be unproblematically physical in the sense of *physical*$_2$, or even a more restrictive physical excluding the chemical, which was precisely the account of the physical to which Block's worry was supposed to be an objection. In this case it would follow that the law could operate outside the brains which Block supposes to be their normal context, as long as the relevant physical conditions obtained. Block's

---

[28] Nagel (1963: 336-397), for example, discusses a range of constraints on a worthwhile reduction none of which are explicitly required by Block (1980).

concern would then collapse into a worry about the context in which certain laws were discovered, which need pose no difficulty for the Completeness Thesis, and hardly merits consideration along with the possibility of strong emergence.

If, though, the laws were not thus translatable, perhaps because they could not be formulated in a way which was free of reference to categories such as 'brains', 'living', 'intelligent', etc., then those laws would properly be regarded as strongly emergent. Then they would be evidence against the completeness of physics, not an objection to the *physical*$_2$ account of what it is to be physical. Block's suggestion that there could be problematic boundary cases between the physical and the strongly emergent is under-motivated, and in what follows I stick with the account of physical adopted in the preceding chapter, and focus on the challenge of strong emergence.

In one straightforward sense the consequences for the Completeness Thesis of strong emergentism about the non-physical are essentially the same as those of dualistic interactionism: both are committed to the view that there are non-physical causes of physical outcomes.[29] While a positive argument for the Completeness Thesis presents a challenge to emergentism and dualism alike, though, the cases made on behalf of those doctrines, and by implication against the Completeness Thesis typically differ. Emergentism tends to be advanced as an empirical thesis specifically related to the results of certain experiments, and considerations relating to known scientific theories. Dualistic interactionism is typically defended by a combination of arguments for the immateriality of mind and arguments for the consistency of some model of the process of interaction which is less directly connected to working science than the cases made by emergentists. Direct consideration of dualism is outside the self-imposed scope of this enquiry, and in what follows I will focus on emergentism, bearing in mind that what is good for the Thesis is bad for dualism.

3.3. The Alleged Incoherence of Emergentism

Although the question of the truth of emergentism is generally taken, as is the case in the present work, to be an empirical matter, there are some attempts to argue that it is either philosophically refutable or demonstrable. Bhaskar's argument for emergentism is discussed shortly, for the time being I want to look at one attempt to argue that the doctrine is philosophically incoherent.

Peper (1926) argues that emergentism is faced with two unacceptable dilemmas, depending on whether it is formulated in terms of an account of emergent 'levels' of properties, or in terms of the nondeducibility of certain properties from those of fundamental constituents. In each case, that is (a) whether the effects at a level are the cumulative consequence of those at lower levels, or (b) whether the properties are in fact deducible from fundamental properties, Peper argues that the emergentist has a choice between endorsing the view that the

---

[29] Lowe (e.g. 1993) explicitly defends emergentism as a dualist response to the problem of mental causation.

properties in question are non cumulative or non-deducible, and having also to accept that they are epiphenomenal, or allowing that they are in fact both deducible and predictable.

What is wrong with all this from my point of view at least is clear enough: Peper's argument takes the truth of the Completeness Thesis as a premise,[30] going so far as to describe predictability as a property of 'any physical change' (1926: 241). This is required to establish the epiphenomenality of the putatively emergent properties, and is further evidenced by his worry about the 'conflict' between possible emergent causes and 'physico-chemical laws' (Peper 1926: 244). Worries about the specific conclusion Peper draws concerning epiphenomenalism need not detain me here, since they are not essential to my enquiry. Suffice it to say on this point that he ignores the possibility of any kind of ontological identity thesis, or what would later be called a functionalist account of higher level properties which at least in principle avoids epiphenomenalism without committing itself to reductionism.[31] The most important matter is that I am attempting to work out how to decide whether there is good reason to support the Thesis in the face of various challenges including emergentism, and it simply will not do to follow Peper and reject emergentism in the name of the Completeness Thesis.

More recently something like Peper's argument has resurfaced in the consideration of emergentism. Kim (1992) for example argues that emergentism is best seen as a species of nonreductive physicalism, and consequently prone to the danger of epiphenomenalism. I have remarked more than once here that Kim's analyses take the Completeness Thesis as a premise, whereas it is the acceptability of that premise in other debates which is at stake here.

Although the supposed philosophical or formal arguments against emergentism seem to fail, there is also the possibility of arguments for the doctrine. It is to this which I now turn. Before dealing with Bhaskar's argument for emergence it will help to set out the main outlines of his philosophy of science, some aspects of which are opposed to Cartwright in the discussion of laws of nature in Chapter Four below.

4. Bhaskar's Philosophy of Science

In *A Realist Theory of Science* (1975, 1978) and *The Possibility of Naturalism* (1979, 1989) Bhaskar set in motion a programme in the philosophy of natural and social science which came to be called critical realism.[32] He diagnoses a recurrent error, dubbed the 'epistemic fallacy,' (1978: 16) in which ontology is

---

30 Meehl and Sellars (1956) reject Peper's analysis for essentially the same reason, even though holding for their own part that emergentism is false on empirical grounds.

31 Some of these questions are returned to in Chapter Three below (section 6) and in the conclusions of the present work.

32 Bhaskar (1975) defended 'transcendental realism', Bhaskar (1989a) 'critical naturalism'. The two terms were subsequently combined (e.g. 1989b) to form 'critical realism'.

reduced to epistemology, in other words where questions about the real are pursued as though they were questions about our knowledge of the real. Opposing this allegedly anthropocentrist move he argues, against empiricism and positivism, that we may deduce from features of scientific practice that the world is structured, differentiated and changing. A differentiated world contains open and closed systems, where closed systems alone manifest the regularity (constant conjunction) classical empiricists assume to be ubiquitous, and which they take as the paradigm of causality. A structured (or 'stratified') world, on the other hand, is one where different orders or levels of causal power operate, not always obviously manifest in the actual flow of events.[33] Thus it is, he says, that in some cases scientists may causally intervene in nature to produce a closed system, and the consequent experimental regularity still be partly informative about nature itself. More specifically, as is explained below, Bhaskar thinks that we may conclude by means of transcendental argument from the fact of differentiation itself that the world is stratified, i.e. that physics is not complete, and furthermore that the ways in which the world is stratified are susceptible to change.

Bhaskar's overall argument is extremely ambitious and has a rich selection of targets. One feature of that argument, which may be a consequence of this extreme ambition, is that it is not always clear that particular theses Bhaskar advances are capable standing solely on the strength of the reasons given, but appears instead that, as presented, they stand only in the network of other conclusions and arguments he offers. Such 'all or nothing' situations are philosophically odious, and my conclusion in what follows is that read as an attack on the Completeness Thesis Bhaskar's argument fails. Nonetheless I contend that it is interesting in its own right, and in any event connects in significant ways with issues common to some of the other major positions in the philosophy of science which I will be considering here, notably those of Dupré and Cartwright. I also argue, against Bhaskar, that his approach to the question of the status of laws identified under experimental conditions can in fact be used to construct a *defence* of the Completeness Thesis against lines of anti-completeness arguments which are relevantly similar to his own.

Bhaskar's general purpose in *A Realist Theory of Science* is to develop a 'systematic realist account of science' which will 'provide a comprehensive alternative to … positivism' (1978:12). His starting point in *A Realist Theory of Science* is the distinction between transitive, that is socially constructed, and intransitive, that is entirely independent of social life, aspects of the world. He states that:

> Any adequate philosophy of science must find a way of grappling with this central paradox of science: that men in their social activity produce knowledge which is a social product much like any other ... [but] that knowledge is 'of' things which are not produced by men at all ... If we can imagine a world of intransitive objects without science we cannot imagine a science without intransitive objects...(1978:21-2).

---

[33] Bhaskar uses the term 'structured' or 'stratified' in two distinct ways, a point discussed below in section (5).

In the face of this alleged paradox[34], Bhaskar's starting point is to take it as given that science *does* make sense, and in particular that the activity of the modern style experiment[35] is at least sometimes epistemologically defensible. He then asks transcendentally what sort of world would we have to inhabit for this defensibility to be philosophically intelligible. More specifically Bhaskar regards his argument as a transcendental deduction from the 'possibility' of science (1978: 26). There is good reason to be cautious about transcendental arguments, since they depend heavily on the proper identification and interpretation of whatever is used as a premise, and are open to alternative analyses of either that premise itself or of the conditions needed to render it possible. Furthermore it is at least unclear what precisely is meant by the claim that science is 'possible'. I return to the question of the interpretation of this premise in section (5) below, and for the time being concentrate on Bhaskar's philosophy of science.

4.1. Experimental Activity

Experimental activity is of particular interest here precisely because, with the exception of celestial movement, almost all empirical regularities occur only in situations characterised by experimental control.[36] This is simply to say that what is above called the 'differentiation' of the world into open and closed systems just is an empirical fact.

Bhaskar contends that the constant conjunction is the empiricist/positivist paradigm of causality,[37] and that empiricism and positivism have a common commitment to what he calls 'actualism', and which he defines as the doctrine of the actuality of causal laws; that is, to the idea that laws are 'relations between events or states of affairs' (1978:64). He proposes to show the defects of this view, and the strengths of his own, by analysing what he calls the 'conditions for closure'. As noted above he calls systems in which constant conjunctions do obtain 'closed' and those in which they do not 'open'. Since actualism identifies natural laws with constant conjunctions, it presupposes closed systems. The analysis of the conditions for closure should then make clear the extent to which actualism is sustainable. In opposition to actualism Bhaskar proposes to show that, contra Hume, a constant conjunction of events is neither a sufficient nor even a necessary condition

---

34 It may certainly be a *puzzle* but it seems infelicitous to call it a paradox.

35 Bhaskar himself is not entirely clear about his target here, although the types of experiment he discusses (see section 4.1 immediately below) suggest that his paradigm examples are classic experiments by, e.g., Galileo and Boyle which isolate specific regularities or the responses of single variables under controlled conditions. Dear (1995) gives an exceptionally clear and careful account of the historical development of these types of experiment.

36 This fact has also been commented on by Nancy Cartwright (1980, 1983) and is discussed further in Chapter Four below, especially section (5.3).

37 Bhaskar makes clear that in different ways it is Hume, Mill and Kant who he has in mind here. Whether his exegeses are fair is a question which is not central to my purposes here, although failings with his reading of Mill are an important consideration in section (6) below. What is most significant for my purposes, though, is not whether Bhaskar's exegeses are accurate, but whether the positions he attacks truly permit the conclusions he draws from their failings to be established.

for a scientific law (e.g. 1975: 33-35). Given that closed systems are, at least *prima facie*, not the norm, the analysis of experimental activity just is the analysis of the special conditions for the credibility of actualism.

Bhaskar points out that 'an experiment is necessary precisely to the extent that the pattern of events forthcoming under experimental conditions would not be forthcoming without it' (1978: 33). This observation is fairly common-sensical, drawing attention to the fact that experiment would serve no purpose at all unless by engaging in experimental activity something was made to happen which would not otherwise have happened. At least one possible purpose of an experiment is to be informative, which is to say that in some sense it should 'tell us something' which but for the performance of the experiment we would not have come to know. If, however, we are to make proper sense of our commitment to regarding experimental work as scientifically informative then Bhaskar thinks we are required to accept some significant assumptions:

> It is a condition of the intelligibility of experimental activity that in an experiment the experimenter is a causal agent of a sequence of events but not of the causal law which the sequence of events enables him to identify (Bhaskar 1978: 12).

Bhaskar's claim here is that *without* the assumption which he calls a 'condition of intelligibility' we would have to regard the outcomes of experiments as either arbitrary reflections of experimental intention (so that the experimenter could generate *any* sequence of events, and the result be completely uninformative) or useless (so that the result is informative *only* about the experimental situation rather than the world at large). At this point it is worth noting that the term 'causal agent' can be read relatively weakly, indicating simply that the experimenter was not causally redundant with reference to the events which occurred in the experiment. There are places in Bhaskar's argument, though, where the notion of a causal agent is given a much stronger metaphysical reading involving real bearers of causal powers, a conception which raises a number of difficulties some of which are discussed below. If, for the time being, we take the weak sense of causal agent and regard the assumption as acceptable, that is to say if we reject the alternative possibilities that experiments are either uninformative or useless, then it follows immediately that there must be a categorical distinction between events and whatever it is that the laws of science have as their objects.

In other words: attempting to justify the minimal assumption that experimental activity is at least potentially informative about the world, and noting that the world consists of both open and closed systems where it is generally constitutive of an experimental system that it be closed, it follows that whatever the object of science is, it cannot be patterns of events. If the object of science is events, then either science can tell us nothing about most of reality, since any laws relating purely to events would apply for the most part only in

the laboratory where the patterns of events have to be worked for anyway, or science has discovered no laws at all, since no general law describing events can stand up to the existence of open systems.[38]

Bhaskar, as we know, has already decided that that experimental work *is* at least possibly informative about affairs outside the laboratory and so concludes at this point that such work tells us about the 'generative mechanisms' (Bhaskar 1978: 14) which give rise to events rather than the events themselves. The term generative mechanism is not supposed to carry any great weight at this stage, and it rather marks the 'categorical distinction' noted above, which is to say that a generative mechanism cannot itself be an event, and that generative mechanisms are those aspects of the world in virtue of which the events which do occur occur.

On the view Bhaskar is defending experimental activity is seen as the planned disruption of what is going on in some part of the world with the intention of getting the flux of events, by being empirically regular, to give some indication of the properties of a particular generative mechanism. At times he describes the function of experiment as being to bring some generative mechanism 'in phase' (1978: 35) with the flow of events so that properly interpreted the events can be seen as informative about the mechanism or mechanisms in question. There is no obvious *a priori* reason to assume that closure will always be possible, which is to say that it is reasonable to suppose that the world could include systems (or perhaps 'levels' of reality although I am wary of the implications of the notion of a *level* here) for which closure is impossible.[39]

Bhaskar goes on to develop a view of natural laws as 'transfactual' (1975: 51) ascriptions of 'tendencies' (1975: 97-100) or powers to things in the world. In a closed system a tendency will manifest itself regularly under equivalent conditions, but in open systems the manifest flow of events will typically be the product of the operation of many tendencies, some of which may be exercised without any eventual effect. Aspects of this view are discussed further in the discussion of Cartwright's notion of 'capacities' (Cartwright 1989) below.

Although Bhaskar thinks that specific enquiry into the nature of the generative mechanisms in the world is a scientific task which cannot be pre-empted by philosophy, he also maintains that the claim that there *are* generative mechanisms at all, and that the world is what he calls stratified, can be made on the basis of philosophical argument. He marks this point by drawing a distinction between what he calls *philosophical* and *scientific* ontology:

> One consequence of the argument which establishes the transfactual and non-empirical nature of laws is that a philosophical as distinct from a scientific ontology is irreducible in the philosophy of science. A

---

38 Bhaskar does not draw a distinction between fundamental and phenomenological or engineering laws in the manner of Cartwright (1980, 1983), whose approach (discussed in Chapter Four below) is to accept the 'limited validity' option as far as laws go, although her later (1989) work on capacities is closer to Bhaskar's thinking on tendencies.

39 This may be the case, for example, with Quantum Mechanics. The question of how to interpret failures of closure is discussed below in sections (5) and (6) while Quantum Mechanics is discussed in section (7.2).

> philosophical ontology will consist of some general account of the nature of the world, to the effect that it is structured and differentiated, whereas a scientific ontology will specify the structures which, according to the science of the day, it contains and the particular ways in which they are differentiated (1989b:150).

One of Bhaskar's major conclusions, then, is that as a point of philosophical ontology he regards the world as stratified in the sense that it contains physically efficacious causal kinds other than physical ones. It is in virtue of his defence of this conclusion that his position is centrally relevant to my consideration of the completeness thesis here.

While there is a great deal more which can be said about Bhaskar's transcendental realism, what has been covered here is sufficient for the purposes of engaging with the issue of the strength of his analysis of experiment as a rejection of the Completeness Thesis. The following section presents an analysis of his argument for emergence, and the section after that develops criticisms of the argument which I regard as sufficient to refute it.

5. Bhaskar's Argument for Emergence

As we will see Bhaskar is in many ways a textbook British emergentist.[40] He makes the usual concessions to the fundamental place of physical laws, takes care to make his position at least superficially compatible with the conservation of energy (1975: 109-112) and so forth. For all that, though, his approach has the noted peculiarity of having at its centre a supposedly philosophical rather than empirical argument for emergence. This argument is interesting for this reason alone, since typically the truth of emergence is taken to be an empirical matter, but also for two further reasons. On the one hand Bhaskar's argument is focussed on the question of the status of fundamental physical laws, an important issue from the perspective of the Completeness Thesis, and (as we will see in Chapter Four when Cartwright is discussed) a hotly contested one. On the other, Bhaskar's argument, or at least its conclusion, has exercised a small but significant grip on the imagination of various philosophers of social science.[41]

As noted above, in the argument of *A Realist Theory of Science* there are at least two ways in which Bhaskar uses the term 'stratified'. It will be important to distinguish them more carefully than he does for the purposes of the analysis and criticism which follows. In the thinner sense, which I will call *weak* stratification, a stratified world is one in which there is a 'real distinction' between sequences of events and the 'generative mechanisms' which give rise to the events (Bhaskar 1978: 29, 35). In the thicker sense, which I will refer to as *strong* stratification, on the other hand, a stratified world is one in which there is some *variety* in the kinds of

---

40 He is rather lax about acknowledging this, making no reference to Lewes, Alexander, Bain, Lloyd Morgan or Broad, and ignoring Mill's discussion of heteropathic laws.

41 For example Margolis (1987: 347-50), Collier (1999), Fleetwood (1998).

generative mechanisms responsible for the flow of events (1978 :169-171). In particular the notion of strong stratification applies to worlds in which causal emergence is possible, i.e. in which some kinds of system may possess causal powers which cannot fully be accounted for in terms of the powers of their constituents, and which enable the system (in at least some cases) to 'act back' on those constituents in the manner usually described as 'downwards causation'. It is, I trust, clear that these types of stratification are quite distinct, and that while an argument for the stronger version must (given the formulation in terms of generative mechanisms) entail an argument for the weaker, the reverse state of affairs need not obtain. Strong stratification is a version of what I am here calling strong emergence.

Bhaskar makes clear by implication that what I am calling a strongly stratified world would be one where physics is not complete. He considers the possibility that laboratory science could replicate the conditions under which life arose, and asks whether such a demonstration would affect the status of biology as a science or family of sciences. His answer is negative, '[f]or a knowledge of biological structures and principles would still be necessary to account for any determinate state of the physical world. Whatever is capable of producing a physical effect is real and a proper object of scientific study' (1978:113). Most revealingly of all he maintains that 'biology would not be otiose' to the extent that 'living things were capable of acting back on the materials out of which they were formed…' (1978:113). This image of 'acting back' is clearly at odds with the idea of a complete physics.

Given the distinction between weak and strong stratification it is possible to read Bhaskar's analysis of experimental activity as a two stage argument, first for weak stratification, and then for the stronger version. This involves reinterpreting and to a significant extent restructuring the argument,[42] rather than simply presenting it, but such an approach enables the relevant features of his approach as far as my interests go to be made more clear. I consider each thus separated argument in turn.

5.1. Weak stratification

Bhaskar's argument for weak stratification has already been described briefly in the overview above. In more detail, though, the argument (Bhaskar 1978: 63-69) runs as follows:

(W1) The world is differentiated.

Given that 'differentiation' is Bhaskar's term for the fact that not all systems are at least apparently closed, this is just another way of saying that closed systems are not ubiquitous. This premise is relatively uncontentious, perhaps especially so, since Bhaskar allows that the way the 'system' is specified in any given case is necessarily relative to some description (1978: 73). This premise then amounts to the claim that under

---

[42] I have done this in a similar way elsewhere, see Spurrett (1997).

some descriptions some events do not fall into relationships of constant conjunction. The characterisations of the events themselves and the ways in which the individuals comprising the system are individuated, are to be regarded as similarly relative to description (1978: 14, 33-36).

A closed system, then, will simply be one in which events of type **A** are regularly succeeded by events of type **B**, under some protocol for describing **A**s, **B**s and the boundary of the system. Bhaskar frames his argument in terms of constant conjunctions, and hence has in mind the eventual rejection of regularity determinism, which he takes as the ontological requirement for the plausibility of the actualist account of laws. It is possible, though, to interpret the notion of closure in probabilistic terms, without substantially changing the way the argument works with respect to the issue of stratification. In such a case we could say that a closed system is one where, under suitable descriptions, events of some type follow events of some other type with a known and definite frequency. Nothing important for my argument will hang on this question, though, and I will continue to present Bhaskar's argument in terms of constant conjunctions.

(W2) True causal laws have general validity.

From Bhaskar's point of view this premise is something which needs to be sustained to do justice to the ways in which experiments are performed, and to our sense that experiments are informative. That is to say that our concept of laws should be such as to make it possible to think that what is found in the laboratory obtains outside the laboratory too.

There are, it must be noted, philosophical positions which would contest this premise. The discussion of Cartwright in Chapter Four below raises some alternative interpretations of the significance and status of experimental regularities, and especially of the fact that such regularities are not generally observed outside experimental contexts. For the time being, though, I concede Bhaskar's second premise without argument. It is worth bearing in mind that Bhaskar's notion of the informativeness of experiment has to do with his view that in experiments it is possible to make progress towards discovering *laws*.

So closed systems for Bhaskar and the constant conjunction theorist alike are important sources of evidence for the claim that some law holds between regularly conjoined events. Bhaskar, though, takes the fact that closed systems are not ubiquitous to point to something important about the nature of the laws which are discovered in them.

(W3) So laws must relate to something other than events, i.e. the world is weakly stratified.

This follows from the first two premises. If laws are to be generally valid, and some systems are open, then laws cannot be statements of empirical regularities.

In Bhaskar's vocabulary the claim here is that if we want to sustain the universal character of the laws discovered in the laboratory, then we should accept, given the fact of differentiation, that a 'sequence of events

can only function as a criterion for a law if the latter is ontologically irreducible to the former' (1978:65). Making the same point slightly differently he says 'if there is a real distinction between open systems and closed and our intuitions about the rationality of science are to be preserved there must be a real distinction between structures and events. In this respect the differentiation of phenomena still provides the best argument for the stratification of the world' (1978: 66-7).

More precisely Bhaskar concludes that what laws have as their objects are the generative mechanisms which are the sources of events. Subject to a restrained reading of 'generative mechanism' as 'those features of the world in virtue of which the events which occur do occur' Bhaskar's conclusion follows from his premises.[43]

Bhaskar's case for the generative mechanism view of laws is driven by the attempt to uphold our apparent ability to explain in cases where we may be unable to predict, and he notes that:

> Leaving aside astronomy, it is only under conditions that are experimentally produced and controlled that a closure, and hence a constant conjunction of events, is possible. The empiricist is now caught in a terrible dilemma: for to the extent that the antecedents of law-like statements are instantiated in open systems, he must sacrifice either the universal character or the empirical status of laws. If on the other hand, he attempts to avoid the dilemma by restricting the application of laws to closed systems ... he is faced with the embarrassing question of what governs phenomena in open systems. (1978:65)

Finally, Bhaskar contends that something like the notion of weak stratification is essential to sustain any attempt to distinguish between accidental or uninformative sequences of events, and necessary or scientifically useful ones (1978: 46). In a world consisting entirely of events all sequences of events are in some sense equally relevant to science, in a world of generative mechanisms only a subset of such sequences are likely to be empirically useful: those ones in which the flow of events is 'in phase' with some mechanism or set of mechanisms.

It is important to note at this point that there is no *prima facie* reason why weak stratification could not be entirely compatible with the completeness of physics. If the generative mechanisms which give rise to physical phenomena are themselves exclusively physical then physics will be complete even in the total absence of closed systems. It is of course a separate matter to argue for this possibility, a task to which I turn later, but for the time being all I wish to do is *note* it, and the consequence that there is no need for me to contest the

---

[43] Chalmers (1987: 82) argues that Bhaskar's realism can be presented without the distinction between laws and constant conjunctions, subject to regarding laws as conditional constant conjunctions, that is as regularities only manifest in particular types of situation. I am not convinced by this proposal, since it reduces the force of Bhaskar's argument that there is a way to regard laws as operating even in the absence of constant conjunctions, and there is no independent reason to expect that laws should be regarded as relating events, or taking the form of statements of event regularities. Since my main interest here is in the argument for emergence, or strong stratification, I will proceed without getting into this point, but return to the question of the status of laws in Chapter Four below.

claim that the world is weakly stratified. For the time being I want to consider Bhaskar's argument for strong stratification.

5.2. Strong Stratification

Bhaskar's argument for strong stratification can be seen as continuing directly from that for weak stratification. As noted already, the boundaries and constituents of the 'system' which may be either closed or open and the criteria used to identify the individuals which the system comprises, are matters of description. It follows that it is quite possible that there could be those who accept the differentiation of the world contingently, i.e. as related to some classes of description, but who further maintain that relative to some other real or possible set of descriptions, the world is genuinely closed.[44] Bhaskar calls the view that there is some form of description in terms of which closure is ubiquitous 'strong actualism', a view which he argues entails the rejection of strong stratification.

(S1) Strong actualism assumes the ubiquity of closure relative to at least one form of description.

'Strong actualism' is simply Bhaskar's label for the extreme form of the view that the laws discovered by science have events and relationships between events as their objects. To have a hope of succeeding it is clear that the strong actualist programme is, in the limit, committed to the view that relative to *some* description every system is, or can be, in principle, closed.

Strong actualism is not equivalent to nor does it automatically entail the Completeness Thesis, since as it stands it is compatible with any combination of event types (including mental, economic, biological ones) being regularly related. (I return to this point in section (6) below.)

(S2) The world is weakly stratified.

Although, as noted, Bhaskar seems at times to equivocate two distinct sense of stratification, when his argument is reformulated into two separate strands, that for strong stratification relies on that for the weak version.

(S3) Confronted with open systems the strong actualist may engage in various strategies for redescription and investigation so as to restore closure.

In the face of an apparent failure of closure anyone who endorses strong actualism, and is hence committed to the view that the failure *must* ultimately be apparent only, may embark upon various courses of action in the hope that the expected regularity can be made manifest. Each programme involves a kind of redescription (assuming that the failure lies with our knowledge rather than the world) dedicated to making

---

[44] There are various examples of views of this sort. Oppenheim and Putnam's account of the Unity of Science discussed in Chapter Three below (especially section 2.1) is one, while the view Cartwright (e.g. 1994) calls 'fundamentalism' may be another (see Chapter Four sections 2 and 3).

manifest the regularity which it is assumed must exist. These redescriptions will typically refer to available data and theory, and may lead to the formulation of new experimental procedures to decide between the description under which the system is apparently open and the description where closure is restored.[45] Bhaskar maintains that an analysis of the redescriptive programmes produced in this way reveals the limit conditions for closure in the sense of conditions which, if satisfied, would have to produce closure, or the strong actualist position thereby be refuted (1978: 70). That is to say that if there was some set of conditions the meeting of which would result in constant conjunctions, the failure of a conjunction to manifest in such conditions would be a counter to the constant conjunction view of laws.

Let us say, for example, that two different philosophers when placed under the same conditions do not respond in the same ways, a moderately plausible scenario. There are, Bhaskar argues, two principal strategies for redescribing this or any other apparently open system so as to restore closure. One approach is to say that the conditions were not in fact the same. On this view some interfering factor may have been responsible for the apparent failure of regularity, and the description of the system needs to be expanded to include it.[46] Another approach (the two are not exclusive, and can be deployed in concert) is to say that the two philosophers were not in fact the same for causal purposes, but that some internal difference is behind their different responses. In this case the level of description which picks out 'philosophers' may need to be replaced by one at the level of the *constituents* of philosophers, in the hope that such a description will be more regular in virtue of not taking for granted the previously un-analysed internal states of the lovers of wisdom. Both approaches can be repeated, or iterated, which is why Bhaskar refers to them as 'regresses' calling the first 'interactionism' and the second 'reductionism' (1978: 74-79).

The functioning of the interactionist and reductionist regresses requires, Bhaskar argues, a further assumption: that the systems in question have additive principles of organisation. The properties of a system *with* additive principles of organisation are exhausted by the properties of the individuals which the system contains. The requirement that this exhaustion be possible is the requirement that full causal reduction be possible or conversely that real emergence, where the properties of the system are *not* exhausted by those of any set of individuals comprising the system, be impossible.[47]

Bhaskar's discussion of the interactionist and reductionist regresses distinguishes between 'dominant' and 'recessive' criteria (1978: 76), where the recessive version calls only for possible interferences to remain

---

45 The search for determinism-restoring additional variables in Quantum Mechanics (e.g. Bohm 1952) would be an example of this.

46 There is room for variation here, since the external influence can be included, or its operation can be prevented in some way. For the purposes of the argument here there is no important distinction between these cases, and in the practice of science both are used. Cartwright discusses both in (1989: 39-90).

47 I argue shortly that Bhaskar is incorrect about this, and hence that his argument relies upon a spurious dichotomy.

stable while the system is being observed, whereas the dominant case insists that interference be completely *absent*. In the present narrowly ontological discussion I consider only the dominant case.

In the limit the first 'interactionist' regress requires that the entire universe be included in the antecedent state-description. This appeal to the incompleteness of the description of the system always allows the actualist to reject criticism directed at the failure of the model to provide accurate predictions by reference to practical difficulty, although such a response potentially leaves the actualist on empirically shaky ground. The second regress has distinct but similar effects. In the limit it requires that regularities be thought of as obtaining only between true atoms, that is between individuals with no internal structure and consequently with no possibility for change.

If all three of these conditions *are* met, then it should be clear that in the event strong actualism was true, that fact would be manifest. Bhaskar's contention, though, is that it is impossible that these conditions *could* be met, given certain features of our world.[48]

(S4) 'Natural possibility statements' are impossible for strong actualism.

'Natural possibility statements' is a Bhaskarian term of art. What Bhaskar means by a natural possibility statement is, for example, the claim, made of an aeroplane presently in a hangar, that it 'can' fly at 600 m.p.h. (1978: 87). His claim, then, is that the strong actualist cannot account for the distinction between what things do, and what they can do (1978: 78, 87). (I note in passing that Bhaskar's claim here is strikingly similar to some made by Dupré, discussed in Chapter Three section (3.1) below, as well as some made by Cartwright, discussed in Chapter Four.)

This premise is, to say the least, contentious. The interpretation of propositions including terms like 'can' and 'could' is notoriously vexed, and lies at the heart of some of the debates over the possible compatibility of freedom and determinism. Bhaskar is at least ignoring the possibility of an actualist account of such claims in terms counterfactuals, which is to say that there is nothing to prevent the actualist saying of any outcome that *were the conditions different* the outcome could have been different too. A 'could' claim which did not imply the truth of at least one 'would' claim would be likely to be false itself, so an exegesis of 'could' along the lines of 'would … if …' has considerable initial plausibility: were the Boeing fuelled, controlled by a pilot, given clearance for take-off, and properly serviced, say, it would fly.

Bhaskar's defence for (S4) is as follows:

---

[48] Bhaskar's argument is supplemented (1978: 79-90) by a discussion of what he calls the 'Classical Paradigm of Action', which is a further articulation of the limit conditions for closure. Since I argue shortly that what he has said already is needlessly extreme and misguided there is no need to follow him down this road. See Spurrett (1997) for a more detailed discussion, which comes to conclusions the shortcomings of which I detail shortly.

(S4.1) For natural possibility statements to be possible we need a view of the world as containing agents with various powers and capacities, so as to permit situations of dual and multiple control.

Bhaskar does not maintain that natural possibility statements involve any departure from causality, but rather that the key notions are those of an unexercised power, and a power exercised either without effect or with modified effect (Bhaskar 1978: 87f, 97f). He contends that for such talk to make sense we need to suppose that agents possess a collection of powers, and that it be possible for some of these powers to override others, or to trigger or inhibit their exercise. He states that in situations of 'dual and multiple control' the world or system will be open at the level of at least one of the kinds of cause in operation (1978: 112).

Stripped of the jargon and neologisms a portion of this makes sense, and is defensible. The claim that if more than one causal power is in play in any given system, that system will not behave as expected from the point of view of either power taken in isolation is not contentious. Cartwright makes a similar point, although to a different purpose, in her discussion of the behaviour of bodies which are both massive and charged.[49] This, though, is not all that Bhaskar calls for, and his metaphysically loaded talk of agents (Bhaskar 1978: 105, 108, 110) is both unclear and under-motivated.

(S4.2) This would require strong stratification, i.e. the failure of reductionism.

This is one of the two most contentious premises in Bhaskar's argument. I argue shortly that it is false, but note here that Bhaskar explicitly refers to the metaphysical view necessary for strong stratification as being a kind of 'emergence' which permits causal 'novelty', that is the operation of kinds of causes not reducible to those at lower levels (1978: 72, 109). It is also worth noting that Bhaskar's account of possible reductionist (1978: 114-5) theses is extremely selective, leaving out, *inter alia*, micro-reduction and ontological reduction.[50]

(S5) If reductionism were true science would not be possible, since experiment presupposes agents with powers and capacities (1978: 114, 116).

As well as requiring weak stratification, given the differentiation of the world, Bhaskar holds that experimental work requires that experimenters think of themselves as agents in the sense of (S4.1) above. He specifically contends that the possibility of science requires the notion of a 'law-governed world independent of man' (1978: 26) and that scientists have the power of 'bringing about effects which but for their action would not have been realised' (1978: 239).

---

49 Cartwright (1980: 77-80). See also Chapter Four below, especially section (5.4).

50 Bhaskar's shortcoming here is shared by Dupré (1983, 1993) who also considers a drastically limited range of reductionist theses. See Chapter Three below for more detailed discussions of reductionism, especially sections (3.1) and (6).

For Bhaskar the possibility of science involves the informativeness of experimental work as required for premise (W2) in the argument for weak stratification above, i.e. such that laws discovered by science are at least possibly universally valid.

Even allowing, for the moment, Bhaskar's account of reductionism in terms of which strong actualism requires the truth of a form of reductionism which permits prediction of what is reduced on the basis of what it is reduced to, the present point is questionable. Marx asserted that 'all science would be superfluous if the outward appearances and essences of things directly coincided' (1972: 817). This seems reasonable enough, but it is not the same thing as what Bhaskar is saying here, which is that unless some inflated account of agency[51] was true science would be not merely unnecessary but impossible. The chief danger of this step in Bhaskar's argument is the suppressed premise that the only two available options are the restrictive kind of reductionism he considers, and his own inflated view of agency.

(S6) But science is possible, so the world is strongly stratified.

A strongly stratified world is one where physics cannot be complete, one where we are 'unable to predict future physical states of the world without referring to comprehensive entities and principles of behaviour specific to them' (1978: 114).

If this argument works, or can be repaired, then it counts significantly against the Completeness Thesis. In what follows I discuss a number of detailed criticisms and possible Bhaskarian responses, then develop an argument for the thesis that all of the phenomena of a strongly stratified world are compatible with the view that physics is complete. This second argument does not, of course, show that physics is indeed complete, and in the final element in my criticism of emergentism I ask, what the empirical credentials of strong emergentism are.

6. Criticism of Bhaskar

Bhaskar's argument for strong stratification presents no shortage of critical targets. His case for weak stratification, though, has a lot to recommend it, and is in crucial respects the same as a thesis advanced by Cartwright to the effect that what she calls the 'facticity' view of laws is unsupportable, and hence that true causal claims cannot be about event regularities (Cartwright 1980, 1983).[52] The strength of the case for weak stratification presents no significant difficulty for my purposes, though, since as noted that view is entirely compatible with the Completeness Thesis. What can be said against the argument for strong stratification?

---

[51] Bhaskar's views on agency are similar to some of those of Dupré on free will, discussed in Chapter Three below, especially section (3.3).

[52] See Chapter Four below, especially sections (3) and (5).

To begin at the most general level, there is no reason to accept an argument which has the form '**A** is untenable, therefore the *opposite* of **A**' without some reason to think that these are the only options. In Chapter Three below we see how Dupré (1993) argues from the failings of a highly attenuated form of reductionism to the defence of an equally extreme position he calls 'promiscuous realism'. Bhaskar's case and Dupré's alike stand in need of some justification for the premise that the alternatives which they consider are the only ones available, and in both cases I maintain that the suppressed premise is in any event false.

Matters are made worse by the fact that the target at which Bhaskar directs his fire is a complicated conglomerate of extreme theses which I am not convinced even forms a candidate for a viable position. (It does not save Bhaskar that he himself maintains that the strong actualist position is ultimately untenable, since the issue here is whether his argument relies appropriately on the contrast with the position he has manufactured.) So, for example, Bhaskar's paradigm strong actualist is supposedly committed to a certain form of reductionism incompatible with strong stratification. Bhaskar's argument is that strong actualism implies the reductionist denial of strong stratification, but that the required kind of reductionism is incredible (I return to this contention shortly), a fact which reflects badly on strong actualism. This, though, just cannot be right, at least not if Mill's philosophical position discussed above (e.g. section 2.1) is coherent, since Mill is both a strong actualist, a point which Bhaskar makes frequently enough (e.g. 1978: 70), and also an emergentist, or a proponent of strong stratification, a fact of which Bhaskar seems on the textual evidence to be entirely unaware.[53] Bhaskar's argument, then, is established by means of a contrast with a position which simply fails to have the consequences which he claims to detect.

Mill's case is doubly damning, since it both shows the untenability of Bhaskar's claim that the strong actualist must be anti-emergentist, and also gives the lie to the suppressed premise that the only options worth considering here are strong actualism and his brand of emergentism. This alone is reason to reject Bhaskar's argument, and move directly to the consideration of the empirical credentials of strong emergentism, but there is more to be said.

A further important criticism, hinted at already, is that Bhaskar's taxonomy of reductionist theses is importantly incomplete. In fact it is limited in the same type of way as Dupré's (discussed below in Chapter Three), which is to say that they both leave out forms of reductionism which are immune to their own criticisms. In Bhaskar's case the key demand he makes of reductionism is that it take a form which implies that predictions of the behaviour of the objects described by the reduced science are in general possible by the reducing science. This is a extraordinarily strong demand, for which Bhaskar gives no significant justification. Not even archetypal reductionists such as Nagel (1961) and Oppenheim and Putnam (1958) go so far as to

---

[53] This is, to say the least, somewhat remarkable, given that Bhaskar does have things to say about Mill's views concerning the composition of forces (e.g. 1978: 99).

maintain that successful reduction will permit prediction in all cases, and the Oppenheim and Putnam approach to reduction in terms of micro-reduction explicitly enables, if defensible, the rejection of strong stratification without commitment to the general possibility of prediction.

Further considerations arising from chaos theory make clear that determinacy need not entail predictability. In the cases of both theories and real systems the distinguishing feature of the chaotic is *dynamical*. Specifically the systems in question have aperiodic, i.e. non-repeating, time development properties. (Thus the pithy accuracy of the title of Lorenz's (1963) groundbreaking paper: *Deterministic Nonperiodic Flow*.) While aperiodic, the systems have finite possibilities of action. So, if every dynamical aspect of a system is assigned to a variable and the set of *n* variables allowed to define a multi-dimensional 'phase space' in the usual manner the infinite series of states realised by the system will fall entirely within some finite volume of *n* dimensional phase space. The only way for an aperiodic (non-repeating), and therefore in the limit infinitely long, time development to fit into a finite space of options is for it to be 'fractal', i.e. infinitely finely structured, since if at any point the graph of the successive states meets up with an earlier part of itself then *by definition* the system is periodic. Consequently, for chaotic systems finite volumes of phase space may be passed through by an infinite number of segments of the phase graph. Lorenz made this very point by concluding, with reference to the phase portrait of his artificial weather system, that 'there is an infinite complex of surfaces, each extremely close to one or the other of two merging surfaces' (1963: 140).

It is this remarkable fine-structure which creates the well known problems of prediction of the behaviour of chaotic systems, since any attempted measurement of a real system described by such a dynamical model cannot fail to fall foul of the impossibility of adequately (that is with absolute precision) establishing the actual values of the relevant variables, and consequently producing inaccurate predictions. (No finite digital computer can store the values of any continuously varying quantity with the precision required for predictive success.) The actual development of the system and the development predicted on the basis of variables with necessarily limited accuracy may be wildly divergent, so that a tiny initial difference may entail enormous differences in outcome. The celebrated 'butterfly effect' is no more than the recognition that even such apparently irrelevant factors as the flapping of a butterfly's wings in Brazil may not be considered so small as to be of no consequence in predicting the likelihood of a hurricane in Tokyo. The very accident that led to Lorenz's discovery of aperiodic determinism was consequent upon an attempt by him to repeat a weather simulation with figures rounded off to three decimal places, instead of the six to which they were computed. The evolution of the second simulation diverged wildly from that of the first because in Lorenz's non-linear simulation no difference was too small to be irrelevant.

So a chaotic theory is a deterministic one which has the formal property of aperiodic dynamical evolution. That such dynamical systems exist shows that the earlier belief that only stable and periodic

dynamics were possible was mistaken, and that real deterministic systems need not be obviously deterministic, since in some cases the fact of determinism is compatible with apparently random behaviour. *Prima facie* this presents a challenge to Bhaskar's interpretation of differentiation, since here was have systems which are deterministic, and hence closed by his own definition, but unpredictable, and therefore also open by his definition. To this the transcendental realist may respond that there is in fact no difficulty here, since even without predictive success in all cases, the analysis of chaotic systems conforms to the requirements of Bhaskar's discussion of the conditions for closure set out above.

The reasoning here would be that whatever supports the assertion that chaotic systems are deterministic, it is not predictive success. The reason seems rather to be that for chaotic systems it is believed that the conditions for closure are satisfied, and therefore that the systems are closed, and hence repeatable in principle, even though they appear to be stochastic. Certainly chaotic *models* are quite clearly closed, in the sense that chaotic simulations run from identical starting values (which can easily be done) produce identical results. Thus when Lorenz's false weather was run with values which were not rounded down, the original results were exactly reproduced. So, what we can now call chaotic systems, even though they are not superficially closed, do not really count as a direct objection to Bhaskar's analysis of closure and the context of experiment.

The point at stake here, though, is that chaotic and related systems are both entirely compatible with the truth of the Completeness Thesis, and hence incompatible with strong emergence, but also open under significant descriptions, and hence a challenge to any argument which, like Bhaskar's, attempts to draw a conclusion about the real dynamical properties of the world from particular features of our knowledge of it. If a computer calculating the evolution of some set of dynamical equations cannot be used to predict what another computer running the same sums, but storing values to one more decimal place, will do then the burden of proof is not where Bhaskar assumes. It is now squarely on anyone who wants to argue from failures of predictability to some ontological conclusion, since at least some cases of unpredictability are entirely innocuous as far as the stratification question goes: What can Bhaskar say to the critic who demands that he show that this state of affairs does not obtain in all cases?

In Chapter Three below, in connection with Dupré, I show by reference to research on cellular automata how the possibility of all manner of large-scale descriptions and taxonomies does not in any way count against the possibility of the systems so described being causally closed at some fundamental level. That argument complements the present one, and also points to the final, and perhaps most damning, criticism of Bhaskar's argument from my point of view, which is that weak stratification can account for everything to which he refers in his argument for strong stratification.

Given the notion of a generative mechanism (read in terms of tendencies, causal powers or even Cartwrightian (1989) capacities) and also the possibility that more than one such mechanism can be in play in

any given system, it is possible to make sense of the distinction between open and closed systems. Not only does this endorsement of weak stratification already involve the denial of strong actualism, again refuting Bhaskar's assumption that there are only two viable options available, it also enables the distinction between accidental and necessary sequences of events to be drawn, and a rationale to be given to experimental science. Furthermore, since it does not follow necessarily from weak stratification that prediction is in general possible, Bhaskar's dominant contrast between reductionism and strong stratification is defused. Prediction is not automatically possible for the simple reason that it is contingent whether, even given knowledge of the laws describing the operation of the generative mechanisms at work in a given system, we are capable of doing the required sums, or of isolating the system from possible external influences or accounting for the effect of those influences.[54] Bhaskar has simply not made a case for the thesis that the stronger form of stratification, and his inflated account of agency, are needed to accomplish the things which can manifestly be accomplished with weak stratification, so his argument is simply incomplete.

I conclude that the failings of Bhaskar's case for strong stratification, or strong emergentism, are such that the argument fails dramatically, and is left beyond repair: It relies on unsupported and untenable contrasts, undermotivates key premises, and can be refuted by counterexamples at too many significant points. This does not make emergentism false, any more than Mill's account of heteropathic laws makes it true: Whether emergentism is supportable or not remains an empirical question.

7. Further Reflections on Evidence

If the preceding arguments are correct then the current situation is that emergentism has been found to be a coherent doctrine which may turn out to be false. What I want to do in the present section is argue that false is exactly how, on balance of evidence, it turns out to be. The key evidence in question, I argue, arises in nineteenth century biology and chemistry, and to a lesser extent with quantum mechanics in the early twentieth century.[55] Following a discussion of this evidence and how it favours the Completeness Thesis over emergentism, I look at a few very contemporary treatments of emergence, so see whether there are any new lines of defence which merit consideration.

7.1. The nineteenth and early twentieth century

The nineteenth century biological research to which I refer is typically regarded as part of a dispute between mechanists and vitalists. This is not quite the same thing as a dispute between defenders of the Completeness Thesis and strong emergentists, since mechanism is not equivalent to the Thesis, but is rather a

---

54 This point is discussed further in connection with Cartwright in Chapter Four sections (5.2) and (5.4).

55 This places my interpretation of the evidence partly at odds with McLaughlin (1992) who places far more emphasis on quantum mechanics. Quantum mechanics is discussed towards the end of section (7.1) and in section (7.2).

very restricted version of it which is committed to the view that all causality is mechanical.[56] Vitalism is also not automatically equivalent to emergentism, since vitalism by itself requires no commitment to the hierarchical form of strong emergentism, even though both entail the denial of the Completeness Thesis. Before discussing the research in question, then, a little background may be useful. From the perspective of earlier forms of mechanism, especially seventeenth century mechanism, the question of emergence hardly arises. This is because, whether plenist or corpuscularian, most mechanists were agreed that material action took place by impact, and that the 'quantity of motion'[57] was conserved. So Descartes, except for the matter of his commitment to dualism, set out a causally closed physics in the *Principles of Philosophy,* where he says of his law of collision that:

> All the individual causes of the changes which occur in [the motion of] bodies are included under this third law, or at least those causes which are physical; for I am not here enquiring into what kind of power the minds of men or Angels may perhaps have to move bodies; I am reserving that matter for a treatise *on man* (1991: 62).[58]

Whether or not Descartes' physics was supposed to be causally closed, Leibniz took pains both to suggest that it wasn't, and that a reason to prefer Leibniz's own approach was that it was so closed. As noted in Chapter One above (section 2) Leibniz's approach can be seen as holding two Completeness Theses to be true: one for the physical and one for the mental. In his *Monadology* Leibniz argued as follows, essentially berating Descartes in the name of the Completeness Thesis:

> Descartes recognised that souls cannot impart force to bodies because there is always the same quantity of force in matter. However, he thought that the soul could change the direction of bodies. But that is because the law of nature, which also affirms the conservation of the same total direction in matter, was not known at that time. If he had known it, he would have hit upon my system of pre-established harmony (1989:223).

In time, though, mechanism gave way to a Newtonian approach to physics and mechanics, which was framed in terms of forces. Early reception to Newton was partly characterised by mechanist derision over his

---

56 Some remarks made by Broad make clear how mechanism is more restricted than the Completeness Thesis, at least in the minds of emergentists and vitalists: 'On a purely mechanical theory all the apparently different kinds of matter would be made of the same stuff. They would differ only in the number, arrangement and movements of their constituent particles. And their apparently different kinds of behaviour would not be ultimately different. For they would all be deducible by a single simple principle of composition from the mutual influences of the particles taken by pairs; and these mutual influences would all obey a single law which is quite independent of the configuration and surroundings in which the particles happen to find themselves' (1925:45-6).

57 This is not generally what we would recognise as momentum, i.e. the product of mass and velocity, which is a concept that can only properly be formulated after Newton.

58 It is of historical interest that in the course of many protracted discussions concerning the question of interaction, Descartes never once explicitly allowed any break with the laws of motion which he saw as holding for normal physical causation.

talk of forces, which many of them saw as an unwelcome regression to explanation by reference to occult powers.[59] In one sense the worry voiced by those mechanist objectors was well founded: Newtonian physics did open the doors to something potentially similar to explanation by occult powers, and that was the possibility of other kinds of force including non-physical ones operating in the world. Newton had, after all, given the law for the action of only one force, gravity, and had also himself established that the cohesion of bodies could not possibly be explained by reference to gravity alone (which is far too weak) and hence that it was likely that there were other forces to be found. In fact a reference to the problem of short range attraction between the parts of cohesive bodies is mentioned pointedly at the very end of his *Principia Mathematica* (1934: 547).

The call for further forces was answered with enthusiasm, and the variety of processes for which specific forces were proposed and after which enquiries were made was large: additional forces of attraction and repulsion for known phenomena such as electrostatics, magnetism and the cohesion of bodies; forces of irritability and sensibility[60] to account for perception and more general responses of living things to changes around them; forces to explain fermentation,[61] the origin of micro-organisms,[62] chemical bonding and more besides.

A fair summary of the empirical fortunes of these various proposals is to say that the physical forces, in some form or another, found a place in later science and that the non-physical forces did not. In the late eighteenth century (c1785) Coulomb successfully established the inverse square law by which charges attract and repel one another and did so similarly for forces due to magnetic polarity (Coulomb in Shamos 1987: 59-77). By the early 1830s Faraday had published his major research on electromagnetic induction, which also showed the unity of apparently different sorts of electricity whether electrostatic, induced or from batteries (Faraday 1952: 265-360, Faraday in Shamos 1987:128-158). In the 1840s Joule did pioneering and important research showing the quantitative equivalence between heat and mechanical work (Harman 1982: 35-41, Shamos 1987: 166-183). Joule's work, and to some extent that of Faraday, is perhaps especially important for my purposes because it integrates what until then had been treated as relatively distinct and independent phenomena, showing ways in which they are interrelated and also extending and refining the concepts which

---

59 See, e.g., Hall (1981: 310f) on the resistance of Huygens and Fontanelle, both on broadly Cartesian grounds, and Westfall (1977: 155f) especially on Leibniz. Newton was not admitting occult powers, of course. He urged the distinction between explanation by force and by occult power at various points, notably including Query 31 of his *Optics* (Newton 1951: 401f).

60 These forces were proposed by von Haller, see Singer (1959: 370-372) and Coleman (1977: 17-21).

61 Pasteur maintained that fermentation was a vital rather than chemical process, see Coleman (1977: 137) and Singer (1959: 444-446).

62 Singer reports that Leeuwenhoek and other early microscopists had recourse to talk of additional forces or causal principles to account for the appearance of micro-organisms visible under the microscope in apparently clear samples of water, and that microscopic biology initially favoured rather than disputed the Aristotelean notion of spontaneous generation (1959: 437-444). See also Hall (1981: 186f).

led to the principle of the conservation of energy.[63] The discovery that apparently different classes of phenomena and process were related was a crucial stimulus for the research which went on to show that the types of energy and causation operating in living systems were not different in kind to those operating outside them:

> The steam engine daily demonstrated the conversion of heat into mechanical work. Current electricity, produced by chemical processes in the newly invented Voltaic pile, gave rise to heat and light. That same electricity, when suitably administered, brought about chemical dissociation. The interconvertibility of electricity and magnetism was demonstrated as was that of mechanical work and magnetism. These phenomena strongly suggested that most and presumably all physical processes were convertible into one another and many surmised that that basis (energy) might be expressible in precise numerical terms (Coleman 1987: 122).[64]

It is important to bear in mind that the principle of the conservation of energy was investigated and developed in the biological researches which followed, rather than simply taken as a premise. Certainly in one sense the conservation of at least certain forms of energy falls out of Laplace's mechanics, but the very issue at stake here is whether *other* forms of energy conform to the same law, and particularly whether the law holds at all in biological cases. A series of research programmes took this question forward, an especially notable one of them treating respiration on the analogy of combustion.

A crucial figure in the early part of this story was Lavoisier, himself at the forefront of work on the role of oxygen in combustion. Some of Boyle's early air-pump experiments had shown that air was essential to life and flame alike, but Lavoisier went much further than that. In conjunction with Laplace he designed the ice calorimeter, which permitted measurements of heat produced against carbon dioxide emitted by a living creature inside the calorimeter (Lavoisier 1952: 99-103, Coleman 1987: 123-127). Lavoisier concluded that respiration was best understood as a form of combustion, obeying the same conservation constraints.

Intervening enquiries showed that muscular action also involved the consumption of oxygen and the emission of carbon dioxide, that is that muscular action was clearly related to respiration. Notable parts of this work were done by Leibig and Pflüger (Coleman 1987: 128-130, Singer 1959: 377-379, 408-411). Leibig mistakenly built his theory around the expectation that matter rather than energy is conserved in respiration, a

---

63 Prior conservation concepts were applied to matter and quantity of motion (Descartes), matter and momentum (Newton and Leibniz), and matter (Dalton). The modern notion of the conservation of energy was to a significant extent developed in the work of Joule and von Helmholtz. By 1849 the topic was well enough established for Faraday to devote his Royal Institute Christmas lectures (to an audience of children) to combustion and respiration (Faraday 1861).

64 The discussion of the history of biological research which follows is indebted to Coleman's account of the issues, on which I have been unable to improve significantly.

view which was not conclusively refuted until the work of Frankland, who performed detailed experiments establishing the energy gained from the consumption of specific foodstuffs (Coleman 1987: 135-138)

Further research required different tools, and the respiratory calorimeter was the next crucial innovation. Introduced by Regnault and Rieset in 1849 (Coleman 1987: 139) it enabled highly accurate measurements of the consumption and emission of gases under controlled conditions, which data could be correlated with other information about food consumption and excretion.[65] In a series of detailed and rigorous experiments conducted between 1889 and 1894 Rubner combined the ice and respiratory calorimeters to establish once and for all the applicability of the principle of conservation of energy to biological systems. His results also established that energy of known types was conserved, that is to say that the sole sources of energy for the mechanical work done by the living subjects within the calorimeter were ultimately chemical, and his own conclusion, following some investigations lasting 45 days, was that:

> Not a single isolated datum chosen at will out of all of these experimental results can leave us in any doubt that the exclusive source of heat in warm-blooded animals is to be sought in the liberation of forces from the energy supply of the nutritive materials (in Coleman 1987: 142).

Rubner's work collaborated earlier proposals by Helmholtz, who had in turn been influenced by the work of Liebig, and performed various researches on the inter-translatability of different kinds of energy similar to those conducted by Faraday (Harman 1982: 40-44, 60-63). One of Helmholtz's notable contributions was his 1847 essay on the conservation of force (*Über die Erhaltung der Kraft*) which developed the mathematical apparatus for rigorously expressing conservation relations by expressing the translations and conversions of different kinds of energy into a single framework.[66]

This research on the conservation of known types of energy was complemented by enquiries which refuted various claims on behalf of peculiarly biological causal principles to explain a wide range of phenomena. In 1897 Buchner successfully isolated what we would now call an enzyme from yeast, which enzyme promoted fermentation in the absence of any cells: Pasteur was wrong, the theory of the chemical catalyst vindicated (Coleman 1987: 137f).[67] In 1668 Redi's investigations had discredited any talk of sponteneous generation with respect to maggots by observing putrefecation under suitably controlled conditions (Hall 1981: 186) and the experiments of Spallanzani in 1767 with rigorously isolated samples of previously boiled vegetable material showed the inadequacy of Needham's (1748) experiments apparently showing that

---

65 The respiratory calorimeter (or its descendants) are in regular use today. A biologist colleague Chris Lotz (pers. comm. 1999) is presently pursuing research on the metabolism and energy efficiency of passeriform sunbirds, for which research the calorimeter is indispensable.

66 Helmholtz's 'Kraft' is neither specifically 'force' nor 'energy', a distinction which was really only made precise in the course of the research presently under discussion.

67 Schlick (1953: 524) contends that the laboratory production of urea by Wöhler in 1828 'refuted once and for all the doctrine that the synthesis of organic compounds requires a special force.'

micro-organisms did arise spontaneously in boiled and isolated samples (Hall 1962: 233). Coleman notes that by 1897 Bernard was able to state uncontroversially that:

> …there are not two chemistries or two physics, the one applicable to living creatures and the other to inert bodies; rather there are general laws applicable to all substances[s], however [they] might be disposed, and these laws admit of no exception (in Coleman 1987: 126).

Neither this nor a great deal of the same (and there was much, much more) goes to prove emergentism false. There is no limit to the cunning with which those areas not yet given physical explanations can be exploited in the name of the operation of supposed non-physical causes. Nonetheless this research is more than suggestive: not only is energy conserved, energy of known types is conserved. Conservation by itself poses no threat to any but the most idiotic emergentist, and all it takes to make an emergentist position safe, in principle, from criticism based on energy conservation is some stipulation to the effect that the energy books balance when the emergent forces are taken into account.

What we are looking at here is not simply the conservation of energy, though. On the one hand it is the conservation of energy of known types: mechanical, kinetic, electrical, chemical. This means that any supposed additional types of energy have to fit into an increasingly precise and detailed set of constraints imposed by the growing understanding of how known types of energy inter-translate and are transferred. So, for example, when the energy conserving translation from heat to mechanical energy is understood (Joule) there is no need to suppose that any additional principle is at work in biological systems performing the same kind of transfer. On the other hand the research described is, in an important sense, only the beginning of an ongoing series of enquiries into the mechanisms suggested by the discoveries already made. The very table of contents of a much more recent work (even the title is suggestive: *The Physical Basis of Organic Chemistry*) gives a partial outline of some of the outcomes of these enquiries. It includes chapters on thermochemistry, catalysis and isotopes, with more specific discussions of problems related to the shapes of molecules, free energy, forms of entropy, photon-transfers and more, all within the limits of well established science (Maskill 1985: ix-xii). The better these mechanisms are understood, the smaller and less numerous the explanatory gaps which present any hope for the emergentist become, and the overwhelming trend has been for physical mechanisms to be found in all cases.

My reference to the recent work on organic chemistry partly jumps the gun, though, since there is an important intermediate step between the nineteenth century researches I have described and the work represented by that volume, which is quantum mechanics. In the face of all of the work I have described above, a single and promising life-line for the turn of the century emergentist was chemistry. Although a great many chemical laws had been discovered, there was nothing which counted as an even *prima facie* worthwhile explanation of chemical bonding in terms of the more fundamental physical constituents of atoms and molecules, nor

anything except speculation as to what the mechanisms of chemical bonding might be. Quite naturally, and along the lines we have seen already, some of the speculation included hypotheses to the effect that there were peculiarly chemical forces. So, in 1909, Nernst noted that there was no acceptable answer to the question about what underlay chemical bonding, and opined that it seemed 'reasonably certain that we should admit the existence, not only of electrical and therefore polar forces, but of nonpolar natural forces somewhat of the nature of Newtonian gravity' (Quoted in McLaughlin 1992: 55).[68]

Some of the early emergentists had anticipated the possibility that chemical bonding could be explained by reference to the properties of electrons.[69] Nonetheless, at the time they were entertaining these scenarios, the electron view was certainly no more plausible than the view that there were peculiarly chemical forces. The development of quantum mechanics showed, though, that the electron, properly understood, was equal to the task.

In a series of theoretical and experimental enquiries which there is no need to go over in detail here the atom was made the object of detailed physical understanding.[70] Before the end of the nineteenth century Thomson had discovered alpha and beta rays. By 1911 Rutherford had published his nuclear theory of the atom, and soon after met Bohr. And very soon after, in 1913, Bohr had developed a model of the atom considering the nucleus and electrons together which remains the basis of physical understanding of atoms and molecules. A key test for the theory was whether it could predict the emission spectra of excited atoms, reduced to a descriptive formula for some elements by Balmer, but until Bohr not given anything like a satisfactory causal explanation. The Balmer formula was a guide to Bohr in a similar way to Kepler's laws of planetary motion were to Newton: a description which gave no guide to dynamics even while being a measure of the success of any dynamics. In three papers that year Bohr first detailed his theory with respect to the hydrogen atom, then with respect to atoms heavier than hydrogen and the structure of the periodic table, and finally with respect to the structure of molecules (Pais 1991: 146-152). It was through Bohr's work that the concept of the valence electron entered physics, the role of the outermost electron orbit in all chemical bonding was first properly understood, and physical chemistry born.

As with any significant scientific achievement, Bohr's left plenty to be done: by clearly separating atomic/molecular and nuclear physics the questions peculiar to nuclear physics were thrown into starker relief

---

68 McLaughlin quotes the passage from Banz, D. A. (1980) 'The Structure of Discovery: Evolution of Structural Accounts of Chemical Bonding' in Nickles, T. (ed), *Scientific Discovery: Case Studies*, Dordrecht: Reidel. Broad (1926: 65) referred to chemistry as the 'most plausible' candidate for an 'example of emergent behaviour'.

69 For example Lewes (1875: 414) concedes that at some point in the future the phenomena he urges us to recognise as emergent could be found amenable to mathematical treatment in a physical theory, and hence not be emergent in the strong sense.

70 A useful survey of this history is presented in Pais (1991: 117-223), and several of the key papers are anthologised in Van der Waerden (1968).

than they had been when thought of along the lines of Thomson and Rutherford. For all its genius Bohr's work on hydrogen and helium did not generalise beyond simple atoms as directly and clearly as desired, and it was by no means the case that the entire periodic table fell simply and unambiguously out of his model. Nonetheless chemistry had been made physical: no causal principles other than physical ones were required to explain chemical processes.

With this, I propose, the final nail was driven in the strong emergentist coffin: biological processes could already be understood within the framework of a nineteenth century theory of energy and work, as being essentially chemical and mechanical. Now chemistry, the last outpost of serious emergentists hopes, and the site of their greatest empirical strength, had been shown to have an exclusively physical basis.

There was, and still is, more work to be done. The present thesis could have been more than filled with a simple list of the problems which have fallen to physico-mathematical treatment in the intervening time: the application of the theory of self-organisation (without 'extra' causal principles) to the morphology of living things,[71] and to such specific problems as how the arrangements of seed in a sunflower tends to a configuration which can be numbered as a Fibonacci series;[72] this type of research in conjunction with fractal geometry showing how intricate structure can arise from local adherence to simple rules, the use of cellular automata to model the ways in which ants find food and form feeding chains, or how flocks of birds maintain coherence without collisions, all without an overall plan or guidance;[73] the discovery of the molecular basis of inheritance, the list goes on. Morgan (1933: 1) stated boldly that it was 'beyond the wit of man to number the instances of emergence'. My conclusion is the opposite.

The above argument is not essentially inductive. It would certainly be possible to reason that there have been a succession of candidates for non-physical forces which one by one have been found to be unnecessary and to conclude on that basis that future candidates of the same type are likely to go the same way. I do not object to such an inference, but that is not the best way to think of the argument and account of the evidence against emergentism which I have just offered. My claim is not a generalisation over the history of putative non-physical forces, but an account of the overall state of science, to the effect that there is neither *place* nor *need* for non-physical forces.[74] Contemporary emergentists may or may not be likely to be wrong given the fates of earlier emergentists. The point I have been trying to make is that all emergentists appear to be wrong on the basis of the available evidence. Physical forces are enough.

---

71 e.g. Mandelbrot (1977).

72 e.g. Schwabe (1979) Schwabe and Clewer (1984).

73 Koza (1992), Reynolds (1987).

74 This point came out of a discussion with Andries Gouws.

That there are still adherents of strong emergentism, such as Sperry, attests more to persistence on the part of those adherents than to the evidential merits of that doctrine. On any reasonable and balanced evaluation of the evidence the plain fact of the matter is that the only fundamental forces which make any significant demands on our recognition are physical forces. McLaughlin summarises the state of the evidence with the requisite emphasis when he says:

> Given the achievements of quantum mechanics and [...] other scientific theories, there seems not a scintilla of evidence that there are emergent causal powers or laws in the sense in question; there seems not a scintilla of evidence that there are configurational forces; and there seems not a scintilla of evidence that there is downwards causation from the psychological, biological or chemical levels (1992: 54-55).

If the foregoing is correct then there has not been a significant question about whether strong emergentism is true since the 1920s. Certainly discussion of emergentism died out fairly sharply soon after the rise of quantum mechanics, a point also made by McLaughlin (e.g. 1992: 89). It is also the case, though, that the topic has undergone a recent revival of interest, and has also been the object of some serious attempts at defence.

In the following two sub-sections I consider the two most significant lines of defence which I have encountered, one related to quantum mechanics and the other to chaos theory. Before looking at either, though, I reiterate the distinction between weak and strong emergence drawn above set out in section (3.1) above. Recall that weak emergence, whatever its wider significance and interest, poses no threat at all to the Completeness Thesis. Not only that, since in many debates and discussions the truth of some version of the Thesis is largely taken for granted, much discussion of emergence is specifically concerned with the weak variety of that doctrine, i.e. with the emergence of attributes, properties or the conditions for the applicability of certain descriptions.[75]

7.2. Quantum Emergence

It would be somewhat ironic were there to be good reason to think that quantum mechanics suggests the truth of strong emergence, given the historical role of QM in the defeat of traditional emergentism. Quantum mechanics is certainly at odds with much of classical physics, and the prolonged period of theoretical debate and confusion over the possible occurrence, nature and causes of state reduction (the transition from the evolution of several probabilities in the Schrödinger equation to the unique result of a macroscopic measurement) saw no end of extraordinary speculations. No less a physicist than Wigner (1961) entertained the notion that it was *consciousness* which brought about state reduction. If true this would directly indicate the

---

[75] The vast majority of the papers in the Beckermann, Flohr, and Kim (1992) collection, for example, discuss emergence in the weak sense without even mentioning the alternative view.

falsehood of the Completeness Thesis. The days of such heady speculations are largely over, and while the proper way to understand QM and the increasingly venerable measurement problem remains both unclear and contested, there is widespread consensus that no non-physical factors are relevant to micro-physical processes. Put another way, although a probabilistic element to microphysical causation seems irreducible, the only factors relevant to the determinations of the relevant probabilities are themselves physical.

Not only that, some progress is being made in the direction of resolving the measurement problem which faces standard of interpretations of quantum mechanics which are committed to the occurence of state reduction. The GRW conjection (Ghirardi, Rimini, and Weber 1986) for example abolishes state reduction at the expense of assuming that there is a very small (although definite) probability of any particle spontaneously undergoing a state reduction which leaves it in a relatively definite position.[76] This has the twin consequences that spontaneous collapses are extremely rare for isolated micro-particles, but very common for macroscopic objects and especially for macroscopic measuring systems. (More plainly, if the GRW conjecture is correct, then Schrödinger's cat is always alive or dead, never in a superposition of life and death.) The GRW conjecture is not free from problems,[77] and it is not my intention to make my argument hostage to the fate of that ingenious but contested proposal. Rather I want to emphasise that a protracted theoretical debate lasting almost the whole of the 20th century, which at one time gave serious attention to suggestions which implied the denial of the Completeness Thesis, seeks its answers within physics, and is making significant progress.

Nonetheless, there are some who claim to discern emergence in microphysics. Silberstein and McGeever (1999: 182) distinguish between epistemological emergence, which is 'merely and artefact of a particular model or formalism generated by macroscopic analysis' and ontological emergence where there are 'features of systems or wholes that possess causal powers not reducible to any of the intrinsic capacities of the parts nor to any of the (reducible) relations between the parts.' This distinction is very close to that which I urge above, between strong and weak emergence, and it is clear enough that a case for 'ontological' emergence where the emergent power is non-physical is a problem for the Completeness Thesis.

Silberstein and McGeever further maintain that 'fundamental physics provides us with the strongest empirical evidence for ontological emergence' (1999: 184) and take quantum mechanics to provide 'the most conclusive evidence for the existence of emergence'(1999: 187). They attempt to justify this claim by means of a discussion of EPR-Bohm[78] type systems (1999: 187-9). The systems in question are ones where some microphysical event results in there being two spatially separated systems or particles where some quantity is

---

76 A precisely definite position would result in completely uncertain momenta, a prediction which is in conflict with the evidence. Albert (1992: 92-111) contains a helpful and careful discussion of the GRW conjecture.

77 Bohm and Hiley (1993:326-329) discuss some problems, as does Albert (1992: 104-111).

78 The 'EPR' refers to Einstein, Podolsky and Rosen (1935), while the 'Bohm' acknowledges the greater elegance of Bohm's (1951) reformulation of the EPR argument.

conserved across both systems, but where in terms of the standard QM formalism the values of that quantity in each system are not determinate. Bohm's example is of a decay of a single particle with spin-zero into a pair of particles (**A** and **B**) each with spin-half. Given that spin is conserved, a measurement of spin on one of the particles (**A**) after the decay, once the two are spatially separated, should result, it seems, in both spin values becoming definite at once. This, though, is reasonably thought of as a physical change caused by the 'becoming definite' of the spin of **A**, and the requirement that it take place instantly means that the influence from **A** to **B** cannot be Lorenz-invariant, i.e. it runs afoul of special relativity by requiring 'non-local' influences. The EPR argument was originally intended as a *reductio* of the thesis that quantum mechanics could be a complete description of microphysical systems, and an appeal for work on additional variables interpretations of QM. Later work by Bell (1964) showed that, subject to certain conceptual restrictions, and modification of QM with additional variables which restored determinism would also have to be non-local if it was to equal the empirical accuracy of standard QM.[79] The upshot is that the measurement correlations in experiment conform extremely closely to what QM predicts, but that our best theory of the systems in question presents us with the threat of non-locality whether we take it as it stands, or attempt to repair it in order to salvage determinism.[80]

Silberstein and McGeever take up the issue of the correlations as follows:

> There appear to be only three plausible explanations of the agreements: (i) the parts of the system agreed in advance on which outcomes they would yield for each kind of measurement; or (ii) the parts non-locally influence one another so as to come to agreement when measured in the same way; or (iii) the parts exhibit a holistic (emergent) correlation property possessed by the system but not locally carried by the separate parts (1999: 187).

What can be said about Silberstein and McGeever's three options? The notion of the parts of a microphysical system 'agreeing in advance' is somewhat sloppy, suggesting that active agency, not to mention conspiracy, is being attributed to sub-atomic particles. Stated more carefully, though, option (i) indicates the possibility that some additional variables interpretation of QM is true, where the additional variables restore determinacy to microphysical processes. Theoretical and empirical considerations suggest that this scenario is wildly implausible, either because the number of such variables would have to be vast, or because no version of QM which restores determinism can produce the same predictions as unmodified QM, the empirical credentials of which are impeccable. Silberstein and McGeever are, then, correct to reject this option, and also correct when

---

79 More precisely Bell's result indicates that modified QM would have to produce different predictions from those of standard QM. The experimental tests of Aspect and others (Aspect, Dalibard and Roger 1982) indicate very strongly that additional variables theories of the type which Bell considered are empirically ruled out.

80 Bohm's approach to QM restores determinism and has additional variables, but gives the same predictions as standard QM. His approach, though, has difficulties with special relativity which have) at least not yet been resolved. Bohm and Hiley (1993: 271-295) address this question, but as Albert (1993: 160-1) notes there is not yet any general relativistic extension of Bohm's approach, rather an incomplete collection of partial solutions to the problem.

they reject option (ii) on the grounds that it exacts a prohibitively high price: non-local influences of the sort required would result in physical processes which could not be Lorenz-invariant, generating conflict with the special theory of relativity, a theory which strongly demands our assent on various grounds. Silberstein and McGeever conclude:

> Therefore, since there is every reason to think that special relativity is true and we do not want to give it up, holism must be the correct explanation for EPR–Bohm systems (1999: 188).

Having made it this far, Silberstein and Mcgeever run out of focus for a while, and conclude their discussion of quantum mechanics with two rather vague paragraphs about entanglement, parts and wholes, failures of mereological supervenience and the like. The most defensible of the claims which they advance is their suggestion that one way to interpret the material to which they refer is as a failure of atomism,[81] although there are less sure about the significance of this thesis. They also note that their views are in some respects close to those of Teller (e.g. 1986) and even to Bohm, since despite the latter's adherence to an additional variables model, his view also attaches great importance non-separability and entanglement (e.g. Bohm 1980).

What kind of emergence is this? Is it, in fact, any kind at all? Certainly that there is ample evidence for such processes counts against the 'classical picture of the physical world' (Silberstein and McGeever 1999: 189) but that is not what is at stake here. Non-separability is not strong emergence, which requires the notion of independent causal powers, not that of wholes with causal powers not themselves further decomposable. The failure of atomism is no challenge to the Completeness Thesis in cases where there is no doubt about the physical nature of the factors relevant to the causal conduct of whatever it is that cannot be further divided.

Although I am unwilling to proliferate distinctions, one way of separating what Silberstein and McGeever are concerned with and what is of central interest to me here, is to say that if their analysis shows any kind of emergence at all, it is 'horizontal' emergence, which is quite innocuous from the perspective of the Completeness Thesis. 'Vertical' emergence, that is to say emergence which indicates downwards causation is another matter entirely, but it is precisely that kind of emergence which so grievously lacks empirical credentials. Teller (1986) to whom Silberstein and McGeever refer is quite careful to make this clear, first by avoiding talk of emergence in favour of what he calls 'relational holism' and also by emphasising that his analysis does not call for the abandonment of physicalism, merely a shift from 'local' to 'global' versions of that doctrine, where the global version is not narrowly tied to any expectation of atomism.

There are no difficulties for the Completeness Thesis here.

---

[81] This is an interpretation for which I have argued (Spurrett 1997) although at the time I took the failure of atomism to be circumstantial evidence for emergentism, since it meant that, in Bhaskar's sense, the conditions for closure were not met. The criticism of Bhaskar in section (6) above should make clear that I am no longer of this conviction.

7.3. Chaotic Emergence

As noted at several points in this Chapter, chaos theory and the study of cellular automata and self-organising systems are at the centre of much of the revival of interest in emergentism. Most of this work concerns what I am calling weak emergence, a point emphasised again and again by most of those in the fields in question. It will be worth spending a little time explaining why this is so, since thus far I have said little to defend this claim, and the frequent references to 'emergence' by computer scientists, mathematicians, and others call for a more pointed engagement on my part.

The principle reason that chaotic systems pose no difficulty for the Completeness Thesis is that there is no doubt that they are causally closed. What is interesting about chaotic systems is precisely that they are deterministic yet nonetheless not typically predictable. Any emergence has, then, to be weak emergence. In a recent paper Newman (1996: 245) argues that 'chaotic systems provide concrete examples' of emergent properties, and in particular contends that the property of being in the basin of attraction of any strange attractor is an emergent property. A brief look at his argument will make clear that it is weak emergence that he has in mind.

Newman discusses Broad's[82] view of emergence which he takes to be centrally concerned with the question of the conditions under which it might be possible to predict the 'existence of a property or the possibility that a system might have that property' (1996: 247).[83] He quotes an important passage from Broad:

> The emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be deduced from the most

---

82 There are a number of lacunae and errors in Newman's account of the history of emergentism. I do not want to delay the main argument more than necessary, but note the following: His account of Broad's emergentism has Broad's central concern as being non-deducible properties, whereas there is ample evidence that Broad was (like Mill and other British emergentists before him) concerned with *processes* and the *behaviour* of composite wholes, while Mill and the rest were also concerned with the question of whether the instantiation of a property was predictable. (As we saw in section 2.1 above, Mill makes use of the example of the taste of a chemical compound which is qualitatively distinct from the tastes of its constituents.) Newman notes Broad and Mill and says of emergentism that the 'other major advocates were George Alexander and C. Lloyd Morgan' (1999: 247) for some reason leaving out Lewes and Bain. McLaughlin (1992) gives a far better account of these matters.

83 Newman regularly and unfortunately refers to the question of the 'existence' of properties. Whether a property is emergent or not has, of course, nothing at all to do with whether that property 'exists' (whatever that might mean) but relates rather to the circumstances in which the property might be *instantiated*. Except when quoting Newman I will put things in terms of instantiation.

> complete knowledge of the properties of A, B, and C in isolation or in any other wholes which are not of the form R(A,B,C). (Broad 1926: 61, in Newman 1999: 246).[84]

Newman discusses a number of different historical approaches to the question of understanding emergence, and decides that Broad's is the most promising. He offers his own souped-up version of Broad with the following definition:

> A property designated by a predicate **P** in an ideal theory **T** is emergent if and only if the following conditions are met:
>
> (1) **T** describes a class of systems **SG** which are structured aggregates of entities described by **T′**; **T′** is an ideal theory of those entities, and the entities described by **T** strongly supervene on those described by **T′**.
>
> (2) Occurrences of the property designated by **P** are epistemically impossible to identify with occurrences of any property finitely describable in **T′**.
>
> (3) Each occurrence of the property designated by **P** is an occurrence of one of a set of properties **PC**, which is modelled by **T′**. Each member of **PC** is epistemically indistinguishable in **T′** from some other member of **PC** (1996: 251).

Newman is making fairly heavy going of all this, but the point is clear enough: his notion of emergence is one of weak emergence, or in his term 'epistemic'. This way of putting things is close to that of Silbertstein and McGeever, whose account of chaos (1999: 190-198) is more careful than their treatment of QM. They say that the study of chaos and complexity alike suggest only epistemological emergence, their reasoning for this being that the systems in question are causally closed, and also that they operate with a single level of causation, usually some abstraction from a physical theory. As Kellert puts it chaos theory 'introduces no new postulates about the physical world at all' (1993: 41, quoted in Silberstein and McGeever 1999: 191).[85]

This way of thinking of things is, it must be noted, at odd with that of some who write on chaos theory and complex systems theory. Scott, for example, argues that:

> '...*coherent structures must be treated as independent dynamic entities*. Each emergent structure has its own regularities of propagation and interaction, which must be respected.' Scott (1999: 22).

Scott, though, is simply incorrect. The coherent structures he speaks of, for example when he discusses the work on the dynamics of equipartition which Ulam and Fermi had done on the MANIAC computer at Los Alamos (1999: 15), are all the reproducible results of local applications of deterministic rules in the causally closed systems of digital computers. No matter how *interesting* it is that coherent and relatively stable

---

[84] Newman quotes from an American 1929 edition (Harcourt Brace) which is otherwise the same as the first (1925) English edition to which I refer throughout.

[85] That chaos and complexity involve no departures from causal closure at the relevant micro-level is a point repeatedly made in Holland (1998).

structures can occur under these conditions, it is not *evidence* of the action of independent causal powers over and above those of the constituents of the systems in question. Holland marks the distinction which needs to be drawn when he contends that what emerges in these cases is the possibility of certain descriptions in terms of higher level rules or laws, which may be useful for some purpose which we have, but nonetheless where what happens is the result of a 'small number of rules or laws' (1998: 3-4).

In a wide ranging treatment of the theory of complex systems and a variety of applications, Auyang (1998) conceives of emergence in terms of the overall organisation of the systems under study, but takes great care to warn against any notion of strong emergence: 'According to our best experimentally confirmed theory, all known stable matter in the universe is made up of three kinds of elementary particle coupled via four kinds of fundamental interaction' (1998: 1). Although giving a great deal of attention to the emergence of these organisational structures and properties, Auyang is careful to emphasise that this does not require any relaxation of commitment to physicalism as an ontological thesis, i.e. to the view that there are no 'nonphysical vital' forces, no 'spiritual' substances (1998: 45) and that the study of complex systems shows that there is no need to suppose that the system has any 'independent funds of causal power' (1998: 54).

Looked at more correctly, chaos and complexity not only fail to be a source of evidence for strong emergence, they are a powerful argument against it. This is so because the ongoing study of this kind of system shows again and again that extraordinarily complex and robust structures can be the result of simple local interactions between uncoordinated elements, which is to say that the case for there being any need of additional casual principles to account for these structures is systematically weakened by the study of chaos and complexity.

8. Conclusion

The upshot of this chapter is that emergentism poses no significant threat to the Completeness Thesis. The attempt to establish that doctrine by means of philosophical argument was seen to fail, leaving the question an empirical one. A survey of the available evidence showed how all of the major phenomena to which emergentists typically draw attention in order to urge the necessity for additional causal principles in order to account for those phenomena, showed that the science of the past century and more has time and time again shown that no such causal principles are required.

For the reasons noted above in section (3.2) this conclusion is just as bad for the would-be dualist as it is the strong emergentist. Both want grounds to think that physical outcomes can have at least partially non-physical causes. Neither can find the evidence to be especially congenial company. In the sense of 'physical' adopted and defended at the end of the preceding chapter, the reasonable conclusion appears to be, on balance of the available evidence, that physics is indeed complete.

Chapter Three: Dupré and Disorder

'All science is either physics or stamp collecting.' – Rutherford.[1]

1. Introduction

Although the philosophical and evidential failures of emergentism noted in the preceding chapter are good news for the defender of the Completeness Thesis, the emergentist approach is by no means the only serious attacking position available. The present and following chapters are substantially concerned with a line of criticism which can be called the 'disorder' approach. Although the disorder approach forms, in many ways, a single body of thought, it is nonetheless convenient to separate it into two separate challenges, one directed mainly at reductionism and taxonomy, and exemplified in the work of John Dupré, and one focussed on the notion of physical law with which the Completeness Thesis is naturally associated, to be found most strikingly in the work of Nancy Cartwright. The present chapter is primarily focussed on Dupré, while a treatment of Cartwright is offered in the following chapter.

It is worth noting that the disorder approach is significantly different from emergentism, although from the perspective of the Completeness Thesis the two do have significant common features (see section (3.3) below). The most crucial difference is that defenders of the disorder approach are, in general, suspicious of the hierarchical implications of emergentism, which they would regard as conceding too much to the traditional image of the generality and scope of physical laws. Cartwright, for example, is emphatic that we should see the world not as a hierarchy but as a *patchwork* consisting of 'tens of thousands of patches, cut up in no particularly logical way, exhibiting tens of thousands of different regularities of countless different forms' (1994: 281, 298). She also protests against the emergentist that the view that macro-properties 'come out of nowhere' falls afoul of the fact that there is 'nothing of the newly landed about these properties' (1994: 290),[2] and makes clear that she wishes to go beyond the simple rejection of *downwards* reduction, which rejection is typical of emergentism, and assault the possibility of 'cross-wise reduction' (1994: 281) in a way which

---

1 This infamous remark was made by Rutherford in 1908 upon accepting The Nobel Prize in Chemistry. See Bernal (1939: 9).

2 In partial defence of the emergentist here it is hardly fair to attribute to emergentism the view that any properties at all 'come out of nowhere'. Not only that, fairly simple considerations, broadly following from thermodynamics and evolution, suggest that there have been living things longer than conscious ones, and that there have been things at all longer than living things. If nothing else this indicates that there is more to be said for the idea that some classes of properties are first instantiated relatively recently to other classes than Cartwright is prepared to allow. Recall Popper's evolutionary emergentism briefly discussed in Chapter Two above (section 2.2).

directly undermines the defensibility of the view that there even *are* nomological or ontological foundations of a sort congenial to either the Completeness Thesis or emergentism.[3]

As noted at the outset of the present work (Introduction section 1.2) the thesis that physics is complete underwrites a popular way of understanding what has been called the unity of science. The notion of the unity of science, whether as a regulative principle or a supposedly attainable ideal, commands less widespread philosophical support, at least explicitly,[4] than it did a few decades ago. Nonetheless, given the historical association between completeness and some programmes for the unity of science, it is not *prima facie* unreasonable for one opposed to unity to perceive in the Completeness Thesis a common foe, which is exactly what Dupré does.

More specifically, for the purposes of the present chapter, Dupré's assault on the unity of science is relevant to the present study of the Completeness Thesis in a number of related ways. At the most general level there are two positions which he defends which amount to criticisms of the notion that physics could be complete: first his defence of 'promiscuous realism' (1993: 7) involving the endorsement of a cross-classifying network of causally capable natural kinds and entities reminiscent of Cartwright's patchwork, in which only some of the kinds are physical ones, and, second, his claim that only the 'most innocuous form' of materialism, one shorn of any suggestion of causal closure, is supportable (1993: 87). I consider each position in turn, and argue that neither of Dupré's cases contains anything which need seriously trouble the defender of the Completeness Thesis. This conclusion is subject to the crucial qualification that on some issues (relating to the status of fundamental physical laws) where Dupré and Cartwright make common cause,[5] I postpone critical discussion until Chapter Four, since Cartwright's approach is more directly focussed on the question of *physical* laws, and also represents the most sophisticated challenge from the disorder line of thinking on this particular point.

Before getting down to the details of Dupré's position it will help to spend some time setting out the main features of two lines of thinking at which he directs a major portion of his philosophical fire, the Oppenheim and Putnam (1958) account of the unity of science and the Kripke–Putnam theory of natural kinds.

---

3 The discussion of the consequences of Dupré's promiscuous realism for the Completeness Thesis in the present chapter (section 4 below) and the account of fundamentalism, emergentism and patchwork realism in Chapter Four (especially section 2) expands on these preliminary contrasts.

4 A point also made in Margolis (1987: ix). Even so, the ideal of the unity of science does come up from time to time, notably in the scientistic tub-thumping of writers such as Gross and Levitt (1997) and Sokal and Bricmont (1998).

5 When I asked Dupré (in May 1999) whether he inclined to emergentism or favoured a Cartwrightian approach to laws, his answer was 'I'm with Nancy on that one'. This is relatively unsurprising since emergentists (at least those who broadly follow Mill) tend to have no quarrel with scientific unity. Cartwright and Dupré have in fact made common cause in one paper: Dupré and Cartwright (1998).

2. The Unity of Science and Natural Kinds

In much of the literature, perhaps especially around the middle of the twentieth century, the Completeness Thesis, or the notion of the possibility of a comprehensive fundamental level of both being and description, was associated with an expectation which can be glossed as 'order at the top: order at the bottom'. What I mean by this is that various thinkers of physicalist, reductionist, and other persuasions tended to think that order, whether taxonomic or in the form of law-like regularities, at just about *any* scale in the world would stand in some kind of relatively stable, and itself typically law-like, relationship with the basic ontological order imagined to entail the truth of some set of fundamental physical laws. There are many examples of this tendency, including theories of type-type identity, causal theories of natural kinds, programmes for inter-theoretic reduction based on bridge laws and the like.

This association is not, as we will see, a necessary one, but it has the effect of presenting two likely targets to the opponent of the Completeness Thesis. One approach would be to make a direct attack on the view that there is order at the bottom at all, or that there even is a coherent sense to the required notion of a 'bottom'. This is, roughly, Cartwright's line, and is discussed in the following chapter. Another approach would be to defend the view that 'disorder at the top' or failures of law-like relations between top and bottom, weakens any inference to order at the bottom, which is, for the most part, the approach taken by Dupré. (Having said that, both of these champions of disorder develop both of these lines of criticism in their respective work, albeit to differing extents.) As we will see both the Oppenheim and Putnam treatment of the unity of science, and the Kripke–Putnam theory of natural kind terms involve commitment to some version of the 'order at the top: order at the bottom' ideal, which partly explains why Dupré should find both positions worthy of attack.

2.1. Oppenheim and Putnam on the Unity of Science

A classic and influential account of the notion of the Unity of Science[6] can be found in Oppenheim and Putnam (1958), for whom the notion of the Unity of Science is a 'working hypothesis' which they argue commands more respect than the alternatives. In the present sub-section I give a compressed account of the main features of their argument, since so much of what Dupré says is directed against it. Oppenheim and Putnam distinguish between three senses of the Unity of Science, which are, in increasing strength: first, Unity of Language, where 'all the terms of science are reduced to the terms of some one discipline' (1958: 3); second, Unity of Laws, which implies Unity of Language, which is 'attained to the extent to which the laws of science become reduced to the laws of some one discipline' (1958: 4); third, a form of Unity of Science which in turn

---

[6] Oppenheim and Putnam consistently capitalise their references to forms of unity, a convention I follow for the purposes of the present sub-section of this thesis only.

implies Unity of Laws, and requires that 'the laws of science are not only reduced to laws of some one discipline, but the laws of that discipline are in some sense "unified" or "connected"' (1958: 4).

Oppenheim and Putnam further note that this third sense of unity, i.e. Unity of Laws plus whatever might meet the demands of their gestures at unification and connection, seems difficult to render precise, and in any event decide not to impose it in their discussion,[7] but rather to focus on the Unity of Science, first, as 'an ideal state of science' and, second, to 'a pervasive *trend* within science, seeking the attainment of that ideal' (1958: 4).[8]

In their argument Oppenheim and Putnam make use of a specific notion of reduction which they refer to as 'micro-reduction' which is developed from an account of reduction appearing in Kemeny and Oppenheim (1956). The key development in the notion of reduction in the earlier paper is the formulation of reductionism *without* commitment to the Nagel type reductionist programme based on bridge principles or 'co-ordinating definitions'.[9] The further developed version used by Oppenheim and Putnam says, given two theories $\mathbf{T_1}$ and $\mathbf{T_2}$, that $\mathbf{T_2}$ can be said to be reduced to $\mathbf{T_1}$ iff:

(1) The vocabulary of T_2 contains terms not in the vocabulary of T_1.

(2) Any observational data explainable by T_2 are explainable by T_1.

(3) T_1 is at least as well systemmatized as T_2. (1958: 5)

This account of reduction between theories (**T**) is applied by extension to *branches* of science (**B**) and the key notion of a micro-reduction here is that 'the branch $\mathbf{B_1}$ deals with the parts of the objects dealt with by $\mathbf{B_2}$' (1958: 6), which presupposes both a universe of discourse for each branch, and the part whole notion (**Pt**)[10]. To count as a micro-reduction[11] there must be a decomposition of entities (or objects in the universe of

---

7 It is also worth noting that they distinguish their approach from that associated with the ideal of a Unity of Method (involving canons of evidence and investigative procedure common to *all* science), and also with proposals that a unification can be effected by *logical* means, perhaps by relating science to phenomenal predicates (1958: 4-5).

8 Feigl (1953) distinguishes three sense of unity of science, where the first is unity of language along the lines of logical positivism, the second naturalism which Feigl glosses as the view that 'the explanatory constructs of all sciences will not go beyond the spatio-temporal-causal frame', and finally '*physicalism in the strict sense*' which postulates the 'potential derivability of *all* scientific laws from the laws of physics' (1953: 382). What he calls 'naturalism' can be expressed as the view that what he later called *physical*$_1$ would turn out to be complete. (See Chapter One above, especially sections 5.1 and 6.4) There is some irony in Feigl's observation that unity of language is the 'basic idea of the Encyclopaedia of Unified Science' given the title that series carried in 1962: Kuhn's *The Structure of Scientific Revolutions*.

9 Nagelian reduction, and the Kemeny and Oppenheim development, are discussed briefly in section (6) of the present chapter.

10 See also Hempel and Oppenheim (1953).

11 The micro-reduction relation is regarded by Oppenheim and Putnam as transitive, irreflexive and asymmetric (1958: 7). That is to say that if **A** micro-reduces **B** and **B** micro-reduces **C** then **A** micro-reduces **C**, that no theory or branch can micro-reduce itself, and that micro-reductions always go in one direction.

discourse) of $\mathbf{B}_2$ into proper parts all within the universe of discourse of $\mathbf{B}_1$. Oppenheim and Putnam contend that micro-reductions, unlike theoretical reductions more generally, are *always* indicative of progress in the direction of the Unity of Science, involving both progress with respect to the Unity of Language and the Unity of Laws (1958: 6).

Oppenheim and Putnam divide the world and science into six levels (The levels are 'Social groups', '(Multicellular) living things', 'Cells', 'Molecules, 'Atoms' and 'Elementary particles' (1958: 9).), and argue that a complete unification of science would depend on each layer being micro-reduced to that below (1958: 8), so as to make *all* scientific laws ultimately stand in deductive relations with those of fundamental physics. (i.e. It is a condition for the plausibility of their working hypothesis that there be *some* fundamental level of being and description.) They maintain that while it is not typically possible to 'skip levels' of micro-reduction, the transitivity of the micro-reduction relation means that we can conceive of a succession of micro-reductions being effected, say between psychology and biology, biology and chemistry, chemistry and atomic physics, and then on to the fundamental physics of elementary particles:[12]

> If this is achieved, then psychological laws will have, *in principle*, been reduced to laws of atomic physics, although it would nevertheless be hopelessly impractical to try to derive the behaviour of a single human being directly from his constitution in terms of elementary particles (1958: 7).

Oppenheim and Putnam take certain steps to rule out trivial micro-reductions where some property is added to the reducing level in an *ad hoc* manner, and also note that a *prima facie* difficulty for their model is presented by the existence of aggregates of things not uniquely belonging to one of their levels. They consider the example of a 'man in a phone booth', which is an aggregate of a level three object ('phone booth') and a level five object ('man'). Few passages better capture the spirit of their paper than the following paragraph:

> The problem posed by such aggregates is not serious, however. We may safely make the assumption that the behaviour of "man in phone booths" (to be carefully distinguished from "men in phone booths") could be completely explained given (a) a complete physicochemical theory (i.e., a theory of levels up to 3, including "phone booths"), and (b) a complete individual psychology (or, more generally, a theory of levels up to 5). With this assumption in force, we are able to say: if we can construct a theory that explains the behaviour of all the objects in our system of levels, then it will also handle the aggregates of such objects (1958:11).

Oppenheim and Putnam are careful to maintain that the question of the truth or falsity of their hypothesis of the Unity of Science is an empirical one, depending on the fortunes of specific research programmes. I note, though, not quite in passing, that they seem to have importantly misconstrued certain

---

[12] Feigl (1953) also argues in terms of reduction of successive levels of explanation. He also (1953: 383) anticipates some of Nagel's points about when reductions might be 'premature' and also unhelpful, a point discussed in section (7) below.

remarks made by Mill in his *System of Logic*. As noted above in Chapter Two (section 2.1), Mill regards the structure of science as ideally deductive, a point which does not escape Oppenheim and Putnam:

> John Stuart Mill asserts [...] that since (in our wording) human social groups are wholes whose parts are individual persons, the "laws of the phenomena of society" are "derived from and may be resolved into the laws of the nature of the individual man." In our terminology, this is to suggest that it is a logical truth that theories concerning social groups (level 6) can be *micro-reduced* by theories concerning individual living things (level 5); and, mutatis mutandis, it would have to be a logical truth that theories concerning any other level can be reduced by theories concerning the next lower level. As a consequence what we have called the "working hypothesis" that unitary science can be attained would likewise be a logical truth (1958: 11-12).[13]

Oppenheim and Putnam go on to berate Mill for being 'not so much wrong as [...] vague' (1958:12) where the substance of the charge is that Mill is apparently guilty of not specifying what is to count as the 'nature of individual man'. It is difficult to know what to make of this allegation. I am happy to concede that the *System of Logic* is a long book, but this proposed reading of Mill is plainly untenable, and founded upon scant acquaintance with the facts. The passage to which Oppenheim and Putnam refer occurs in Book VI of the *System of Logic*, in Chapter vii, entitled 'Of the Chemical, or Experimental Method in the Social Science.' Mill's contention is that in social phenomena we do *not* see the types of non-deducibility he takes to be characteristic of the chemical mode of causal combination,[14] and hence that social science is properly seen as free from heteropathic laws: 'In social phenomena the Composition of Causes is the universal law' (Mill 1972: 879). Mill's claim is not supposed to be 'logical' at all – he considers it to be a contingent and empirical truth that social phenomena are the additive effects of individual human actions, deducible from proper empirically won comprehension of human dispositions and natures. As was noted in Chapter Two (section 2.1) above, Mill's commitments to both deductivism and determinism are entirely consistent with his emergentism, a feature of his thinking which Oppenheim and Putnam, like Bhaskar, appear not to notice at all.

That aside, although I return to Mill more than once in what follows, there is the question of the *evidence* both in favour of and against Oppenheim and Putnam's 'working hypothesis'. Half of their paper (1958: 16-27) is devoted to a survey of various empirical results and research programmes which they contend show both the credibility of the hypothesis and the weaknesses of alternatives. They admit that evidence at the level of social groups is weak,[15] but are optimistic about the prospects for a micro-reduction of economics

---

13 Referring to the 1843 edition, Book VI, Chapter 7.

14 In Oppenheim and Putnam's terms we could say that Mill takes it to be an empirical fact that in some cases even given a complete decomposition of some entity in terms of the **Pt** relation, micro-reduction would fail because the behaviour of entities at the higher level could *not* be exhaustively accounted for by reference to the properties of the entities at the lower level.

15 Although they *are* somewhat desperately encouraged by evidence showing the effect of levels of testosterone on the 'pecking order' manifest in groups of domestic fowl (Oppenheim and Putnam 1958: 16, referring to

based on a theory of rational choice, and for some form of social theory favouring methodological individualism. At all other levels they think the available evidence is stronger, taking it as established that all living entities consist of cells, and favouring at the time relatively new work on cellular automata[16] and neural networks for attempting to understand some of the functions and behaviour of cells. From there on down they refer to a variety of results concerning DNA, molecular biology more widely, evolution, chemistry and fundamental physics, all with varying degrees of enthusiasm, and conclude that they have shown that belief in the 'possibility that all science may one day be reduced to micro-physics (in the sense in which chemistry seems today to be reduced to it),[17] and the presence of a unifying trend towards micro-reduction through much of scientific activity', as a *working hypothesis* has been shown to be both '*justified*' and '*credible*' (1958: 27-28). As will become clear I am not convinced that the balance of evidence really favours Oppenheim and Putnam's confidence in the prospects for unification through micro-reduction. This point is not central to my purposes, though, since the main burden of argument in this chapter is to separate the question of completeness from that of scientific unity.

The relationship between all this and the Completeness Thesis is relatively clear. If the Thesis was false, then science could not possibly be unified in the ways Oppenheim and Putnam suggest. This follows easily enough, since their vision of layers of micro-reductions requires that there be *some* fundamental level of description, and they make clear that they want that level to be physical. Three qualifications need to be noted at once, though. The first is that the Completeness Thesis need not be tied to any particular conception of the physical cashed out in terms of microscopic constituents, the important thing is that there be a closed system of 'inorganic' causes.[18] So there might be a complete physics which still did not do the work required by Oppenheim and Putnam because of problems with the **Pt** relation, say, or unavoidable macro-physical considerations. The second qualification is that there are ways in which science could be unified, and the Completeness Thesis be false. Mill's emergentist deductivist determinism, which Oppenheim and Putnam seem to misread, is an example of this, and is discussed above in Chapter Two. Oppenheim and Putnam are opposed to this form of unity, and in favour of an anti-emergentist version tied to a combination of micro-reductionism and the Completeness Thesis.[19]

Guhl 1956). Bizarrely the crackpot emergentist Mott (1948: 153-6) mentioned in the preceding chapter takes the same research (he refers without giving details to a discussion in LIFE magazine) as evidence of the operation of configurational influences of some sort.

16 It is not clear what they would say about viruses. Cellular automata are discussed in section (6) below.

17 I note that Oppenheim and Putnam share the quantum mechanics inspired confidence in the implausibility of emergentism diagnosed by McLaughlin and discussed in Chapter Two (section 2 and 7.1) above.

18 Recall the discussion of 'physical' in Chapter One above, especially sections (5.1) and (6.4).

19 Besides the features of their argument already described, I note that they say that if there were 'presently unknown attributes of a more radical kind (e.g., attributes which are relevant for explaining the behaviour of living, but not of non-living things)...' then science could not be unified as they suggest (1958: 13).

Not only that, the independence of completeness from unity arises in the other direction as well, which is the third qualification which I want to note here: physics might well be complete and yet Unity of Science in the sense of Oppenheim and Putnam still be quite unattainable. It is a part of the purpose of later sections of the present chapter to argue for the plausibility of precisely this state of affairs, order at the bottom without order at the top, or, more briefly: completeness with disorder. For the moment, though, I note that given Oppenheim and Putnam's account it should be clear how someone opposed to either the unity of science *or* the completeness thesis might think that the two were some kind of package deal, and especially that to argue against unity is by implication to argue against completeness. Fodor puts it as follows:

> I think that many philosophers who accept reductivism do so primarily because they wish to endorse the generality of physics *vis à vis* the special sciences: roughly, the view that all events which fall under the laws of any science are physical events and hence fall under the laws of physics. For such philosophers, saying that physics is basic science and saying that theories in the special sciences must reduce to physical theories have seemed to be two ways of saying the same thing, so that the latter doctrine has come to be a standard construal of the former (1974: 51).

As we will see Dupré, for the most part, takes things to stand in just this way, and argues at length from the failure of reductionism to the falsehood of the Completeness Thesis, in the form of what he calls the principle of causal completeness. He does, it must be noted, have independent arguments against the Thesis, but his case against unity is, I argue, by far the most compelling part of his argument.

I note also that Oppenheim and Putnam could, it seems, have saved themselves a great deal of trouble by arguing directly for the Completeness Thesis or some equivalent first, and then on that basis attempting to defend some version of the Unity of Science by means of separate arguments for the plausibility of reduction.[20] With the exhaustive fundamental level 'in the bag' the task of seeking micro-reductions and part-whole relations presents itself as a *prima facie* more likely prospect than does a huge structure of reductions culminating in some complete fundamental description. I am certainly not endorsing this way of going about things, since, as will become clear in the remainder of this chapter, I am in broad agreement with Dupré about the prospects for scientific unity, even while disagreeing about the defensibility of the Completeness Thesis.

2.2. The Kripke-Putnam view of Natural Kinds[21]

Another target of Dupré's criticism is an account of natural kind terms due, independently, to Putnam (1975, chapters 8 and 12, 1988 Chapter 2) and Kripke (1972). In what follows I will focus primarily on

---

[20] Kim (1998), for example, argues for reduction while taking the Completeness of Physics essentially for granted. See the introduction and conclusion of the present work.

[21] This discussion gives some of the detail omitted from the account given in section (6.1) of Chapter One in connection with Daly's criticisms of Snowdon's account of 'physical'.

Putnam's version. The major ancestor of twentieth century thinking on the question of natural kinds is Locke, who drew a distinction between nominal and real essences, where the nominal essence of something was, roughly, our criteria for picking some things out as members of a relevant group, which Locke saw as related to an idea 'in the mind' (1975: 415) while the real essence was whatever it was that those things essentially had in common, the 'real constitution' (1975: 442) of the things in question. For his part Locke was pessimistic about the chances of real essences ever being known to us, saying that one who tried to 'regulate himself' according to real essences would be 'at a loss' (1874: 416) and hence regarding our classification of the things in the world, including living species, as being the 'Workmanship of the understanding, taking occasion from the *similitude* it observes amongst them' (1975: 415).

Although thinking it unlikely that we could ever come to know the real essences of things, Locke did make clear what he thought a real essence was, that is 'that real constitution of any Thing, which is the foundation of all those Properties that are combined in, and are constantly found to co-exist with the *nominal Essence*; that particular constitution which every Thing has within itself' (1975: 442). Putnam and Kripke are among those who do not share Locke's comprehensive pessimism concerning knowledge of real essences. Before explaining further, I need to note that the notions of a real essence and a natural kind naturally seem to apply only to non-arbitrary collections, unlike, say, all the objects presently in my desk drawers, or everything bigger than a FIFA regulation football. Common candidates for being natural kinds are pure substances, such as gold, and biological species. Being a proper member of a natural kind typically means sharing some key feature or set of features with other members, and entails certain restrictions on which properties the member can and cannot instantiate: Gold in its solid state is ductile, but cannot be nutritious for humans, rabbits cannot fly unaided. For a Lockean the (unknown) real essence is supposed to coincide with the nominal essence in all cases – if some object fits the nominal essence then it cannot fail to be a member of the relevant kind, and if it does not fit it then it cannot be a member.

In strikingly similar ways Putnam and Kripke take up but also modify and extend the Lockean way of thinking. Putnam's[22] view, specifically, is that meanings in general, and natural kinds in particular, are best understood by means of a four-fold scheme, to be thought of as the components of a vector, of syntactic marker, semantic marker, stereotype and extension, where the extension is typically a microstructural or theoretical feature of the world, and where suitable causal contact with these features enables successful reference. The stereotype for Putnam does basically the same heuristic work as the nominal essence does for Locke - it is that in terms of which we typically pick out examples of the kind in question. Since Putnam is prepared to grant that we can come to have knowledge of the real essence he is in a position to grant that not all members of the kind need meet the criteria for individuation specified in the stereotype. The stereotype will

[22] I am indebted to Longworth (1999) for my understanding of Putnam's position on these questions.

typically emphasise common but not essential features, one of Putnam's examples being the stereotype of 'lemon' (1975: 140-146) which might include being yellow in colour, having a tart taste, etc. If the stereotype was strictly to fix the reference, though, certain difficulties would arise:

> The most obvious difficulty is that a natural kind may have *abnormal members*. A green lemon is still a lemon – even if, owing to some abnormality, it *never* turns yellow. A three-legged tiger is still a tiger. Gold in the gaseous state is still Gold. It is only normal lemons that are yellow, tart, etc.; only normal tigers that are four-legged; only gold under normal conditions that is hard, white or yellow, etc. (1975: 140).

Putnam's approach, which deals with the problems created by the identification of kinds with their stereotypes, is to regard kind membership as related to common possession of some essential feature, or Lockean 'real essence' where (with Kripke) he maintains that if some entity has the relevant essential feature then it is metaphysically necessary, although *a posteriori*, that it is a member of the kind. If something is H_2O then it cannot fail to be water. Ordinary linguistic competence does not presuppose knowledge of the extension, for example H_2O, of the word 'water', since that term could conceivably be used by the inhabitants of a world (the proverbial twin earth) or community in just the same way and attended with the same psychological states, although to refer to a superficially indistinguishable but actually distinct substance, say XYZ. Rather, according to Putnam:

> It means that (*we* say) the extension of the term 'water' as *they* (the speakers in question) use it is *in fact* H_2O (Putnam 1975: 269).

Competent speakers do, though, according to Putnam, need to be masters of the markers (semantic and syntactic) and the stereotype, which in the case of 'water' might include 'colourless; transparent; tasteless; thirst-quenching; etc.' (1975: 269). On this view, then, there will be a real fact of the matter about the extension of a natural kind term which it may or may not be possible to discover, and which may or may not be known by any member of the community using the term.[23]

Although there is no general need for ordinary language to change itself to keep in step with the process by which science improves classifications, it should in principle be possible that normal usage adapt itself, as it did, for example, with what Dupré calls the 'dewhaling of fish' (1993: 24). Even if ordinary language remains resolutely unruly, it seems plausible to expect that scientific usage could and should adapt itself. Putnam's view, then, shows one vision of how science could eventually come to furnish an ideal set of classifications by identifying suitable paradigms and sameness relations grounded in essential features, and refining the set of syntactic and semantic markers correlated with them.

---

[23] Kripke's position is more closely tied to the question of *names* refer, but is nonetheless very close to the Putnam view described here.

As was the case with the Oppenheim and Putnam account of the unity of science, there are few strong connections between this approach to natural kinds and the Completeness Thesis. If the Thesis was true, *and* the Putnam-Kripke approach the best way of understanding natural kinds, then it seems plausible that the essential features shared by natural kinds would be physical. But there is no general requirement that they be physical, and it is quite clear that there could be Kripke-Putnam natural kinds without a complete physics, as long as there were essential features of some kind. Not only that, there could be a complete physics without the right kind of essentialism being sustainable, in the event it turned out that at least some kinds did not have essential features of whatever sort. It does seem as though the Putnam approach is more likely to succeed with respect to kinds which are themselves fairly physical to start with, 'water', say, rather than 'omnivore'.

Having set out the main features of the two positions against which so much of Dupré's argument is directed, it is possible to turn to his case for promiscuous realism.

3. Dupré's Promiscuous Realism

In *The Disorder of Things* (1993) Dupré maintains that the standard image of science, according to which the world is 'profoundly orderly' (1993: 2) depends on commitment to three metaphysical assumptions: essentialism, reductionism and determinism. Essentialism, according to Dupré, is the view that there are essential differences between kinds along the lines suggested by Putnam, which hold out the possibility of objectively grounded taxonomies. Reductionism, with Oppenheim and Putnam as the foil, is the view that proper understanding of any entity or process is best arrived at by study of the fundamental constituents of that entity or process, while determinism, which, as we will see, Dupré does not sharply distinguish for the purposes of his major argument from a position he calls 'probabilistic uniformitarianism' (1993: 172), is the view that there is always some total causal truth about any situation, as a result of which any outcome was either necessitated or had some definite probability.[24]

Dupré opposes all three assumptions, his two principal positive theses being the disunity of science, or the view that science neither constitutes nor 'could ever come to constitute, a single, unified project' and the complementary metaphysical assertion of the 'extreme diversity of the contents of the world' (1993: 1). Besides the specific rejection of compositional materialism and causal completeness, partly via analyses of causality and reductionism, both discussed below (section 6), Dupré's major argument concerns taxonomy and the question of natural kinds and essentialism. In opposition to the notion that a major, or even ultimate, goal of science is the furnishing of a single consistent set of sharp and found (rather than imposed) classifications of the things in the world, he contends that 'there are countless legitimate, objectively grounded ways of classifying objects in

[24] Davies (1996: 5-6) gives an essentially similar overview.

the world … [which] may cross-classify one another in indefinitely complex ways' (1993: 18). I begin with his argument for the inescapable disunity of our taxonomic practices.

3.1. Disorder, Disunity and Reduction

A significant portion of my criticism of Dupré relates to what I argue is a selective and perhaps even somewhat misleading way of setting up the reductionism question. My account of his argument for disorder begins, then, with an outline of his account of reductionism. In Dupré (1983: 321) the question with which he begins his discussion is that of whether science can reasonably be 'expected at some time to reach a terminus.' He counterposes two views. On the first, his favoured vision, science is a 'loosely connected collection of more or less independent theories designed to meet particular theoretical or practical interests.' Alternatively, according to the second view, science is 'in the process of being unified' by means of a succession of reductions of parts of science to parts dealing with the next lower level of 'structural components' culminating in the reduction of all of science to particle physics. On this view, apparently, science would reach its culmination in a state where 'the only *fundamental* scientific laws would be those of particle physics', and any new scientific problem would, in principle, be amenable to treatment by means of microphysical enquiry. For good measure we are advised that the notion of the unification of science would 'require the extremely incautious prediction that we will at some point cease to acquire any new interests' (1983: 321).[25]

There is some additional baggage in this account not typically to be found in statements on the subject by reductionists themselves. Neither the idea that the unity of science is a completion of *all* scientific endeavour, nor that its achievement requires the disintegration of curiosity and the cessation of the process of coming up with new practical objectives, are standard unificationist or reductionist fare.[26] Indeed, thus described the reductionist, or unificationist, position is sufficiently implausible that it seems redundant to go on and spend time refuting it, but in the interests of doing so Dupré distinguishes more finely between a range of reductionist theses. He begins by making what he says is possibly the most 'fundamental' distinction, that between 'reduction in principle and reduction in fact' (1983: 321, see also 1993: 95). By reduction in principle he has in mind the view that a reduction in some case is theoretically possible, although practically impossible. The obstacles may themselves be either theoretical or practical, and in some cases be difficult mixtures of both, as, for example, where effecting the reduction would apparently require greater computational power that would be available if every particle in the universe were used in some mighty and maximally efficient number cruncher. Dupré's main concern, though, is with reduction 'as a substantial claim about the prospects for science' (1983: 322). He consequently sets aside most discussion of reduction in principle, bypassing the

---

25 I return to this issue in the discussion of 'stamp collecting' in section (7.3) below.

26 Neither Nagel (1961), Kemeny and Oppenheim (1956), Oppenheim and Putnam (1958) nor Causey (1977) either predict the occurrence of or make any approving noises about the prospect of such outcomes.

question of obstacles by redescribing it as 'theological reductionism' – roughly a thesis about the reductive capabilities of a divinely endowed reducer, and proposes to return to the question of the credentials of theological reductionism towards the end of his paper.[27]

The next major distinction to be drawn is that between synchronic and diachronic reduction, where diachronic reduction is a supposed relation between some theory and its 'historical successor', whereas synchronic reduction involves questions about the 'relations between coexisting theories dealing with different levels of organisation' (1983: 322, also 1993: 94-5). Dupré's major concern is, unsurprisingly, with diachronic rather than synchronic reduction.

Finally, within the class of practical diachronic reductions, Dupré distinguishes between reduction by derivation and by replacement (1983: 323, 1996: 98). The idea here is that replacement reduction involves the gradual rendering obsolete of whole branches of science by the 'expansion of physics into areas of ever increasing structural complexity' while derivational reduction rather claims to show how existing branches of science can be 'shown to consist of perhaps very complex deductive consequences of the laws of particle physics', subject to the development of suitable 'bridge principles'[28] relating the ontologies of the sciences so reduced with physics. With reference to this distinction, Dupré is interested in showing both options to be untenable, but proposes to begin by dealing with derivational reduction, of which he takes Oppenheim and Putnam (1958) as a paradigm example. I will postpone direct criticism of Dupré's account of the various forms of available reduction for a while (see section 6 below), and bearing his taxonomy in mind, turn to his arguments for promiscuous realism.

3.2. Promiscuous Realism about Kinds

Dupré contends that Putnam's position regarding natural kinds is 'entirely untenable', ultimately because it is overwhelmingly implausible that any possible set of experts could deliver the required goods. In particular Dupré argues that with respect to biology Putnam's view 'founders on the complete absence of the sameness relations its application would require' (1993: 24-5) and that it is simply incredible that ordinary usage, various kinds of specialist usage, and scientific usage could ever be brought in step with one another, at least without unacceptable losses in descriptive and explanatory power.

---

27 For my own part I return to the question of 'theological reduction' towards the end of the present chapter in section (7.3).

28 I note briefly here that the Oppenheim and Putnam approach has no place for bridge principles at all, and is committed neither to their possibility nor usefulness. One version of the 'Unity of Language' view, which they mention at the outset of their paper (1958: 3) *does* rely on biconditionals, but their modification of the Kemeny and Oppenheim (1956) version of reduction is part of a tendency to move *away* from reduction on the Nagelian model. These questions are discussed further in section 6 below.

Dupré's first line of argument draws attention to the richness and diversity of taxonomic practices which are not usually thought of as 'scientific' – the terms of ordinary language, including relatively specialised portions of it.[29] He opposes promiscuous realism to what he calls 'taxonomic realism' (1993: 27) which involves commitment to both (a) some form of essentialism about natural kinds, and (b) to the view that there is a single correct system of classification of the natural kinds.[30]

In defence of promiscuous realism Dupré discusses a wide range of classificatory terms in ordinary and various types of specialist, although not specifically 'scientific' usage, all of which might *prima facie* seem to be candidates for natural kind terms. His major point is that the relationship between these terms and the findings of scientific enquiry are quite unlike those supposedly anticipated in Putnam (1975), where scientific work could apparently be expected to refine non-scientific taxonomies. Rather Dupré maintains that 'it is far from universally the case that the preanalytic extension of a term of ordinary language corresponds to *any* recognized biological taxon' (1993: 27). For example the English term 'lily' typically picks out a selection of the various genera of the lily family, a family (Liliaceae) which includes garlic and onions as well. Dupré is surely correct in saying that to suppose, along Putnamesque lines, that the extension of 'lily', given our ordinary stereotype, actually includes garlic and onions would 'amount to a debasement of the English term' (1993: 28).

Similar difficulties abound, Dupré's examples including 'cedars', which do not form a biologically respectable category despite its usefulness in the practice of carpentry, 'rabbits' and 'hares', which are far more sharply distinguished by ordinary usage than by major scientific accounts of species, 'trees', which are not even a scientific kind at all, 'moths' and 'butterflies', 'hawks', 'fish', etc. (1993: 27-31). Dupré notes that most of the features of ordinary taxonomies are driven by human interests, that the consequent classifications are 'overwhelmingly anthropocentric' (1993: 34), but also maintains in a convincing manner that there is a wide variety of interests and practices which might be served by such taxonomies, which means that it is both unsurprising that there are a variety of pragmatically effective schemes of classification, and unlikely that this state of affairs would fail to endure.

I do not wish to go any further into the details of Dupré's case for the point currently at issue, since it is my intention to concede the force of much that he says (how much will become clear below). His provisional conclusion after his preliminary survey is that it provides evidence for his promiscuous realism, since nothing he has said 'either about scientific kinds or about the kinds of ordinary language, suggests that these kinds are in any sense illusory or unreal' (1993: 36). He concedes, though, that there might be those who would be

---

29 This is suggestive of the later Wittgenstein, and Dupré occasionally refers approvingly (e.g. 1993: 10, 242) to Wittgenstein's remarks concerning family resemblance.

30 See also Wilson (1996: 306-308).

happy to grant what he has said, but still hold out for the prospect of science providing an 'orderly refuge from this extrascientific Babel' (1993: 36) and goes on to argue that this prospect too founders on pluralistic realities.

The problem here is a version of the problem with ordinary language: that one type of systematic and scientific classification can produce a taxonomy at odds with another. A portion of Dupré's argument (1993: 38-44) relates to the question of the ontological status of species, and an ongoing debate over whether they are best seen, from a metaphysical point of view, as individuals (so that all of the members of a given species would constitute a single individual on which selection pressures would operate) or natural kinds. With Wilson I regard this difficulty as essentially irrelevant to the wider question of the defensibility of promiscuous realism about the results of scientific work on classification. Wilson points out that even if we do as Dupré suggests, and allow both classification of species as individuals and as natural kinds to go on side by side, what we will have is 'not a form of pluralism about natural kinds at all' (Wilson 1996: 311).[31] Real support for promiscuous realism will come, if it comes from anywhere, from difficulties in classifying species *no matter what* we might think the proper ontological status of a 'species' is. What, with Wilson, I regard as the more interesting and significant part of Dupré's analysis (1993: 44-59) considers evidence for just this.

There are various scientific ways of laying down criteria for species membership. The three most significant are morphological (also the most venerable), biological (or reproductive) and phylogenetic (or genealogical). Roughly the first classifies species according to structural or morphological features of the members and candidate members (such as having livers, or feathers, or flowers), while the second favours considerations relating to whether candidate members successfully reproduce with acknowledged members, and the final scheme inclines to the view that evolutionary history is the proper guide to what is and is not related, including the question of what is to count as a conspecific.

It takes no great effort to see that these three forms of enquiry need not produce compatible results, that strikingly similar life forms might not be suitable mating partners or have evolutionary histories which are as close as those of life forms which are strikingly different, or that strikingly different looking ones might be closely related in terms of genealogy, and perhaps be viable mates, etc., etc (Dupré 1993: 44-59). Matters get worse rather than better when the criteria for species membership are driven by ecological considerations, where

---

[31] Furthermore, even if there was general agreement in favour of either the species as kinds or species as individuals, way of dealing with the ontological question, the problem of setting boundaries and of determining criteria for species membership would remain as pressing as ever. Wilson (1996) in criticism of Dupré argues for the species as individuals approach, but does not rule out the possibility of the kinds approach having a role to play. *Pluralism* about such a metaphysical question presents a confusing prospect at best, but it does seem as though Dupré's *tolerance* is the preferable approach.

what is crucial is how some life form relates to others in the relevant area: for example on what does it predate, and by what is it predated upon?[32]

Dupré's conclusion here is a more forceful version of the pluralism already suggested regarding the taxonomies of ordinary language, arguing that his analysis shows that 'there is no metaphysical guarantee that whatever accounts for the coherence of species and their distinction (such as it is) from other species will be the same for all living things. [...] Thus both ontological and pragmatic grounds support a pluralistic approach to the species question' (1993: 51).

Certainly it is clear that given the present state of knowledge there seems to be very little evidence in favour of the view that a unified account of species will be forthcoming, and that it would consequently be unacceptable to allow one theoretical orientation to trump others in the interim. Furthermore, it would be question begging to endorse some kind of molecular-genetic approach simply because that one seemed most congenial to the expectations of the Putnam account of natural kinds.

As with the argument for pluralism about common sense taxonomies, I do not want to get into more detail than necessary. Two crucial points do have to be noted, though, before I can move on to consider the specific terms of Dupré's realism, and the consequences of all this for the Completeness Thesis.

The first point, and one which Dupré himself notes, is that it would be possible to argue that biological species cannot be natural kinds at all, for the reason that they lack essences. If so, the main force of Dupré's criticism of Putnam's approach to the natural kinds question would evaporate entirely, partly as a result of Dupré's own assault on essentialism, even though the impact of his criticism of scientific unity would remain largely unaffected. This would follow since an argument that species don't have essences is an argument that Putnam's approach to natural kinds does not apply to them at all.[33] Dupré's claim, though, is that he is showing *defects* in Putnam's view. For his own part Dupré notes that he wants to continue to think of species as natural kinds, but to do so in a non-essentialist way (1993: 53, 60-84).[34] I return to this issue in section (5) below.

The second is that Dupré goes on to extend his case further, by considering what seems to be a plausible candidate for something besides a species which might be expected to relate to some determinable essence, the category of sex (1993: 68-73), and again finds that no essentialist account can do the required work. (So that

---

32 Dupré's own discussion of ecology arises in the course of his argument against reductionism (1993: 87-167, especially 107-120) rather than essentialism (1993: 17-84), although for the purposes of my discussion (not to mention his) the two are merely different aspects of one fundamental issue.

33 That is to say that if Dupré is correct, and 'dogs' no more have a common essence than 'objects bigger than a FIFA regulation football' do, then neither of them would be examples of natural kinds at all.

34 This broad commitment to the view of species as natural kinds, albeit without essences, need not contradict Dupré's above noted pluralism about whether species are kinds or individuals.

some but not all male-female distinctions correspond to a difference in chromosomes, in some but not all species the female plays a disproportionately large role in nurturing young, etc., etc.)

Given all of the reference to human interests in the development of taxonomies, the multiplied analyses of how different approaches classify the things in the world in different ways, one might expect Dupré to be making a case for some or other form of pragmatism, where the only differences which make a difference[35] are ones which bear on some or other human project or practical concern. Such an impression would be seriously misleading though, to the extent that while Dupré's approach takes pragmatic considerations seriously, these anthropocentric aspects of his position are to be tempered by realism, and his ultimate concerns are metaphysical. Promiscuous realism is not just a name; we are supposed to take the realism part of it very seriously indeed. This is despite his partial acceptance of relativism (1993: 262) and the fact that passages like the following seem to invite a pragmatist reading:

> There is no God-given, unique way to classify the innumerable and diverse products of the evolutionary process. There are many plausible and defensible ways of doing so, and the best way of doing so will depend on both the purposes of the classification and the peculiarities of the organisms in question, whether those purposes belong to what is traditionally considered part of science or part of ordinary life (1993: 57).

Dupré is equally emphatic, though, that these commitments are quite compatible with realism, saying that he sees 'no possible reason why commitment to many overlapping kinds of things should threaten the reality of any of them' and also that, while remaining agnostic about many specific candidates for being counted as real, he does *not* see 'why realism should have any tendency to cramp one's ontological style'. He is also adamant that he is *not* relativist in the sense that all scientific theories, claims, models, etc., are equally worthwhile, and he makes clear that he wants to hold short of the relatively indiscriminate tolerance advocated by Feyerabend, while still sympathetic to much of Feyerabend's philosophy of science (1993: 262-3). His particular realism involves, it appears, the notion that there are independently existing things, which may, in the ways described, play any of a number of roles in various taxonomic practices:

> A certain entity may be a real whale, a real mammal, a real top predator in the food chain, and even a real fish (1993: 262).

It is not entirely clear what Dupré has in mind here, since *prima facie* this last remark can be read as being at odds with other parts of his case for promiscuous realism. This is so, because it would make sense to regard the 'entity', whatever it is, as in principle describable in some sparse and neutral way, and the accounts in terms of 'mammal', 'top predator' and 'fish' as various *descriptions* of this entity. Davies (1996 12) takes this kind of remark as evidence that Dupré is committed to there being some minimal ontology, a 'determinate

---

[35] The slogan is a gloss of James, see, e.g. (1907: 43-81).

set of objects that can be sorted in different ways'. On the other hand, if there is to be such an ontology then it seems that our ontological style would thereby be 'cramped' in just the ways Dupré is anxious to avoid, and it would also be odd to call a position aiming primarily to defend the propriety of certain forms of description not aimed at this minimal ontology a *realism*. In what follows I take Dupré's claim to be realist about 'many overlapping kinds of things' and his refusal to allow his ontological style to be cramped at face value.[36] In due course I turn to the question of criticism, but for now want to consider what the consequences of all this might be for the Completeness Thesis.

4. Consequences for the Completeness Thesis

It will be reasonably clear that the promiscuous realist position Dupré is defending does not sit particularly well with the notion of the completeness of physics. If we are to be realist about a richly varied collection of natural kinds, only a subset of them obviously physical and those in all likelihood overlapping with non-physical kinds, then it seems unlikely that we could possibly regard physics as complete. Before the question of completeness can even arise we need reason to be confident that there is some stable and well defined domain of the physical, about which the question of causal closure can then be posed. It is, admittedly, possible to imagine that all of the things about which we were promiscuously realist could supervene[37] on physical features of the world, be type or token identical with physical features, or be decomposable into proper physical parts, but that is not Dupré's intention, and he develops further arguments to head off that possibility, arguments which are discussed below (section 7). Even without those arguments, though, it is clear that he thinks the credibility of any idea of a unified fundamental science has been weakened by his case for promiscuous realism at non-fundamental scales. I turn to the question of the acceptability of this inference in due course, but for the time being simply note that Dupré wishes us to understand the variety of kinds he urges us to recognise as being, at least in general, causally capable, and it is here that the conflict between his views and the Completeness Thesis is most clear.

This is perhaps most striking in his essay 'The Solution to the Problem of the Freedom of the Will' (1996b).[38] Opposing traditional neo-Humean forms of compatibilism he insists that a proper account of agency requires that we do no significant violence to the 'natural intuition that humans are, sometimes, causally efficacious in the world around them' (1996b: 386). He argues also that the threat to agency is not specifically

---

36 Note that Dupré endorses Quine's principle of ontological commitment, see Dupré (1993: 94) and Quine (1960: Chapter 7).

37 Subject to some satisfactory account of supervenience being available, a question discussed briefly in section 6 below, and returned to in the conclusion of this thesis, where Kim's arguments concerning supervenience and epiphenomenalism are discussed.

38 Although the major points are also made in the course of part II of (1993, especially 87-106) albeit not with specific application to the free will question.

determinism, since lapses of determinism in the direction of indeterminism are not good news for agency *per se*, but rather what he calls the 'thesis of causal completeness' (1996b: 386, also 1993: 171-4). What I am calling the Completeness Thesis is clearly a version of what Dupré has in mind, which is the notion that for every situation there is some total causal truth, where it is not important whether that causal truth fixes the likelihood of some outcome, or makes it necessary. Certainly, whatever the correct analysis of freedom might be, I am no *more* free with respect to some occurrence if its objective probability of occurring was fixed but less than unity than I am if that probability was equal to unity.

In striking contrast with traditional incompatibilist approaches which tendentially concede the main features of the causal completeness image, but attempt to find ways of, perhaps partially, exempting[39] human agency from absorption into that nexus, Dupré maintains that 'causal order is everywhere partial and incomplete' (1996b: 386) and that human agency, rather than being characterised by exceptional *breaches* of the causal completeness he rejects, is rather best understood as producing a particular kind of enhanced, and purposive, local causal order brought about by the exercise of independent human causal powers. This alone clearly shows how Dupré's views would require the rejection of the Completeness Thesis, which in his terms amounts to one possible version of the thesis of causal completeness. Some of his further remarks, though, make clear that it is *precisely* the Completeness Thesis which he has in mind.

In a section called 'Microphysical Determinism and the Causal Inefficacy of Everything Else' (1996b: 387-389) he argues that given some situation where there is (a) a bottom level of microscopic entities, i.e. entities which are not further decomposable, (b) out of which entities all existing objects are either made, or on which such objects are wholly dependent, and with respect to which entities we have (c) some causally complete account, then:

> Although heroic attempts have sometimes been made to deny it, it seems to follow inevitably from this set of assumptions that the behaviour of everything is fully determined by the laws at the microlevel (1996b: 387).

It is not my purpose to get into the question of deciding between various competing analyses of freedom here. Nothing crucial for the purposes of my argument will hang on whether freedom can satisfactorily be squared with some kind of thesis of causal completeness, or whether the incompatibilists have it right. What I do need to do is get clear on Dupré's position, which he describes as follows:

> The alternative picture I would like to advocate denies causal completeness at any level. Objects at many, probably all, levels of the structural hierarchy have causal powers. One of the reasons why these causal

---

[39] Taylor (1985: 164-186) calls such approaches 'exceptionalist' and quite properly regards them as metaphysically objectionable.

> powers are never displayed in universal laws (deterministic or probabilistic) is that objects at other levels often interfere with the characteristic exercise of these powers (1996b: 388).

This passage is reminiscent of one in *The Disorder of Things* where Dupré argues that if we take the pluralism he defends seriously we need to recognise causal relations between the kinds we can pick out, and that this makes 'causal completeness at the micro level impossible' (1993: 101). At times Dupré's position seems more than suggestive of some form of emergentism, especially in the paper presently under discussion:

> As objects are united into integrated wholes they acquire new causal properties (perhaps that is exactly what it is for a whole to be—more or less—integrated). I see no reason why these higher level wholes should not have causal properties just as real as those of the lower level wholes out of which they are constructed (1996b: 388).

Nonetheless he maintains that he favours a Cartwrightian approach[40] to the question of causal relations, perhaps in part because emergentism seems to concede too much to the generality of fundamental physical laws in cases where emergent laws are not operating, which is one reason why I am treating the two together over this and the following chapter.

The disorder approach constitutes a relatively radical line of criticism of the Completeness Thesis. One indication of this is the fact that, that unlike some positions, such as emergentism, or dualistic interactionism, which require the Thesis to be false, but nonetheless typically allow there to be *some* total causal truth about any situation, proponents of the disorder approach would typically tend either to deny this, or at least to question the terms in which the notion is formulated. The Meehl and Sellars/Feigl category of *physical*$_1$ discussed in Chapter One above clearly invites the interpretation that our total causal knowledge could at least be added up in some fashion to give, in principle, a full explanation of any event, even if aspects of the thus constructed totality superficially made some outcomes appear over-determined. From Dupré's point of view it would appear that the image of adding up causal accounts is questionable at best, and that the notion that such knowledge can be accumulated rides roughshod over important differences and inconsistencies between various different possible explanations and attributions of causality.

As noted above, Davies (1996) finds evidence in Dupré (1993) for the view that the world does have a determinate basic ontology, resulting in a collection of objects which we can sort in a multitude of ways. He quotes, as evidence for this, Dupré's remark that: 'A certain entity might be a real whale, a real mammal, a real top predator in the food chain, and even a real fish. Many, perhaps all, of these designations might be the appropriate characterisation of that object for some legitimate scientific purpose' (1993: 262). If Davies is

---

[40] Personal communication (1999), and other references in (1993: 110, 187), (1996b: 401 n6)

correct[41] then this points to another way of stating the relationship between Dupré's position and the Completeness Thesis. If Dupré admits a minimal ontology for the world, he nonetheless denies, in effect, that this ontology of the world is exclusively physical, and also that it manifests *any* form of causal closure.

5. Criticism of Dupré

There are a number of ways of criticising Dupré's position. On the one hand one might attempt to shore up the ideal of the unity of science in general, or at least to pick away at details of Dupré's argument for disunity. Alternatively it might seem worth breaking the association between the Completeness Thesis and the unity of science, and seeing what is left to worry about. It is this second approach which I follow in the following section of this chapter, but before getting into that I would like to note a few criticisms of the first type. In the main I think that they fail to defend the unity programme against Dupré's attack, and hence that it is overwhelmingly likely that any credible version of the Completeness Thesis will have to abandon commitment to a general reductionism at all similar to either the Oppenheim and Putnam, or Nagelian, versions. Nonetheless both Dupré's position and the common criticisms of it share failings which indicate that a somewhat different approach is needed. I argue that just such an approach is to be found in the discussion of completeness with disorder which follows in section (5) below. A third type of criticism which I consider briefly in the present section relates to failings in Dupré's position which are not central to the question of completeness, but which nonetheless show significant weaknesses in his approach.

Existing criticisms of Dupré (e.g. Wilson 1996, Ereshefsky 1995) focus for the most part on the question of biological taxonomy, and the general unity of science, rather than on the question of the completeness of physics, or the possible completeness of any fundamental form of description. Wilson (1996: 313-314), for example, argues in favour of the biological species concept (BSC), and more generally supports an integrative rather than pluralistic approach to the species problem and science more widely. He compares the current epistemic situation in biology with that in astronomy in the late sixteenth-century, 'where much integrative conceptual and empirical work lay in the future and the endorsement of a pluralistic realism would have been at best premature' (1996: 312).

In reply to Wilson Dupré (1996a) admits that it is plausible that there be *more* integration in the future than at present, but notes that the type of difficulties he explains in his (1993) still stand in the way of allowing the BSC to become the decisive criterion, notably because there is no guarantee that its taxonomy will correspond to that deriving from phylogenetic considerations, and some reason to expect that the two would fail to coincide in at least some cases.

---

[41] I think he probably is, except that putting things this way can make it look as though Dupré would consider himself close to functionalism, which he signally does not. On this point see Davies (1996: 12-17) and section (7.2) below.

Wilson's astronomical analogy fails to get at the heart of the matter, since in the astronomical case there is, and was even in the height of the seventeenth century period of transition, powerful reason to think that some single account will be the best. While it is possible to be uncertain what that account might actually be, it seems as though there has to be some single fact of the matter, that there is one Earth, one Sun, one Moon and that each follows a single path. Indeed, for this very reason, it is far from clear what might be involved in a Dupré-style realistic pluralism about astronomy. In the biological case, on the other hand, especially given the role of anthropocentric and pragmatic considerations, taxonomic disunity seems more likely to be endemic in the very ways Dupré suggests.

I think that both arguments for and against unity, i.e. those of critics of Dupré and Dupré himself, are prone to a common difficulty, which is that they attempt to extract a metaphysical conclusion about the constitution of the world from an epistemological argument concerning the features of our knowledge. Even if the state of our knowledge was unambiguously either ordered or disordered, no direct transition to a metaphysical conclusion concerning order would be acceptable.[42] It is partly for this reason, and also because I have no particular interest in the order question, that I attempt to sever the connection between completeness and order in the following section of this chapter.

For all that, it *is* possible to say a number of things in defence of the Putnam account of natural kinds, and against Dupré's attack on it. These points have the effect of saving much of Putnam's account, even while leaving the force of most of Dupré's argument for disunity in place. It has been noted already that if biological species are not natural kinds at all, then they present no significant difficulties for Putnam's approach. Although Dupré wants to go on thinking of species, for at least some purposes, as natural kinds, it is clear that he has in mind something importantly different from Putnam. Defending his anti-essentialist view of natural kinds, Dupré says the following:

> If cedars, for example, do all serve a common function for the carpenter, then relative to that human practice I see no reason to deny the naturalness of the kind they form: it is a natural fact, after all, that there exists a class of things, albeit botanically diverse, suited to this role (1993: 63).

Naturalness *relative to a practice* is just not the same thing as natural in the sense of being, ideally, entirely independent of anthropocentric and pragmatic considerations. On the Putnam view the natural kinds are just those which have this independence - relative to any number of practices H_2O and XYZ may be pragmatically interchangeable, and form a 'natural kind' in Dupré's sense, while still forming two natural kinds in Putnam's sense.

---

[42] Davies (1996) makes this point against Dupré and also against some arguments for unity. It does seem as though there is an asymmetry here, since a powerful unified system of knowledge does suggest that the world itself is unified. Whether this inference is supportable is not central to my purposes here.

The most significant point to be made against Dupré's assault on Putnam, though, is that it appears to be misdirected in an important way. On Putnam's view a stereotype will, if it is to refer at all, fix an extension which will not, in general, coincide exactly with the set of things conforming to the stereotype (recall the unripe lemons.) Not all H_2O is 'thirst quenching', since some samples are too small, some are in vaporous form, etc. 'Jade' turns out to be satisfied by jadeite *and* nephrite - distinguished by science but not by the stereotype.

Dupré's objections, though, typically point to mismatches *between* classifications. So, many people think that there are 'trees' but botanists form a group for whom, if Dupré is to be believed, the term has no rigorous use (1993: 26, 35, 63). For good measure yet other people regard 'cedars' as being a distinct species of tree, while others still maintain that the group called 'cedars' by the carpenters includes several different vegetable species. It is surely not contentious to point out here that each group has a different stereotype, including those who don't think that 'trees' form a coherent group at all, and that it is consequently unsurprising that their taxonomies cross-classify.

So, for example, 'cedar' as used by carpenters might fix a conjunctive extension based, as Dupré suggests, more on 'a kind of timber than a biological kind' just as 'jade' fixes the conjunctive extension jadeite and nephrite, based on considerations of appearance and suitability for certain kinds of carving, rather than geology. It could well be perfectly 'natural' for groups of people to classify in this way, but the acceptability of those practices does not devalue the distinction between kinds which are natural in Putnam's sense from those which are in Dupré's sense. The confusions which arise if the two sense of natural under discussion are allowed to be equivocated are intolerable. What if we developed a new chemical flavouring, and then at least some of us classified vegetables on the basis of whether they tasted good with it. Such a collection would be a fair candidate for a natural kind in Dupré's sense, but would we really want to say that a new natural kind in *Putnam's* sense had come into being?[43]

Another consideration, not noted by Dupré, is likely to arise with at least some candidate kinds. This is that the stereotype in some cases might well be vague, as, for example, with the boundaries between 'tree' and 'bush' or 'shrub'. If the stereotype is vague, though, we should not be surprised to find it a poor aid to fixing an extension, and it would seem uncalled for to adopt a promiscuous realist response to such cases.

What becomes clear here is that Dupré's own position is closer to Locke's than he lets on, and when that is seen the failings of his approach can be seen to be similar to those of Locke's. Recall that for Locke the nominal essence was all that we had access to, that a kind was the collection of things picked out in terms of

---

[43] Note that I have no general brief to try and defend either externalism or Putnam's account of natural kind terms here – neither are essential to the Completeness Thesis. The point is rather to show that Dupré seems to have misidentified his target.

their conformity to the nominal essence to be found 'in the mind' (Locke 1975: 415). So for Locke we cannot be mistaken about whether something is a member of a kind if it fits the nominal essence. But this is, surely, just the kind of thing about which we *can* be mistaken. Not all that glitters is gold, and, furthermore, not all gold glitters.

A final point worth noting here is that there is evidence that Putnam himself was aware of at least some of the difficulties involved with thinking of biological species as natural kinds. Consider his discussion of the Twin Earth creatures which are similar to earth 'cats' (i.e. they fit the stereotype). Even though superficially similar, Putnam invites us to contemplate cases where 'cats$_{TW}$' either can or cannot interbreed with earth cats, where they have different local phylogenetic properties, etc. (1988: 35). Not to mention and his discussion of the vagaries of semantic markers in (1975: 266-268) with reference to the hypothetical robot tigers from Mars. In these and other cases what Putnam's account can do justice to and Dupré's not is precisely the case of something which looks, smells and carries on like a cat, or water, or a lemon, but where that is not all there is to the question whether it is one.

In conclusion, Dupré has not really shown that a single non-vague stereotype fails to fix a definite extension, only that we should not expect pragmatic and anthropocentric considerations to evaporate in the face of knowledge about kinds which are natural in Putnam's sense. What I would like to do now, with these remarks concerning Putnam in hand, is return to the line of criticism I mentioned above, that which involves breaking the association between completeness and unity.

6. Completeness With Disorder

A common feature of many arguments which are supposed to count against either physicalism or the completeness of physics is an inference from the manifest variety of kinds of things and relations between them at macroscopic scales to the alleged impossibility of any relatively simple and uniform level of microscopic constituents. For the purposes of this claim 'macroscopic' can be taken quite generally to mean simply 'non-microscopic', that is, any level of being or description above the fundamental, assuming there to be a fundamental level. Examples, besides Dupré, include Bhaskar (1978). In other words, there is evidence for a tendency to argue from disorder at the top to the implausibility of the view that there is or could be order at the bottom.

Contemporary physicalists, who are almost by definition committed to the Completeness Thesis,[44] are on the whole more than willing to grant this macroscopic variety, but insist nonetheless that doing so need not entail any relaxation in their commitment to a fundamental physical level of description. Indeed it is generally the case that over the past few decades many of those who defend the view that there is some kind of true

---

[44] See the Introduction above, especially section (1.1).

fundamental, and usually physical, description of the world have abandoned the hope of finding bridge laws, type-type relations and other expressions of the 'order at the top: order at the bottom' ideal. Examples of this include Fodor's (1974) account of the special sciences, Davidson's (1970) account of the mental as 'anomalous', and Hellman and Thompson's (1975) treatment of physicalism. Hellman and Thompson, for example, note that while physicalism was, for a while, typically associated with some kind of overall reductionism, 'of late there has been a growing awareness, however, that reductionism is an unreasonably strong claim' (1975: 551)[45]. They go on to set out the terms of a version of physicalism not committed to reductionism, and maintain that 'the truth of physicalism is compatible with the utter absence of lawlike or even accidental generalised biconditionals connecting any number of predicates of the higher-level sciences with those of physics' (1975: 564). On the other hand Papineau maintains that it is in general implausible that type-type relations would hold between truths of a complete physics and other truths (1992: 12).[46] Given all this, it is far from clear today, rather than twenty five or thirty years ago, what position is really the target of an attack on the expectation of order at the top based on the Completeness Thesis, since, by and large, defenders of the completeness thesis are no longer wedded to the 'order at the top: order at the bottom' ideal.

It is one thing, though, for physicalists to admit that large-scale taxa may supervene on at best wildly disjunctive sets of lower level constituents, but another to show that simple interactions between simple constituents actually can give rise to the very variety often used as a premise in arguments denying the possibility of a uniform fundamental form of description. An argument on this point is also worth making since the Completeness Thesis has been, as we have seen, associated with the ambition to relate all macro-level taxonomies to micro-level descriptions, or perhaps more weakly to see fundamental descriptions as essentially superior from an epistemological point of view.[47]

6.1. A Look at Life

I propose that the argument which is needed here can be made in the first instance by reference to the field of computing which constructs and studies what are called cellular automata. Cellular automata were originally proposed by Von Neumann (e.g. 1966) as a manageable way of modelling abstract features of physical processes,[48] and in particular (at least in the original research programme) attempting to formally treat the phenomenon of replication. Wolfram nicely captures the research utility of cellular automata as follows:

---

45 In Hellman and Thompson this remark is followed by a footnote to Putnam on 'Reductionism and the Nature of Psychology' and Fodor on 'Psychological Explanation', which is to say that they have functionalism in mind. See section (6) below.

46 In addition Lewis (1983: 361) on the subject of what he calls 'Materialism', notes that it carries no implication of translatability of non-physical predicates into physical ones.

47 Examples include Quine's 'double standard' (1960: §33, §45).

48 Wolfram (1983: 602): 'Cellular automata are mathematical idealizations of physical systems in which space and time are discrete, and physical quantities take on a finite set of discrete values.'

> Cellular automata are sufficiently simple to allow detailed mathematical analysis, yet sufficiently complex to exhibit a wide variety of complicated phenomena. Cellular automata are of sufficient generality to provide simple models for a very wide variety of physical, chemical, biological and other systems (1983: 601).

A typical cellular automaton consists of some *n*-dimensional array of cells each of which can be occupied or be vacant. An occupied cell may either be simply occupied, or be in one of a number of states specified by the parameters of the system. Further refinements are possible, so it is not required that all the occupants be of the same type, and some cellular automata have any of a number of different types, which may in their turn be in a variety of states. Furthermore the space of cells may have additional attributes which locally affect the rules governing what happens with the cells in that region. For the purposes of the main argument here I will not consider any such souped-up automata.

Time in cellular atomata, typically but not necessarily, takes the form of a sequence of discrete generations, and the state of a cell at any time is a function of its state in the previous generation and rules relating this state to the states of cells in the immediate neighbourhood of the cell in question, where by common convention every cell is not considered to be a neighbour of itself.

In an important paper on general features of cellular automata Wolfram classifies cellular automata into types based on their dynamical properties. He divides them into four types (1984: 5), those, class I, which converge quickly onto and remain in a homogenous state, those, class II, which result in separated simple configurations of occupied cells and/or converge on a periodic pattern of states, those, class III, which result in some or other chaotic pattern, and finally those of class IV, which produce 'complex localized structures, sometimes long-lived.' These options correspond roughly to standard dynamical possibilities, in that the first class amounts to a system with point attractor, the second to a periodic attractor or limit cycle, and the last two to 'strange attractors',[49] or at least to configurations and sets of rules which take a very large number of generations to converge on some more stable state of class I or II. Classes III and IV are of most interest here since they manifest the most interesting structures and also have the property that prediction of what structures will result from the application of the rules to a given starting situation is not typically determinable by any means except direct simulation of the rules and configuration (Wolfram 1984: 31).

It will help to use a particular example, although most of the results of work on cellular automata can be readily generalised. Perhaps the most famous ever cellular automaton is John Conways' set of rules and parameters specifying a system modestly known as *Life*.[50]

---

[49] Lorenz (1963) is the classic reference here, although he himself does not use the term 'strange attractor' in this paper, referring to the time development of his artificial weather as 'quasi-periodic' instead.

[50] Barrow (1992: 239) and Levy (1993: 52-56) report that *Life* was originally 'played' on 'Go' boards at Cambridge University, and that a number of key results and proofs had been established before the use of

Before examining *Life* I note that it is increasingly common for some features of cellular automata to be described as 'emergent' (e.g. Forrest 1991, Soucek et al 1992). In the terms of the preceding chapter, as is argued below, these phenomena are 'weakly emergent', which is to say that there is no suggestion of downwards causation or causal autonomy of any sort at non-fundamental levels. My thesis in the discussion of *Life* which follows is that there is more than sufficient evidence for the view that simple local interactions between simple elements can give rise to effects which are sufficiently diverse to present all the key taxonomic worries Dupré considers important evidence for disorder, but in ways involving only, and at best, weak emergence.

In its canonical form *Life* takes place on a two dimensional array of cells where each cell may be in one of two states: occupied or vacant. Each occupied, or 'live', cell is always simply alive – unlike some cellular automata the rules do not distinguish more finely between states of occupied cells. Whether a cell will be occupied or not in any given generation will be a result of applying the following rules to its state in the preceding generation:[51]

> *Death:* Iff an occupied cell has 0, 1, 4, 5, 6, 7, or 8 neighbours (i.e. less than 2 or more than 3), that cell will cease to be occupied.[52]
>
> *Survival:* Iff an occupied cell has two or three neighbours, the cell will be occupied in the next generation.
>
> *Birth:* Iff an unoccupied cell has three occupied neighbours, it becomes occupied.

From the strikingly simple rules of *Life* a remarkable variety of outcomes can follow. First a few general comments. Most sets of initial conditions tend to develop rapidly towards either a steady state of some sort, or to a cycle of states with a relatively short period (between two and five generations), i.e. are members of either Wolfram's class I or class II. But many exhibit class III and IV behaviour, even some which start from configurations with a few as five occupied cells at generation zero.[53]

Besides these (and some other) relatively fundamental characteristics of the evolution of *Life* worlds given different initial conditions, though, there are a large number of categories of structure and behaviour which are picked out by reference to their relatively macroscopic or, to be more pointed, 'functional'[54] properties. These include configurations which, under some interpretation, store and add binary numbers,

---

computers radically transformed cellular automata research. A classic account of *Life* (Conway himself has never published anything on it) is Gardner (1970).

51 See, e.g., Serra and Zaranini (1990: 61-64).

52 In the original spirit of the game the idea was that too few neighbours would cause death through 'loneliness', and too many through 'overcrowding'.

53 For example the figure sometimes called the 'r-peptomino', which takes around 1000 generations to stabilise (e.g. Serra and Zanarini 1990: 63).

54 I return to functionalism below in section (7.2).

generate prime numbers, exhibit various rates of growth, properties of space-filling, forms of 'interaction' with other configurations, and so-on, including instantiating the rules of *Life* itself on larger scales.

One such class of structures are so-called 'spaceships' defined as collections of active cells which over some period result in an exact copy of themselves, but translated to a different portion of the automaton-space.[55] The number of generations which this translation takes is the 'period' of the spaceship. Since the fastest possible rate of 'movement' or 'growth' in Life is one cell per generation this speed is referred to as 'light speed' and abbreviated as 'c'. The simplest spaceships (and also the first to be discovered) are called gliders, constructed out of five occupied cells in a three by three space, and which translate diagonally at c/4 with period 4. The three next smallest (the light, medium and heavy weight space ships) all move orthogonally at 2/c and with period 4, and are composed of nine occupied cells in a five by four space, eleven occupied cells in a six by five space, and thirteen occupied cells in a seven by five cell space respectively.

Besides these, the 'minimal' space-ships, it is possible for there to be arbitrarily large spaceships which may also have a range of additional properties. Some of them emit streams of gliders or other small spaceships, and are referred to as 'glider guns' or simply 'guns'. (Not all guns, though, are spaceships since many of them produce streams of spaceships without themselves moving.[56]) Furthermore not all moving structures are spaceships, since spaceships conventionally are defined not only by their capacity to translate through the life-space, but are also required to do so without leaving debris (other spaceships, including gliders, themselves do *not* conventionally count as debris). Translating structures which do leave debris are referred to as 'puffers' or 'puffer-trains'.

So far so good. The relationships between these larger structures and the fundamental level of description in terms of cells and rules, is by no means tidy, nor are the boundaries between the classifications themselves as well behaved as the brief taxonomy above might suggest. So, to begin with, the light, medium and heavy weight spaceship all have the same basic structure, but larger versions of the same structure do not result in larger spaceships, but rather in shapes which are non-viable *qua* spaceships. This takes place because 'sparks' or single left-over occupied cells which are produced by the smaller spaceships but which in their cases die off for lack of neighbours, do not die off when the pattern or template of such space-ships is made larger than the heavy weight version. Here, rather, the sparks bring about the collapse of the structure, preventing it from becoming a spaceship at all. The class of 'over weight spaceships', though, need not contain only

---

55 In what follows I discuss only spaceships, and some 'life-forms' directly related to spaceships. This does not even begin to exhaust the range of structures which can be picked out and studied at larger scales in *Life*, nor, for reasons to become clear, will it exhaust the possible behaviours of spaceships.

56 Glider guns occur in remarkable variety as well, including some which emit gliders in irregular streams which are only periodic over thousands of generations. By increasing the distance between interfering parts of such glider guns the periodicity of the emission cycle can be increased arbitrarily, and other properties of the cycle introduced.

disasters, since the destructive sparks can be controlled if the overweight spaceship is 'escorted' by some other suitable spaceship moving in the same direction and with the same speed. This arrangement cannot itself be arbitrarily extended, though, and some over weight space-ships cannot be stabilised even by fleets of smaller spaceships, although at this level some *combinations* of overweight spaceships can be stabilised by smaller spaceships that handle the instability left over by the interaction of the over weight components.[57] It is, I hope, by no means clear that there is an obvious or definite answer to the question whether such structures be counted as single large spaceships or ensembles of smaller entities.

Further to all this, not all spaceships larger than the heavyweight version are overweight spaceships with escorts at all, and those that are not are themselves not all of the same kind. A number of families of spaceships not conforming in any way to the glider template can be constructed out of a selection of basic units in conformity with a space-ship 'grammar' specifying which combinations of elements will yield viable spaceships. Structures called 'tagalongs' exist which follow the path of a spaceship, and are parasitic for their continued existence on the sparks emitted by that spaceship. Since by definition the sparks would disappear anyway it is possible in most cases to regard the spaceship as independent of the tagalong, which can be seen as a special type of spaceship relying on certain environmental conditions to exist. Not all tagalongs are quite so independent, though, and there are cases where it is not possible to say definitely whether a large structure is separable into a spaceship and a tagalong, or whether it is 'really' best thought of as a large spaceship with an additional and possibly interesting structural property.[58]

The difference between a puffer and a spaceship may similarly depend on the presence or absence of some additional element, and, again, it will not be a straightforward matter whether a puffer with some additional structure cleaning up the debris is a large spaceship or a combination of a puffer and a further spaceship which has the additional property of clearing debris. (These additional structures are only 'tagalongs' if they would not be viable *without* the debris.)

Why am I setting out all this detail? In a nutshell, because the problems which arise with the classification of the various configurations which occur in *Life* are just the kind of taxonomic difficulties of which Dupré makes so much, and which he takes to count as evidence against the view that there is order at the bottom. Except that in the *Life* case there is no question at all about there being order at the bottom: that's just what cellular automata *are*. The upshot of all this, to be explained more fully in the following subsection, is

---

57 The literature on some aspects of *Life* is not comprehensive. I tested these claims about overweight space-ships myself after hearing them reported as hearsay in discussion.

58 Various interactions between spaceships and stationary structures are also possible. Gliders, for example, can 'collide' with other structures and either transform or destroy them but continue moving, or bring about some transformation but be lost in the process. Alternatively a glider may be reflected and continue in some new direction, or disappear in the collision. When the shape with which the glider collides is not itself permanently transformed it is classified as an 'eater' for gliders coming from the relevant direction.

not to show that disordered phenomena are *in fact* the net result of large numbers of local and simple micro-processes – drawing such a conclusion would be a *non sequitur*. My conclusion is rather that what the *Life* case shows is that no direct inference from disorder at the top to disorder at the bottom can be acceptable.

6.2. The Meaning of Life

The availability of a fundamental and uncontroversial description in terms of occupied and unoccupied cells and relationships between such cells in a single generation or over a series of generations cannot settle the taxonomic difficulties which arise in *Life* even though it can help in some cases. Spaceships are not picked out by reference to their microscopic properties, and the variety of spaceships suggests that any catalogue of microscopic properties of spaceships would be at least significantly disjunctive, and in any event parasitic for its construction on the broadly functional account or theory concerning what spaceships actually do. While it might be possible to lay down arbitrary criteria distinguishing spaceships from puffers it is far from clear that would give any benefit, since these classifications arose partly in the context of different streams of research into cellular automata including efforts to simulate computational processes, memory, communication and also to explore 'life-physics' (For example by trying to make the fastest possible space-ships, which it has been shown would travel at c/2.) and special evolutionary problems concerning certain structures.[59] Since the terms need to be able to do the work required of them for each of these research programmes, and those programmes have as their objects, by and large, the macroscopic features of certain life-structures, such an imposed taxonomic order would tend to be counterproductive.

A symptom of this is the fact that the typical tools of the trade in *Life* research work with functional and macroscopic characteristics - search algorithms which scout large numbers of configurations for ones which fall into classes III or IV, further tools which generate and filter large numbers of 'mutations' of configurations already identified as interesting, etc.[60]

Any 'cellular epistemologist' is, by and large, not interested in the fundamental description despite its ready availability. In fact, a portion of the work on cellular automata goes on under the heading of 'emergent' computing (Forrest 1991, Langton 1991, Koza 1992) where, under some interpretation, information, structure and interactions are discernible at large scales or 'high' levels of the computing system, yet inaccessible (or in some cases at least prohibitively unwieldy) at the level of the basic constituents of the system.[61] The point here is that even though it might be possible, in some cases, for initiates to argue over whether some configuration

---

[59] For example the conjecture that any life-structure can be the result of a set of interactions between gliders. This conjecture has not been proven or disproven, although a number of structures have been successfully "reduced" to being the consequence of interactions between gliders.

[60] See Koza (1992) for a discussion of genetic programming in a cellular automata-like simulation of the behaviour of a virtual ant colony.

[61] This point is made explicitly in Cilliers (1998) and also critically discussed in Spurrett (1999b).

of cells in Conway's *Life* is a 'spaceship' or a 'puffer train', just as biologists can dispute whether a given creature is a member or one species or another, one thing which is *not* in doubt in the cellular automata case is that there is a fundamental description to which any other description will stand in *at least* a token-token relationship.

In a hypothetical case where the order of epistemological business is reversed, and our cellular epistemologist, or better a community of them, are given the surface behaviour of some automaton but not the rules, then the discovery of the rules would, of course, be very significant indeed. It is plausible to imagine that in the absence of knowledge of the underlying rules, any of a number of non-taxonomic controversies might rage over whether downwards causation was operating, whether a fundamental description was possible, whether the plausibility of such a description was affected by the taxonomic debates, and so forth. Even though the discovery of the rules would, hopefully, close such debates, it would not and could not answer or adjudicate on the taxonomic difficulties for the reasons already given.

Those working in these fields are entirely aware of this, and this recognition is a part of what is at stake in the use of the term emergent in some of the literature pertaining to cellular automata[62]. Forrest (1991:1), for example, refers to emergent computing as having as its object 'computational models in which the behaviour of the entire system is in some sense more than the sum of its parts' adding that in these systems 'interesting global behaviour *emerges* from many local interactions.'[63] Nonetheless there is no suggestion of emergence in the sense involving downwards causation, hence Forrest notes that 'the concept of emergent computation cannot contribute magical computational properties' (1991:3). Rather the point is that *under some interpretation* cellular automata contain and process information not discernible at fundamental levels. Even if one accepts that in principle 'emergent computations can be simulated by a Turing machine, interpreting the resulting patterns as computations is likely to be so difficult as to be infeasible' (Forrest 1991: 3).

This means that despite the noted availability of fundamental descriptions for all cellular automata, it does not follow that such descriptions will be the most informative or useful, or that they will be able to adjudicate taxonomic or other disputes. In a sense this argument is a complement to that developed in Fodor (1974), to which I return below in specific connection with Dupré.

So, in conclusion, although it was noted earlier on that if the Completeness Thesis was false, then science could not possibly be unified in the ways Oppenheim and Putnam suggest, it does not follow that

---

62 'Emergent' computing is not restricted to cellular automata, and includes work on neural networks and other technologies.

63 Forrest refers to Hofstadter's requirements for emergent computation, and notes that he 'stresses that information which is absent at lower levels can exist at the level of collective activities' (1991: 2). Forrest also notes that an explicit practical advantage of emergent computations over traditional methods of programming using explicit instructions are that emergent computations are naturally more flexible, since their flexibility is not a function of any set of explicit instructions (1991: 3).

difficulties with unification pose any problem at all for completeness. The above argument does not, of course, establish that the visible variety in the world at non-microscopic levels is *in fact* the result of simple interactions between simple elements. What it does show, rather, is that no direct inference from macroscopic variety and even heterogeneity to the impossibility of a simple and essentially uniform fundamental level is sustainable. How the world is in fact remains an empirical question, and in due course I will consider the relative strength on empirical grounds of the view that physics is complete.

The present stand off should not be surprising, unless we expect to be able to extract metaphysical conclusions from analysis of epistemological practices and premises. Davies (1996: 5, 8-9) makes the point that such arguments cannot in general be decisive, and that they specifically fail to be decisive in Dupré's case. It is an interesting question where the burden of proof might reasonably be expected to lie in this case. For most physicalists it seems clearly to lie with the defender of disorder, but physicalists by and large are people who accept the Completeness Thesis anyway. There is, though, a case to be made that considerations of minimal mutilation of our normal ways of dividing up and understanding the world place the burden squarely with the would be defender of the Completeness Thesis. For my own purposes the burden clearly lies in this side, for the simple reason that I am engaged in an attempt to see how well the Thesis can be defended against various lines of criticism. In due course I will argue that the correct way out of the present stand off is by means of a defence of fundamental physical laws against the criticisms of Cartwright. Before turning to that task, though, I need to complete the treatment of Dupré by dealing in more detail with the question of reductionism.

7. Dupré on reductionism

Anyone sympathetic to Dupré will feel that what has been said so far does not engage with major parts of his argument. Quite so. The remaining parts of his case may be divided into two parts. On the one hand there are some further exercises in sounding the antireductionist[64] battle call, in this case primarily with respect to ecology and genetics.[65] On the other are some arguments more directly aimed at materialist versions of what Dupré calls the principle of causal completeness, which is, as noted, a version of the Completeness Thesis. In addition, somewhere in between these is an additional salvo in the direction of the philosophy of mind, which combines elements of anti-reductionism with a treatment of causality. If what has been said above does not

---

64 Dupré's anti-reductionism and anti-essentialism are closely related: his claim that there are no essences of a certain kind goes hand in hand with his rejection of those forms of reductionism which would seek to relate higher-level classifications and descriptions to such essences.

65 I do not plan to dispute Dupré's claims on this point, and in both cases his arguments extend those already discussed. Thus in genetics (e.g. Dupré 1993: 123, 126-7, 129f) the problems arise because different accounts of 'gene' pick out different things, and the 'gene for **X**' where **X** is some phenotypical characteristic does not usually stand in any tidy relationship with any given strand of DNA, a plausibly molecular account of a gene. In ecology, on the other hand, (Dupré 1993: 107-120) the most appropriate taxonomies rarely coincide with whose arising from evolutionary, morphological and other considerations.

defuse the threatened inference from disorder at the top to disorder at the bottom then it will not help to repeat myself here. On the other hand if it does, then there is no need to go into any of the details of the antireductionism Dupré defends concerning ecology and genetics: there is nothing about disorder or failures of reduction which need provoke alarm on my part, as long as I can defend the view that there is order at the bottom. Reductivism is not an essential part of the completeness package, and expecting all physicalists or defenders of the Completeness Thesis to hold to the truth of some general reductionism is somewhat like expecting anyone who calls themselves an Aristotelean to believe in the spontaneous generation of frogs.

I note, though, and again, that Dupré takes the two lines of argument which I have just separated as being closely interrelated, and is clearly committed to the view that failures of reductionism are more or less automatically difficulties for the Completeness Thesis, as what he calls the 'thesis of causal completeness' (1996b: 386, 1993: 171-4). He describes his treatments of ecology and genetics as 'among other things, preparing the inversion of the reductionist *modus ponens* (causal completeness implies reductionism) into my antireductionist *modus tollens* (the failure of reductionism implies the falsity of causal completeness)' (1993: 101-2), although he also proposes to develop direct criticisms of the principle of causal completeness. I trust that what has been said above cuts off both the inference from disorder to the failure of completeness and that from completeness to order. In the remainder of this chapter I consider a few further details of Dupré's case against reductionism and his independent arguments against causal completeness. This is partly to tie up a few loose ends, and also to press home the point that the very kinds of diversity Dupré celebrates are entirely consistent with the truth of the Completeness Thesis, as can clearly be seen from the forms of physicalism designed to accommodate them.

7.1. Nagel and After

I noted above (section 3.1) that Dupré's account of the various forms of reductionist thesis was selective and the options he presented needlessly attenuated. I need to say a little more about this now, since the ways in which his account is defective bear significantly on the question of the defensibility of his further arguments against reductionism. This is particularly the case with his treatment of the philosophy of mind, which is in some ways the culmination of his assault on both reductionism and causal closure.

It is genuinely difficult to find real examples of serious thinkers who profess either to entertain or endorse the form of reductionism which Dupré sets out to refute. I have already noted that none of the figures Dupré mentions in his reductionist rogues gallery even come close to expecting that 'we will at some point cease to acquire any new interests' (Dupré 1983: 321). More significantly, though, even Nagel, a committed and energetic defender of reductionism is some distance away from the doctrines by which Dupré seems to feel menaced. Nagel's reductionism does take a relatively extreme form, relying on what he called 'co-ordinating definitions' (e.g. Nagel 1961: 344) later often referred to simply as bridge principles. These were supposed to

establish a correspondence between the statements of one theory and another, Nagel's favoured example being the co-ordinating definition that temperature is 'in some way related to […] mean kinetic energy of […] molecular motions' (1961: 344). With this definition in place, according to Nagel, statements about temperature can be translated into statements about mean molecular kinetic energy, and vice versa. Although in one sense the reduced theory loses some of its autonomy with respect to the reducing theory, Nagel maintains that the reduced theory is not superseded or replaced by the reduction, that the phenomena described by the theory can still be regarded as real,[66] and notes that:

> Indeed, thermodynamics is still usually expounded as a relatively autonomous physical theory; and its concepts, principles, and laws can be understood and verified without introducing any reference to some postulated microscopic structure of thermal systems, and without assuming that thermodynamics can be reduced to some other theory such as mechanics (1961: 343).

Not only that, Nagel nowhere endorses the kind of dogmatic insistence on the possibility of reduction which Dupré makes his opponent. Whether a reduction is possible in any given case is, for Nagel, always a question specific to the particular states of the candidate reducing and reduced sciences, and never a matter for *a priori* resolution. The reason for this is simply that science, not to mention any individual science, has a history. 'Accordingly, the question whether a given science is reducible to another cannot in the abstract usefully be raised without reference to some particular stage in the development of the two disciplines' (1961: 361). Not only that, it is at least an open question how much scientific *value* any given reduction might have, since to count as a 'significant intellectual advance' reductions should suggest new lines of enquiry, perhaps also settling some outstanding questions, and so forth. This need not always happen, and Nagel goes so far as to argue that premature reductions might actually be counterproductive. In a passage which is especially striking in the context of a discussion of Dupré, Nagel states:

> Thus, a discipline may be at a stage of active growth in which the imperative task is to survey and classify the extensive and diversified materials of its domain. Attempts to reduce the discipline to another (perhaps theoretically more advanced) science, even if successful, may then divert needed energies from what are the crucial problems at this period of the discipline's expansion, without being compensated for be effective guidance from the primary science in the conduct of further research (1961: 362).

Even though the arch reductionist Nagel turns out to be considerably more cautious and nuanced than Dupré's attacks on reductionism might suggest, the Nagelian approach does have a range of limitations. These include the fact that it requires both the reducing and reduced theory to be broadly 'true' (at least in the sense of

---

66 '[I]f and when the detailed physical, chemical and physiological conditions for the occurrence of headaches are ascertained, headaches will not thereby be shown to be illusory' (Nagel 1961: 366).

accurately accounting for the observational data in the domain of the theory[67]). This is needed in order for the essentially conservative element of Nagel's view, where the reduced theory survives as a branch of the reducing theory once that theory is supplemented by the co-ordinating definitions, to be sustained.

The Kemeny and Oppenheim (1956) approach, and its descendant in Oppenheim and Putnam (1958) as well as more recent defenders of the unity of science such as Causey (1977) do not make use of a Nagelian model at all, and one crucial purpose of the micro-reductive approach is to avoid some of the difficulties in Nagel's version. Recall the discussion of micro-reduction in section (2.1) above. For a successful micro-reduction it is not, in general, necessary that the reduced theory be approximately true, indeed a micro-reduction could just as well take place where the reduced theory is known to be, or turns out to be, false. The key factor is that the micro-theory be able to account for the 'observational data' in the domain of the reduced theory, not that, given supplementary co-ordinating definitions, the reducing theory enables statements of the reduced theory to be *deduced*.

Reduction on this model is a different matter entirely from what Dupré chooses as his target, since a major portion of his attack is directed at the difficulty with coming up with suitable bridge principles. Since the Kemeny and Oppenheim (1956) and Oppenheim and Putnam (1958) approaches are designed in part to get away from the idea the reduction requires bridge principles at all, it is unclear why Dupré thinks that such an attack will do any damage to the Oppenheim and Putnam view.

Another way of making this point is to say that Dupré is ignoring the possibility of what is now called an *ontological* reduction, which says of some thing or phenomena that whatever it is, it consists of no more than something else. Any given configuration in *Life* is no more than an arrangement of occupied and unoccupied cells subject to the rules of *Life*. This does not mean that the possibility of any configuration can be predicted on the basis of these rules, or that classes of relevantly similar structure will share features describable in the terms suitable for describing single cells. That this is *not the case* is one of the points of Wolfram's account of cellular automata discussed above (Wolfram 1983, 1984, and section 5.1 above).

Thus it is that Klee (1984) mounts an argument against emergence which refers not to reduction or even micro-reduction, but rather to micro-determination, the thesis that the higher level properties of some thing or system are determined by the lower level ones,[68] with no direct implication that any bridge principles or other strict and general correlations might obtain. Smart marks the difference between ontological and theoretical reduction as follows: 'Physicalism as I have characterised it is a reductionist thesis. However, it is reductionist

---

67 Nagel notes that it is extreme to demand that either theory be true, but does require that 'the theoretical assumptions of the primary science be supported by empirical evidence possessing some degree of probative force' (1961: 358).

68 Klee (1984) frames his account of some of these issues in terms of 'micro-determination', rather than micro-reductionism, to mark this very point.

in an ontological sense, not as a thesis that all statements can be *translated* into statements about physical particles and so on' (1989: 81).[69]

The point which needs emphasising here, again, is that no claim about the failure of the search for bridge principles can count against the claim that an ontological reduction is possible, or that micro-determination in Klee's sense is the case. Furthermore, in the event that there is some independent reason to think that such a reduction is possible, all that the failure of Nagel style reduction suggests is that disorder is possible in the case in question, even where causal closure is also the case. If this point stands then from the perspective of the Completeness Thesis, no matter what Dupré might say, his attacks on reductionism are simply not relevant. With this point duly made and emphasised I can now turn to a brief examination of Dupré's treatment of the philosophy of mind.

7.2. Dupré on Functionalism, Eliminativism and Anomalism

The culmination of Dupré's treatment of reductionism is his account of the philosophy of mind (1993: 146-167).His account of the philosophy of mind is, it must be said, selective in a fairly bizarre way. He disposes of functionalism in a single paragraph, noting that it is 'open to debate' (1993: 147) whether the functionalist line has much force against more overtly reductive programmes, and that there are those (e.g. Lewis 1966) who use functionalism to argue for a more reductive form of identity.[70] Following that he devotes a few pages (1993: 148-159) to criticism of Churchland style eliminativism,[71] which criticisms by and large take the form of typical functionalist objections, urging the importance of what things do, noting that the object of a psychological theory is not usually framed in physical terms, and insisting that the 'role in our intellectual economy' of the things detailed in psychological theories has little if anything to do with the physical aspects of these things, even if they are physical (1998: 152, 153, 154). Finally (1993: 159-167) he turns on 'token physicalism'[72] of which he takes Davidson's anomalous monism as representative.

Functionalism is indeed open to debate, a point to which I return shortly. That aside, it would take some genuinely impressive arguments, arguments not to be found in Dupré, to defend the assertion that the

---

69 See also: 'What it is for chemistry to be ontologically reducible to physics, even though one cannot deduce all statements of chemistry from statements of physics is this: any model of a set of true sentences of chemistry will be a model of a set of true sentences of physics' (Smart 1989: 82)

70 See also Kim (1998) who now inclines to a more Lewisian approach after having spent a long time defending various forms of 'nonreductive physicalism'. I return to Kim's recent work briefly in the conclusion to this thesis.

71 Dupré focuses largely on Churchland, P. M. (1981, 1985).

72 Since I'm harping on Dupré's failure to consider functionalism seriously enough, I cannot resist pointing out that this is precisely the term used in Fodor (1974) in defence of a position which is pointedly at odds with anomalism.

Completeness Thesis should stand or fall with the fates of either eliminativism or anomalous monism,[73] neither of them spectacularly popular positions within the ranks of physicalists. Even though Davidson (1970) clearly endorses a form of the Completeness Thesis, his insistence, at least in that text, on both the character of physical laws as being strictly deterministic, and if not strictly deterministic at least establishing strong connections between events, and on the impossibility of laws concerning the mental are far from standard physicalist fare. There are many different ways of thinking of physical causation (the present work has as one of its aims the articulation and defence of a strikingly un-Davidsonian one) and a great deal to commend the view that predictions and some form of nomological understanding are possible regarding the mental, and the actions of intentional beings.[74]

Hellman and Thompson's (1975) account of physicalism is worth bearing in mind here. As noted in passing above (section 5) their version of that doctrine, another micro-determination theory, makes clear that '[p]hysicalism in no way dictates the course of progress in the higher level sciences' (1975: 564) and that '… physicalism without reductionism does not rule out endless lawful connections between higher-level and basic physical sciences' (1975: 552).[75] A carefully understated footnote to the second quotation here observes that this puts their position 'at odds' with anomalous monism. But if this is a perfectly acceptable form of physicalism, which it is, and entirely committed to what I am calling the Completeness Thesis, then it becomes even more clear how Dupré's position relies on the misidentification of the target for his criticisms, which comes to look like more and more of a straw man:

> Put at its simplest, I suggest that the failure of reductionism may be attributed to the following fact. The individuals that would have to be assumed for the derivation of the macrotheory cannot be identified with those that are the subject matter of the descriptive theories at the next lower level, though their relationship to the actual individuals may still be close enough to allow such a derivation to function as an explanation. The possibility of this non-identity is then to be explained by the fact that the individuals at both levels are idealisations. Both models at the macrolevel and descriptive laws at the microlevel are, in the sense explained earlier, abstract. But the abstractions involved are not the same (Dupré 1983: 333).

---

73 Over and above Dupré's objections, it has long been recognised that Davidson's anomalous monism is both dangerously close to epiphenomenalism, that it gives us too little in the way of law like relations outside physics, while also expecting too much from the laws of physics. See Evnine (1991: 160f) and (Kim 1998: 32-35, 57-8).

74 This is something which Dupré (1996) is correct about: some things tend to get more predictable the more structured they are.

75 Kincaid (1990) applies the Hellman and Thompson approach to molecular biology, arguing the error of traditional reductionism is to suppose that 'one theory can do the work of another' (1990: 576) but nonetheless keeping this issue clearly separated from that of micro-determination, and also the question of physical closure.

But who, except Nagel whose version of reductionism is both optional and has furthermore been superseded by a variety of positions not sharing that limitation, would think that the abstractions would co-vary in the ways Dupré is adamant that they won't? This insistence on the disjunction between the classifications and abstractions at different levels is suggestive of functionalism, and it is time to consider the question of the relationship between Dupré's views and functionalism.

As noted above Dupré's engagement with functionalism seems woefully inadequate, although it is worth noting that he is prepared to make use of relatively functionalist arguments to get to points that he himself wishes to make (e.g. 1993: 127). Not only that, despite his occasional protests against functionalism, his own position has a great deal in common with that view. I want to argue here that while it is reasonably clear that Dupré is not a functionalist, and that he would resist being classified as one, the argumentative resources with which he defends promiscuous realism are ultimately insufficient to prevent his position from collapsing into some form of functionalism, since the kind of reduction which functionalism has recently been taken to suggest is not one with which Dupré has engaged, and he has no weapons against it.

Davies (1996) notes the striking similarities between Dupré's (1993) argument for promiscuous realism and Fodor's (1974) treatment of the special sciences, saying that 'one might view Dupré's argument as a more radical version of Fodor's, aimed at establishing pluralism *within* scientific domains where Fodor argued for a measure of pluralism *across* such domains' (1996: 6). There is something profoundly correct about this as Davies' (1996: 6-8) survey of the similarities between Fodor's and Dupré's positions makes clear. As noted above, in section (2.1) Fodor begins his discussion of the special sciences by noting that reductionism and what I am calling the Completeness Thesis are often perceived as mutually supporting. This is how Oppenheim and Putnam see things, as does Dupré. Fodor opposes this, though, arguing as I have already here that the dependency is not symmetrical, and that what has 'traditionally been called 'the unity of science' is a much stronger, and much less plausible, thesis than the generality of physics' (1974: 51). Fodor also makes clear that he wants his point to be understood as relating to the question of natural kinds.[76] He notes, similarly to Dupré's account of essentialism about kinds (see 3.2 above), that reductionism would require 'that *every* natural kind is, or is co-extensive with, a physical natural kind. (Every natural kind *is* a physical natural kind if bridge laws express property identities, and every natural kind is co-extensive with a physical natural kind if bridge laws express event identities.)' (1974: 55). And later says that 'this consequence of reductivism is intolerable' for the following reasons:

> The reason it is unlikely that every natural kind corresponds to a physical natural kind is just that (a) interesting generalizations (e.g. counter-factual supporting generalizations) can often be made about

---

[76] For the purposes of this discussion Fodor regards the natural kind predicates of a science as 'the ones whose terms are the bound variables in its proper laws.' (1974: 54-5).

> events whose physical descriptions have nothing in common, (b) it is often the case that *whether* the physical descriptions of the events subsumed by these generalizations have anything in common is, in an obvious sense, entirely irrelevant to the truth of the generalizations, or to their interestingness, or to their degree of confirmation or, indeed, to any of their epistemologically important properties, and (c) the special sciences are very much in the business of making generalizations of this kind (1974: 55).

Through all this Fodor fully endorses what he calls 'token physicalism', i.e. the view that any single event in a special science is always identical with an event which has a physical description. This means, via his endorsement of the *generality* of physics (1974: 51), that he has no argument with the Completeness Thesis. Fodor argues, though, that the generalisations and laws of a special science are likely not to conform to anything which would count as a generalisation of physics. So, for example, he considers a candidate law of economics, Gresham's law, holding that bad money drives out good, and considers the description of a set of physical events token identical with economically identified events conforming to the law:

> But banal consideration suggest that a description which covers all such events must be wildly disjunctive. Some monetary exchanges involve strings of wumpum. Some involve dollar bills. And some exchanges involve signing one's name to a check. What are the chances that a disjunction of physical predicates which covers all these events (i.e. a disjunctive predicate which can form the right side of a bridge law of the form "x is a monetary exchange [□][77] ...") expresses a physical natural kind? In particular what are the chances that such a predicate forms the antecedent or consequent of some proper law of physics? (1974: 56)

Fodor's claim is that functionalism can save the autonomy of the laws and kind-classifications of the special sciences, *without* having to abandon the generality of physics, i.e. the Completeness Thesis. The kinds of disjunction between higher level laws and taxonomies are basically the same as Dupré's, they are the kinds of thing which make general or type-type reductionism wildly implausible, but then Fodor's position is not reductionist. As he makes clear 'token physicalism is weaker than reductivism. [...] reductivism is the conjunction of token physicalism with the assumption that there are natural kind predicates in an ideally completed physics which correspond to each natural kind predicate in any ideally completed special science.' (1974: 53-4)

A question arises quite naturally here: Why is Dupré *not* a functionalist of some sort? There are two reasons for this, one that he is generally opposed to any kind of theory which enshrines causal closure, of which token physicalism is an example. Another is his noted concern that functionalism can be pressed towards some kind of reductionism. The two are related. Dupré rejects causal closure because he thinks that it goes hand in

---

[77] In Fodor's original this symbol appears as a pair of horizontal arrows facing in opposite directions. Causey (1977) notes that it is unclear how if at all Fodor's symbol is supposed to differ from the more conventional symbol for a biconditional, which I have used here.

hand with reductionism, and his vision of reductionism is of a massively implausible super-Nagelianism. Functionalism grants causal closure, which is reason enough for Dupré to reject it, but it has also been used in arguments for a degree of reductionism, which makes it quite anathemic from Dupré's point of view. Dupré's point of view is defective on this point, though, since the form of reduction for which one can argue on the basis of functionalist considerations is not the kind of super-Nagelianism which worries Dupré.

Recall that Dupré claims that Lewis (1966) 'uses functionalist considerations to argue for an identity theory explicitly committed to classical reductionism' (1993: 167). This is a somewhat bizarre assertion. If 'classical reductionism' means anything like the kind of Nagelian bridge-principle approach which Dupré makes his main opponent[78] then his claim about Lewis is simply false. For a start Lewis (1966) takes account of the fact of multiple realisation (starting from 'functionalist considerations' it's difficult to see how he could fail to) which already places him beyond the pale of the kind of reduction based on bridge principles. Not only that, his theory is an *identity* theory, and identity theories have no need of biconditional bridge laws at all. Staying with Nagel's example, where Nagel would say that temperature is *related* to mean kinetic energy of molecules, an identity theorist would say that temperature *is* mean kinetic energy of molecules. Identities don't need biconditionals to hold them up. The reductionism which Lewis defends is based on multiple realisation, and has it that where functional properties of some sort are realised in the same way, then a reductionistic explanation should be possible, no more and no less.[79] And this kind of reduction is not one which Dupré has even granted the possibility of, let alone given any reason for rejecting, [80] which is to say in sum that his case is damagingly incomplete.

I should make quite clear that I am not setting out to defend functionalism here. I have no particular reason for doing so, and in any event that is well outside the brief of my attempted defence of the Completeness Thesis. The point is merely that there are well-developed and significant physicalist approaches which both entrench the Completeness Thesis and allow all the heterogeneity and diversity which is so important to Dupré. Those versions of these approaches which incline to any degree of reductionism do so in ways simply not covered by Dupré, and the existence of these views shows again that whatever follows from Dupré's arguments, it is not the implausibility of the Completeness Thesis.

I would like to complete the argument of the present chapter by returning to a small point noted in section (2.1) regarding Dupré's image of reductionism.

---

78 If this is *not* what Dupré means, then his (1996) gives no indication of how else to understand 'classical reductionism'.

79 See also Lewis (1980) and Kim (1998).

80 There are a number of difficulties with this approach, some of which are discussed in the conclusion of this thesis.

7.3. Stamp Collecting

For Dupré the expectation that science could be unified would 'require the extremely incautious prediction that we will at some point cease to acquire any new interests' (1983: 321). I have already noted that this extraordinary claim is not one which I can find any supporter of the unity of science endorsing, or indeed even suggesting. Nonetheless it might seem as though were a completed physical science to be achieved, it would follow in the ways expected by Oppenheim and Putnam that all questions would have been thereby shown to be physical ones, and in principle amenable to physical treatment.[81]

Rutherford perfectly captured the stupendous arrogance apparently licensed by the foundational position of physics with his claim that 'All science is either physics or stamp collecting.' In other words, there is, on the one hand, the proper study of real, fundamental, causal relationships, and then there are a whole range of other ways of describing things, which, in virtue of not being fundamental, ultimately amount to expressions of collector idiosyncrasy and preference. One way of saying what Dupré's position is about, is that he does not want to allow non-physical enquiries to take on the status of stamp collecting. The arguments above, especially those of section (5) should show fairly clearly how no such consequences follow from the Completeness Thesis in the event that it is true. The discussion of functionalism just offered reinforces the point in the context of a particular body of physicalist theories.[82]

Although the point may well not stand in need of pressing, I would like to conclude this line of argument by noting that even in mathematics, where there is a widespread consensus that a single unifying theory, namely set theory, has been found which embraces all of mathematics, there is no question at all either that there is no more work to be done, or that significant new relationships between higher level mathematical constructions cannot be found, even though all of them are thought of as in one sense no more than set theoretical entities. Marquis (1998) considers a range of epistemological questions related to the classification of particular mathematical theories and constructs. He notes in ways which are entirely congenial to Dupré's treatment of other sciences that a major problem for many of these theories is one of classification. He notes that even if from the point of view of set theory all sets are 'on an equal footing', it is also the case that 'from the point of view of particular disciplines, not all sets are equally important and significant: the universe of sets has to be carved in ways which are *not* set-theoretical in general. The tools required to carve the set-theoretical world vary from one discipline to another' (1998: 189). In case there be any doubt that the fact of the set-

---

81 If the succession of reductions required to get from psychology to fundamental physics is achieved, 'then psychological laws will have, *in principle*, been reduced to laws of atomic physics, although it would nevertheless be hopelessly impractical to try to derive the behaviour of a single human being directly from his constitution in terms of elementary particles' (Oppenheim and Putnam: 1958: 7)

82 The 'assumption that every psychological event is a physical event does not guarantee that physics (or, *a fortiori*, any other discipline more general than psychology) can provide an appropriate vocabulary for psychological theories' (Fodor 1974: 57).

theoretical underpinning licence any general reductionism he is careful to point out that: 'One thing is clear: these classifications are not reducible to one another in the same way that differential and algebraic considerations are not equivalent in topology. What is still not clear is how these classifications should be related, or whether they should be related at all' (1998: 190).

Recall that Dupré holds to the view that failures of practical reductionism diminish the plausibility of what he calls 'theological reductionism' – the notion that an epistemologically ideal entity could perform the reductions of which we are incapable (1983: 323, 343-345). Dupré maintains that the only plausible support for the notion that theological reduction, or any reduction 'in principle' is possible must be that enough successful reductions in fact have been performed, that 'the evidence for this assumption must be that the behaviour of complex systems can in fact be shown to be derivable from laws governing their constituents' (1983: 345). But even where we have a full account of the behaviour of the constituents, in the amazingly simple yet endlessly surprising artificial world of the cellular automata, no such derivations are generally possible, and there are good reasons why this is so. Having reason to believe that a given system is causally closed does have some effects: in the *Life* case, for example, it makes clear that there is no question which needs answering about where spaceships get their patterns of behaviour and interaction from – they do so from their well understood constituents. But the knowledge that *Life* worlds are causally closed cannot tell us everything they will do, and it cannot permit the real work of classification, analysis and all of the other business of science to be pre-empted. Dupré's worry about theological reductionism is another red herring – no scientist is condemned to being a mere 'stamp collector' in the event that the Completeness Thesis is true.

8. Conclusion

The simplest way to state the conclusion of the present chapter is to say that although Dupré's *conclusions* are incompatible with the Completeness Thesis, neither his evidence nor his arguments present any significant difficulties whatsoever. More specifically all of the forms of epistemic disorder which he draws our attention to are entirely acceptable from the point of view of the Completeness Thesis. There is just nothing there which decides against the completeness of physics, or which ultimately even allows his position to be distinguished from that of one who endorses completeness, most obviously in the case of functionalism.

This does not, and cannot, show that the Completeness Thesis is true. What it shows, rather, is that no argument from disorder at the top can be relied upon to count against the general thesis that there is order at the bottom. Rather, the effect of this brief discussion has been, I hope, to show that disordered phenomena at non-fundamental scales can be saved by defenders of the Completeness Thesis and promiscuous realists alike, which means that the proper way to proceed is directly to consider the issue of order at the bottom. This is the burden of the argument in the following chapter.

Another important conclusion of all this is that no plausible version of the Completeness Thesis can be associated with a general programme of theoretical reductionism. Even in cases where a micro-reduction is defensible, there is no automatic inference to a manageable theoretical reduction. Not only that, given the degree of autonomy of special enquiries, there is no reason why their implied ontologies should not cross-classify with those of other enquiries. Even while recognising a distinction between fundamental and pragmatic ontologies, we need not take that recognition to justify any commitment to the unity of science on a mid-century model.

Given that I have already offered independent arguments against the plausibility of emergentism, it might seem as though it would be possible to stop here: emergentism poses no significant threat to completeness, and neither does Dupré's case. Both the argument against emergentism, and that against Dupré, though, rely upon a fairly traditional image of the best way to understand the laws of nature, and especially the laws relating to the action of fundamental physical forces. This understanding has been radically challenged by the work of Cartwright, and unless her arguments can be answered, the arguments of this and the preceding chapter are in serious trouble. It is to the consideration of Cartwright that I now turn.

Chapter Four: Cartwright and Patchwork Realism

> '*We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances*. To this purpose the philosophers say that Nature does nothing in vain, and more is in vain when less will serve; for Nature is pleased with simplicity, and affects not the pomp of superfluous causes.' – Newton, *Principia Mathematica*.[1]

1. Introduction

In the preceding chapter Dupré's line of attack was seen to founder on the fact that disordered phenomena cannot themselves be decisive against the Completeness Thesis. The argument developed there did not point to the conclusion that disorder was in fact best explained by reference to a uniform fundamental description, although some difficulties with Dupré's promiscuous realism were noted, pointing to the need for a more minimalist ontology. In order to complete the argument against the disorder approach it is necessary to say something about why the view that there is what I have been calling 'order at the bottom' is preferable to the view that there is not. To make this argument requires me to engage with the work of Cartwright.

From the point of view of my present interests, Cartwright's main opponent is a doctrine which she calls 'fundamentalism' (Cartwright 1994), the view that there is some small set of fundamental laws with general validity at work behind all phenomena. There is more to be said about fundamentalism, and more than one way of thinking of the Completeness Thesis. Even so, the Completeness Thesis is naturally enough seen as a version of fundamentalism, holding that the fundamental laws are all physical, and furthermore that they form a causally closed set, i.e. that they are complete in the sense set out in Chapter One above.

If Cartwright's assault on the credentials of the fundamental laws of physics is defensible, then this would have the effect of seriously undermining or even destroying the acceptability of fundamentalism and hence of forms of the completeness thesis tied to fundamentalism about laws. I propose to give reasons for rejecting that line of thinking. My arguments here, since they do not, by themselves rule out emergentism, do not amount to a defence of the completeness of physics, although they make a contribution to that project. If, though, the arguments against emergentism in Chapter Two above hold water, then the fact that the present discussion is focussed specifically on a defence of what I will call moderate fundamentalism need not present

---

[1] This text is the second of Newton's 'Rules of Reasoning in Philosophy.' The first two of these rules appeared in the first (1686) edition of the *Principia* but not as 'Rules', appearing in Book 3 under the heading 'Hypotheses' along with some other material. In the second and third editions (1713 and 1726) the Rules appeared under the correct heading, and with each edition Newton added a further rule. This quotation is from Newton (1934: 398).

any difficulties. If Cartwright is correct then emergentists and fundamentalists alike are in trouble, if I'm right about emergentism and Cartwright, then the Completeness Thesis is what we're left with.

I should note that there is much about which Cartwright and I are in agreement, most notably with reference to the untenability of what she calls the 'facticity' view of laws, a position similar to what Bhaskar calls actualism, [2] and on the matter of realism. Although some of Cartwright's arguments, especially the earlier ones, are framed in terms of realism and anti-realism, she is committed to realism about the entities and causal capacities which feature in successful causal explanations. Where we differ is, for the most part, over the question of the status of fundamental laws, and the crucial issue of their generality. What follows in the present chapter is, then, a dispute within realism,[3] over the sorts of things about which it is reasonable to be realist, about which laws should be regarded as true, and what kinds of truths we take these laws to state.

2. A selection of worlds

What is at issue here can be made clear by considering three competing images of the world in which we find ourselves. The images in question are those of the fundamentalist, the emergentist, and what, for reasons to be explained shortly, we can call the patchwork realist. The fundamentalist holds that there is some finite set of laws (however conceived) which can account for all phenomena, everywhere. A simple illustration of the spirit of fundamentalism can be seen in the wording, especially the first two words, of Newton's first law of motion, the principle of inertia, which reads, in English, 'Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it' (1934: 13). That's the kind of law fundamentalists like: ones which, supposedly, apply to every entity, event or process of some kind, in Newton's case to bodies with nonzero masses.[4] A natural way to understand the view that physics is complete is to read it as a claim about physical laws. Note the prominent place of laws in Papineau's formulation quoted near the beginning of Chapter One above (Papineau 1993: 16). In this case the thesis that physics is complete can be seen as a species of fundamentalism which holds either that there is some subset of the fundamental laws which account for all physical effects by reference to only physical factors, or, more simply, that all fundamental laws are physical, and form a causally closed set. This is not the only way to understand the completeness of physics, but it is the way most relevant to the bulk of the argument of the present chapter, so in what follows when I refer to fundamentalism I will have in mind a kind of fundamentalism which also involves commitment to the completeness of physics. Where there is a danger of

---

2 See the account of Cartwright on Laws in section (3.1) below, Cartwright's treatment of facticity is criticised in section (5.2). Bhaskar's philosophy of science is discussed in Chapter Two section (4) above.

3 This means that it is not part of my task here to defend realism against, for example, the type of challenge mounted by van Fraassen (1980).

4 I am not claiming that Newton himself was a fundamentalist in the sense at issue here, merely that his work includes examples of laws of the fundamental type.

ambiguity I will refer to 'physical fundamentalism'. It is worth noting, though, that physical fundamentalism makes a stronger claim than the completeness thesis, since physics could be regarded as complete without any endorsement of fundamentalism being required for the simple reason that the completeness thesis on its own does not entail any particular commitments about laws. It is a clear enough consequence of this that damage to fundamentalism need not necessarily harm the Completeness Thesis.

As noted in Chapter Three above (section 1), Cartwright describes the world of the fundamentalist as one where all the facts 'belong to one grand scheme' and hence where facts 'legitimately regimented into theoretical schemes' have a special status, but contends that this doctrine is one 'we must resist' and later offers an image of a patchwork world consisting of 'tens of thousands of patches, cut up in no particularly logical way, exhibiting tens of thousands of different regularities of countless different forms' (1994: 281, 298). On the view Cartwright is defending there just is no set of fundamental laws of the '*every* body' type — merely local, even transitory, regions of particular kinds of order. In the essay just quoted Cartwright uses the image of a patchwork of laws, a view she calls 'nomological pluralism' (1994: 288). The corresponding ontology is also pluralist, i.e. she urges that the world described by a patchwork of laws should be seen as itself a patchwork. Here Cartwright is following the principle that we 'best see what nature is like when we look at our knowledge of it' (1983: 13). Her position is in important ways similar to that of Dupré who, in *The Disorder of Things*, argues for the need to recognise the 'underlying ontological complexity of the world' (1993: 7) and, as noted, dubs his preferred view 'promiscuous realism'. Although the two positions are similar in some respects, and also related, what is important about Cartwright for my purposes is her account of laws. To make clear that this is largely independent of the considerations arising from Dupré's assault on unity, I will call Cartwright's position 'patchwork realism.'

What has been said here does not exhaust the options, though. An important intermediate position on these questions is emergentism, discussed in Chapter Two above, and which (in some versions of that doctrine) sees a hierarchy of kinds of law, typically starting with physical laws as a foundation and including, for example, biological laws for those material entities which are alive, psychological laws for those living beings which are conscious, socio-economic laws for those conscious beings which associate and exchange in certain ways, and so-on. There are many examples of emergentist theories but for my purposes here I only need to note that at least some emergentists fall, in a sense, *betwæn* physical fundamentalism and promiscuous realism. The type of emergentist ('strong' rather than 'weak'[5]) I am thinking of here would hold that physical laws are indeed generally applicable in some sense, but also that they fail completely to fix the likelihood even of certain physical outcomes which are rather co-determined by both physical and non-physical laws. As we saw Bhaskar, for example, holds that some basic physical laws such as that of the conservation of mass-energy can

---

[5] Recall the discussion of strong and weak emergentism in Chapter Two section (3.1).

be held to be true, but argues that such laws do not by themselves decide what will happen, only place restrictions on what can happen (1977: 109).[6]

The in-between quality of emergentism can be seen from the ways in which it might appear offensive to either physical fundamentalists or promiscuous realists. From the perspective of the committed physical fundamentalist, emergentism seems like an unacceptable compromise position, raising too many difficult questions about the status of the 'emergent' laws, the empirical content of the claim that there even are such laws, and the proper account of the conditions under which they might be expected to arise. From the point of view of the fundamentalist who thinks that physics is complete, emergent laws are unnecessary, and also undesirable since they suggest the overdetermination of at least some physical effects.[7]

On the other hand a serious patchwork realist is likely to regard emergentism as similarly obnoxious for conceding too much to the fundamentalist, and creating a position amounting to a kind of hierarchical fundamentalism. Cartwright, for example, protests against the emergentist that the view that macro-properties 'come out of nowhere' falls afoul of the fact that there is 'nothing of the newly landed about these properties' (1994: 290). She is willing to grant that her position has at least some common ground with emergentism:

> Reductionism has long been out of fashion in biology, and now emergentism is again a real possibility. But the long-debated relations between biology and physics are not good paradigms for the kind of anti-fundamentalism I urge. Biologists used to talk about how new laws emerge with the appearance of 'life'; nowadays they talk, not about life, but about levels of complexity and organization. Still, in both cases the relation in question is that between larger, richly endowed, complex systems, on the one hand, and fundamental laws on the other: it is the possibility of 'downwards reduction' that is at stake (1994: 281).

She also, though, makes clear that she wishes to go beyond the simple rejection of downwards reduction typical of emergentism, and assault the possibility of 'cross-wise reduction' (1994: 281) in a way which directly undermines the defensibility of the view that there even are nomological or ontological foundations of a sort congenial to either physical fundamentalism or emergentism. And emergentists of the sort at issue here, just like promiscuous realists, would deny that physics is complete, and hence have common foes in physical fundamentalism and the completeness thesis.

The point of this brief survey, getting these three ontological options on the table, is to make clear what at least one of the issues hanging on the status of fundamental laws is. If such laws are philosophically

---

6 Bhaskar's emergentism is discussed in Chapter Two above, especially section (5).

7 For example Kim (1992) who, as usual, takes the Completeness Thesis as a premise and hence regards emergentism as a species of supervenience, prone to the danger of epiphenomenalism (the diagnosis of which occupies much of Kim's recent work). The notion of strong emergence plays no significant role in Kim's account.

respectable then we seem to have a choice between emergentism and physical fundamentalism, perhaps even in the extreme form of the completeness of physics. If, though, fundamental laws lack the required support then, to the extent that we are realists, we had better be patchwork realists.

In the terms adopted in Chapter Three above, what Cartwright calls fundamentalism is a version of the 'order at the bottom' ideal. As also noted in that chapter, there are at least two likely lines of attack against the fuller 'order at the top: order at the bottom' expectation: the use of disorder at the top to try and block any inference to order at the bottom, and direct attack on order at the bottom. Order at the top is not essential to the Completeness Thesis, and is in any event implausible. (Dupré is correct about at least the latter point.) The separation of order at the top from completeness was the burden of argument of Chapter Three. Although, as we will see, some relaxation of and modification to the order at the bottom ideal is called for, my principal task here is to defend the plausibility of order at the bottom against Cartwright, who mounts a direct assault on it. The position I defend here is not full blown fundamentalism – Cartwright has successfully created a view which no-one in their right mind would endorse there, and I for one am uncertain whether any historical individual in the history of the philosophy of science is a fundamentalist of the sort Cartwright sets out to refute. Rather I defend what I call moderate fundamentalism, which involves salvaging a reasonable measure of truth for fundamental laws, but steers clear of the excesses of the facticity view. The distinction between moderate fundamentalism and the extremist version become clear in what follows.

3. Cartwright's Philosophy of Science

The broad outlines of Cartwright's attack on fundamentalism are relatively well known. In *How the Laws of Physics Lie* (1983) and *Nature's Capacities and Their Measurement* (1989) Cartwright develops powerful arguments against the view that we should be especially impressed by theoretical physical laws, and fundamental laws more generally. The first of these two works was sometimes seen as an attack on realism, but Cartwright prefers to put things another way:

> A number of years ago I wrote *How the Laws of Physics Lie*. That book was generally perceived to be an attack on realism. Nowadays I think that I was deluded about the enemy: it is not realism but fundamentalism that we need to combat. (1994: 279)

In the necessarily compressed treatment of her views in what follows I will read Cartwright's work in the spirit suggested by these remarks: as a connected and evolving critical engagement with what she now calls fundamentalism. This approach works well for my purposes, since from the perspective of the Completeness Thesis Cartwright's work can be seen to run continually, with developments and refinements, from the essay 'Do the Laws of Physics State the Facts' (1980) up to the present.

It will be worth looking at a few relatively recent statements by Cartwright which show the relevance of her work for the present investigation especially clearly, before looking at how her position has been developed over the years. A sense of what is at stake in her talk of fundamentalism can be discerned in the following remark, made just after distinguishing two sorts of law-like generalisations, in the first place those 'legitimately regimented into theoretical schemes' and secondly those not so regimented:

> There is a tendency to think that all facts must belong to one grand scheme, and moreover that this is a scheme in which the facts in the first category have a special and privileged status. They are exemplary of the way nature is supposed to work. The others must be made to conform to them. This is the kind of fundamentalist doctrine that I think we must resist. (1994: 281)

This passage is not specifically about the notion of the completeness of physics, but it is simple enough to see how the view that physics is or could be complete could very easily be an example of the kind of fundamentalism Cartwright is describing.[8] The consequences of her position as far as the Completeness Thesis goes are made strikingly clear when she describes her hypothesis of the 'patchwork world' which comprises 'tens of thousands of patches, cut up in no particularly logical way, exhibiting tens of thousands of different regularities of countless different forms' (1993: 298). It is worth noting that there are significant similarities between Cartwright's view and Dupré's, discussed in the preceding chapter, with his talk of the 'underlying ontological complexity of the world' (1993: 7).

Even this type of world is, it must be noted, at least potentially compatible with the completeness of physics as long as either all of the patches are themselves physical, or if throughout the patchwork physical occurrences have sufficient physical antecedents. But this is, *prima facie*, a far from comfortable position for the defender of completeness to be in, since the notion of physics itself as a coherent science or body of sciences is at least threatened by the patchwork image, and there is in addition some reason to suppose that Cartwright would resist such a construal of her position. In at least one place she explicitly refers to the idea of the completeness of physics, and then it is as a notion by which we might be 'misled' (1994: 290). I return to these issues in my discussion of the consequences of Cartwright's work for the Completeness Thesis in section (4) below.

3.1. Laws and Lies

In her first book Cartwright advanced the *prima facie* sensational thesis that the laws of physics are lies. More specifically she adopted and modified an existing distinction between theoretical laws, which are usually very abstract and general in form, and phenomenological laws, which are typically adapted to fit specific and

---

[8] I note in passing that the suggested requirement that all facts should be 'made to conform' to the regimented ones is reminiscent of Dupré's anxiety about the supposed reductionist implications of causal closure. I take it that enough has been said about this issue in the preceding chapter.

real situations in the laboratory or in engineering work. Conceding that theoretical laws can have considerable explanatory power, she nonetheless reverses the traditional epistemological ranking of the two types of laws, favouring the empirically more successful and accurate phenomenological ones over the general and less empirically successful theoretical laws. She notes that 'paradoxically enough the cost of explanatory power is descriptive adequacy. Really powerful explanatory laws of the sort found in theoretical physics do not state the truth' (1983: 3). Phenomenological laws may have limited explanatory power, and also stand in no neat relationship with theoretical laws, but they can, in some cases, state the facts. Theoretical laws, whatever their explanatory power, cannot state the facts, and should hence be seen as 'lies'. She notes that her arguments for this conclusion can be separated into three main strands:

> (1) The manifest explanatory power of fundamental laws does not argue for their truth.
>
> (2) In fact the way they are used in explanation argues for their falsehood. We explain by *ceteris paribus* laws, by composition of causes, and by approximations that improve on what the fundamental laws dictate. In all of these cases the fundamental laws patently do not get the facts right.
>
> (3) The appearance of truth comes from a bad model of explanation, a model that ties laws directly to reality. As an alternative to the conventional picture I propose a *simulacrum* account of explanation. The route from theory to reality is from theory to model, and then from model to phenomenological law. The phenomenological laws are indeed true of the objects in reality–or might be; but the fundamental laws are true only of objects in the model (Cartwright 1983: 3-4).

The argument that the explanatory power of fundamental laws does not sustain any claim for their truth points out that, taken as descriptions, i.e. as accounts of the facts, fundamental laws are manifestly false in all or most situations. Given this it just does not help matters to appeal to the explanatory power of such laws: that power, according to Cartwright, is clearly independent of their possible truth, which must be separately evaluated in each case. (I return to the question of what exactly is at stake in Cartwright's claim that fundamental or theoretical laws do not 'state the facts' shortly.)

The second argument complements the first by analysing the role played by fundamental laws in explanations, and pointing out that real explanations involve major modifications to fundamental laws, including *ceteris paribus* clauses and the use of various other additional features, all of which are needed for descriptive accuracy. Were fundamental laws descriptively powerful to start with, i.e. were they 'true', none of this would be necessary, and the fact that it is indeed necessary shows quite straightforwardly the descriptive poverty of fundamental laws.

The final strand of argument completes the case in an important way, since the first two arguments alone leave open the possibility that the committed fundamentalist will just hold out for better fundamental laws in the future. Even if the explanatory of power of any given fundamental law does not argue for its truth, and,

furthermore, if present fundamental laws can be shown to be false on the basis of the ways they have to be modified to play a role in explanations, it surely does not follow that no possible fundamental law could be true. The simulacrum account of explanation meets this possible counter by arguing that the descriptive weakness of fundamental laws is something that they more or less have to possess in order to do the kind of explanatory work they do.

The argument for this conclusion begins (Cartwright 1983: 143-4) with a nod at Kuhn,[9] who argued that certain general features of scientific practice could serve the function of keeping what would otherwise be rather disconnected and isolated enquiries in some kind of cohesion. Cartwright suggests that this might make sense of why fundamental laws tend to be few in number and also, in her sense, false. Thus the simulacrum account of explanation has it that fundamental laws play very much this role – providing a standard and partly unifying vocabulary enabling scientists to impose a kind of coherence on their heterogeneous rabble of phenomenological models. Although she sees fundamental laws as giving this kind of coherence, Cartwright is careful to make clear that she is generally opposed to any attempt to regard such laws as true, saying that 'it usually does not make sense to talk of the fundamental laws of nature playing out their consequences in reality' (1983: 160).

It is worth noting briefly that the conclusion of the simulacrum account is itself a *non sequitur*. It does not follow from the falsehood (if that line of thinking is defensible at all) of some fundamental laws that all are or have to be false, and it certainly does not follow from the fact that Cartwright simply has a story to tell about fundamental laws that they are all false either. It is one thing to shift the burden of argument onto the fundamentalist or other defender of fundamental laws, but another entirely to require that accurate laws actually be produced in order to justify the belief that such laws could even possibly be true. I return to this shortly, in section (5) below.

In review, then, Cartwright's position is that theoretical laws are very bad at stating the facts, tend to be formulated in terms of various conditions and *ceteris paribus* clauses, and should therefore be seen as lies. This does not make them completely useless, though, since their empirical bankruptcy is also the condition for their explanatory richness. This richness is bought by means of theoretical laws being relatively simple, and also relatively generic, for example the idea of a harmonic oscillator, which is used in cases where it is at least far from clear what could possibly be oscillating in the event that the theory was to be interpreted realistically.[10] This is the basis of the simulacrum account of explanation: relatively simple theoretical constraints impose a coherence and manageability on the possible proliferation of models, but at the expense of empirical accuracy.

---

9 Cartwright refers to Kuhn (1977: 220).

10 This is one of Cartwright's examples, see (1983: 145).

The models can buy greater accuracy, but at the expense of moving away from the simplicity and generality of theory towards the particularity of the real situations being modelled.

Although the bulk of my engagement with Cartwright focuses on those features of her thinking just outlined, it will be worth spending a little time clarifying (a) her realism about entities and capacities, and (b) the extent to which she is prepared to push her anti-fundamentalism. The first is useful because following my criticism of her account of laws I also offer a reformulation of the Completeness Thesis in terms of capacities, while the latter is important for making clear the ways in which her position threatens the Thesis.

3.2. Realism about Entities and Capacities

As noted it would be a mistake to regard Cartwright's position as generally anti-realist. I need briefly to clarify the ways in which her position and mine are not in conflict, so as to make clear precisely what it at issue between us.[11] Cartwright's realism may (with one crucial qualification to be noted in the following subsection) usefully be compared with the position of Hacking (1983). *How the Laws of Physics Lie* appeared in the same year as *Representing and Intervening* and Cartwright herself describes her book as a 'complement' to Hacking's (Cartwright 1983: 20).[12]

What both works share is a commitment to a pro-experiment and anti-theory stance in the philosophy of science, far more willing to endorse realism about the entities and processes referred to by the working experimenter than those of the theoretician. Hacking's dictum 'if you can spray them then they are real' (1983: 22f)[13] has a clear counterpart in Cartwright's assertion that when explaining a change in the observed processes in the same kind of experiment, 'I am inferring from effect to cause, and the explanation has no sense at all without the direct implication that there are electrons or positrons' involved. In fact significant parts of *How the Laws of Physics Lie* are arguments for realism, on the one hand about causes (1983: 74-86, also 1989: 92, 201) and on the other about theoretical entities (1983: 91-93) subject to such causes and entities playing certain roles in explanation, and also to our not imagining that we are thereby committed to realism about the precise descriptions of those causes and entities to be found in theoretical, or fundamental, laws. We can know that

---

11 Gibbins (1984: 392) suggests that Cartwright's early 'anti-realism' may be explicable in terms of different ways of using the term 'realism' in philosophy on opposite sides of the Atlantic, so that for North American philosophers a realist is one who holds of some class of sentences that they are all true, whereas for Anglophone philosophers on the other side of the ocean a realist would be one who asserts that each sentence in the class is either true or false.

12 See also Gibbins 1984: 390. Hacking (1983) makes more direct reference to Cartwright (1983) than the converse, and Hacking's chapter 12 is largely devoted to Cartwright's ideas.

13 The context of this remark is a discussion with an experimenter about the study of quarks. When asked how some part of the experiment works, the answer given was in terms of spraying part of the apparatus with electrons or positrons. Hacking's point is simply that philosophical worries about realism regarding a given entity have no force against the kind of working causal involvement that makes it seem reasonable to speak of 'spraying' such entities.

there is a kind of thing which can cause in certain ways without any fundamental law being true of either the thing or the cause.

About this much Cartwright and I are in agreement. Where the conflict arises is over the question of laws, and especially fundamental laws. Cartwright combines realism about entities and causes (or in the language of her more recent[14] work of 'capacities') with the rejection of such accounts of those entities and causes as are to be found in fundamental laws. The principal justification for this is that fundamental laws, and fundamental accounts of entities have very poor empirical credentials, and Cartwright contends that in these matters empirical credentials are the only ones which matter, hence that confidence in such laws is the mark of inappropriate metaphysical factors playing a role in the philosophy of science.

Cartwright is adamant that these commitments are entirely distinct from and independent of any kind of fundamentalism about laws, and may be held in conjunction with the rejection of such fundamentalism. Whether this position is sustainable is open to dispute, and in due course I argue that the best account of capacities is one framed in terms of laws.

3.3. Blowing in the Wind

It is important to bear in mind just how far Cartwright is prepared to push her opposition to fundamental laws. The best way to make the point needed here is to take a short look at an example, discussed by Cartwright (1994). Having earlier used the case of being able to drop a pound coin from an upstairs window with some accuracy as an example of a kind of reliable knowledge, Cartwright considers a case, which she acknowledges as being due to Neurath,[15] where what is dropped is a $1000 bill, and where, for good measure, it is dropped on a windy day in St. Stephen's Square. She states that mechanics 'provides no model for this situation' (1994: 283) but also that it is characteristic of the fundamentalist that she or he think that there is 'in principle […] a model in mechanics for the action of the wind [on the banknote], albeit probably a very complicated one that we may never succeed in constructing', and further that this belief is 'essential for the fundamentalist' (1994: 284).

Against this faith Cartwright asserts that if 'there is no model for the 1000 dollar bill in mechanics, then what happens to the note is *not determined by its laws*' (1994: 284 emphasis added). The fundamentalist might rejoin at this point that the credentials of mechanics are well established in a range of other situations, and that this justifies any claim to their applying in the case of the banknote, but Cartwright urges as follows:

---

14 See Cartwright (1989: 1) for an explicit statement of realism about capacities. Cartwright explains at the same point that the general argument of *Nature's Capacities and Their Measurement* is not formulated in terms of capacity realism since she wished to avoid debates over realism not central to her purposes.

15 Cartwright refers to Neurath's 'United Science and Psychology' in McGuiness (1985).

> The successes of mechanics in situations that it can model accurately do not support it, no matter how precise or surprising they are. They show only that the theory is true in its domain, not that its domain is universal. The alternative to fundamentalism that I want to propose supposes just that: mechanics is true, literally true we may grant, for all those motions whose causes can be adequately represented by the familiar models that get assigned in force functions in mechanics. For those motions mechanics is a powerful and precise tool for prediction. But for other motions, it is a tool of limited serviceability (1994: 284).

Here too, a fundamentalist rejoinder is likely. Surely, one feels tempted to insist, the note undergoes accelerations, a succession of rapidly varying ones to be true, but a real observable sequence of changes in velocity. And if this is so, then it is subject to forces, which are what cause accelerations. Cartwright has an answer to that line of thinking too:

> Many will continue to feel that the wind and other exogenous factors must produce a force. The wind after all is composed of millions of little particles which must exert all the usual forces on the bill, both at a distance and via collisions. *That view begs the question*. When we have a good-fitting molecular model for the wind, and we have in our theory (either by composition from old principles or by the admission of new principles) systematic rules that assign force functions to the models, and the force functions assigned predict exactly the right motions, then we will have good scientific reason to maintain that the wind operates via a force (1994: 285 emphasis added).

Before long I turn to the question of what criticisms might be raised against this line of thinking, but for the time being need to emphasise that from Cartwright's point of view any confidence in the general validity of some supposed law of nature outside the domains in which it can be justified by prediction is question begging. In case there is any doubt about how strictly she draws the line between cases where we do and do not have credentials for some kind of realism, her recent spin on Hacking's experimental realism is worth bearing in mind:

> Ian Hacking is famous for his remark, 'If you can spray them, they exist.' I have always agreed with that. But I would now be more cautious: 'When you can spray them, they exist.' (1994: 291-2)

If this remark is to count as drawing a distinction between her position and that of Hacking, the most likely interpretation of it is as saying that we are only justified in making claims about the reality of entities in cases where we are in suitable causal contact. Just as thinking that a law applies outside the conditions of its empirical adequacy is question begging, so too is confidence in the existence of things outside cases where suitable causal contact takes place.

If Cartwright's line of argument goes through then the fundamentalist and emergentist alike are left in a very poor state indeed. But how should we go about evaluating this argument? Cartwright's case is elegant, careful and detailed, far more so than the sketch just offered shows, although further detail will be added in the

course of developing criticisms of it. It will certainly not do simply and dogmatically to hold out for better fundamental laws, or baldly to assert faith in their possibility. And Cartwright is surely correct in some sense to say that we 'best see what nature is like when we look at our knowledge of it' (1983: 13). But then the question becomes different: what is the best way to decide what the state of our knowledge tells us about what nature is like? In response to this question I think it is possible to outline a defence of what I will call moderate fundamentalism, which disagrees with Cartwright on a number of key points, and permits the Completeness Thesis to be sustained.

Before doing that I want to spend a little time making more clear what the implications of Cartwright's position for the Completeness Thesis are.

4. Consequences for the Completeness Thesis

In the sense already noted in connection with Dupré, one way of reading Cartwright's position is as a general rejection of the notion of any causally complete model, theory or form of description.[16] From this both conclude that there is no ontological causal closure which we are justified in holding to obtain. Matters are not quite so straightforward, though, since unlike Dupré, Cartwright does not directly deny either any general thesis of causal completeness, or that specifically concerning the physical. It has been reported to me that in debate and conversation she is willing to defend the view that some material changes happen for, e.g., economic rather than any other causes,[17] which would by implication deny the Completeness Thesis, but I have been unable to find any direct endorsement of such a position in her published work.

Despite this, there is a clear sense in which her work has significant implications for the Completeness Thesis, and no attempt to treat the credentials of that thesis within the broad scope of the philosophy of science can afford to ignore her challenge. I suggest that there are two central aspects to this challenge for my purposes:

In the first place a very natural way of thinking of the Completeness Thesis is as a claim about some set of laws, the physical laws, to the effect that they can account for the likelihood of all physical occurrences. If Cartwright is correct about the status of physical laws, then it would seem that any version of the Thesis framed in terms of laws would have to be false.

A further aspect of the point at stake here, is that in Cartwright (1989) and related material (1994, 1995a, 1995b) Cartwright makes no concessions to the view that some capacities might be more fundamental

---

16 Some opponents of physicalism have drawn on Cartwright's criticisms of fundamental laws, for example Wagner (1993: 133-135).

17 David Papineau (pers. comm. 1999). The case under discussion was the movement of some gold bars from one bank vault to another, and I understand Cartwright's position on the matter to have been that economic forces (in some less-than-usually metaphorical sense of that term) are properly regarded as the real and irreducible cause of the physical motion of the gold.

than others, or that any could be genuinely fundamental at all. That electrons are the cause of the tracks in the bubble chamber, that aspirins are the cause of the relief of headaches, that economic factors are a cause of some social unrest, are all simply well-established claims about the existence and operation of some causal capacity. I have already noted her image of the patchwork of laws. What is important to realise is that not all of the patches are, on Cartwright's view, to be regarded as physical.

In the second place the ways in which the Completeness Thesis is used in various debates[18] relies on the notion of the physical as being fundamental, that is as that which, no matter what else any material thing has, it has to have in virtue of being a material thing. Physicalist theses about the relationship between non-physical and physical properties typically rely on the idea that the physical properties are ontologically more basic. The notion that anything could be in this position is seriously threatened by Cartwright's anti-fundamentalism, since she attacks the grounds which are typically used to justify the claim that physical things have this status. I should note here that in some ways her more recent work (1989, 1995a) grants far greater significance and generality to capacities, suggesting a line of defence for completeness. Further below I pursue just this line of thinking, for the time being I note only that even in recent defences of capacities Cartwright remains uncompromisingly opposed to fundamentalism (e.g. 1995a: 155).

The two aspects of Cartwright's thought just noted come together, of course, in her rejection of fundamental laws. Before turning to the criticism of her views, a quotation from a recent paper will help fix the nature of her challenge in mind:

> '...some features of systems typically studied by physics may get into situations where their behaviour is not governed by the laws of physics at all. But that does not mean that they have no guide for their behaviour, or only low-level phenomenological laws. They could fall under a quite different set of highly abstract principles' (1994: 189).

5. Contra Cartwright

Given the startling and contentious nature of some of her conclusions, it is somewhat tempting to view Cartwright's argument as a *reductio* of its own premises. Surprising and perhaps unpalatable theses are not, though, the same as absurd ones, and a proper counter-argument is needed here.

In one sense some of what needs to be said here has already been said in connection with the discussion of Dupré on reductionism in the preceding chapter. The same arguments as were used there can show here that the features of the world which Cartwright claims support the hypothesis of a 'patchwork' ontology, and Dupré argues support 'promiscuous realism' can be accounted for by a much more limited ontology with a definite fundamental level of description of the sort some hope physics will be able to provide. If the arguments used

---

[18] See the Introduction to the present work, especially section (1.1).

against Dupré hold, then pragmatic disunity is entirely compatible with fundamental closure and unity. So, once again, the argument from methodology or epistemology cannot be ontologically decisive. But the argument against Dupré traded on a certain confidence in the credentials of fundamental laws, which is precisely what is at stake here, since something which is distinctive about Cartwright's case is that she explicitly contests certain orthodox views about the status of fundamental laws. If her major line of argument goes through, then the response to Dupré developed in the preceding chapter finds itself on rather shaky ground.

So the pro-completeness argument that a complete physics *could* support the variety and heterogeneity in question stands in need of some kind of argument for the view that a physical ontology might be preferable to a patchwork one. After all, if both positions can indeed explain the evidence then it follows that the evidence, by itself, is not decisive. This bring us to the question of fundamental laws, since although appeal to fundamental laws is not the only possible road open to the defender of completeness here, it is surely a promising, and perhaps tempting, one. I begin, then, with an attempted defence of fundamental laws.

5.1. Future Laws

One obvious response to Cartwright is to accept her account of the present state of affairs, i.e. concede the accuracy of Cartwright's description, but also to protest that the present state of affairs is temporary, that nothing has been said to disqualify fundamental laws in principle, and therefore that we have reason to hold out for the possibility of a more unified physical science in the future.

In section (3.1) above it was noted that drawing the conclusion that any fundamental law would be false from Cartwright's account of the failings of those laws which she analyses would be a *non sequitur*. That is part of the reason Cartwright offers the simulacrum account of explanation. That too, though, was seen to be a *non sequitur*, since there is no reason why what Cartwright calls a 'simulacrum' could not also be true.

Some debate about present and future laws is pursued by means of meta-inductions on the known history of science. Hacking, for example, observes that every year since 1840 'physics alone has used successfully more (incompatible) models of phenomena in its everyday business, than it used in the preceding year. The ideal end of science is not unity but absolute plethora' (1983: 218). But this argument is almost obviously flawed. Poland (1994: 169) calls it an 'instance of a crude, theoretically unsupported enumerative induction' and claims that the value of Hacking's inference is at least unclear without, *inter alia*, an analysis of 'physics' and some reference to the reasons for which various models were introduced, since without them it is likely that the induction is 'performed over a radically heterogeneous sample'. More importantly Poland points out that neither Cartwright nor Hacking has given compelling reasons, or indeed much reason at all, 'for physicists to abandon the ideal of unity within physics' where such unity is regarded as connected to the truth, rather than the demands of the simulacrum account (Poland 1994: 170).

Furthermore it is far from clear that 'absolute plethora' could possibly function as an 'ideal end' for science. Plethora is, after all, far more easily come by than science.[19] But even if we reject the under-supported claim that plethora is the end of science, we have to deal with the more reasonable possibility that it may be an inevitable consequence of the pursuit by scientists of other goals. This is to say that whether or not scientists are actually seeking disunity, as scientists they will tend to produce it.

The problem for Poland's response, and others like it, is that they are based on a misinterpretation of Cartwright's position. She does have arguments for the necessary empirical poverty of fundamental laws, developed in her vision of the simulacrum account of explanation. It is not possible to defeat her position by protesting that it is inductive, since (perhaps unlike Hacking's suggestion noted above) it is not. And even if it was, that fact would not make it false.

A variation on the inductive argument holds that there is a straightforward sense in which theoretical advances at the level of fundamental laws, such as the unification of two formerly distinct bodies of laws, are (subject to certain qualifications) regarded in an uncomplicated way as scientific achievements. Furthermore there is no acceptable inference from the fact that some physicists, no matter how many, work on more phenomenological problems to the illegitimacy of the fundamental work being pursued by others. (See Poland 1994: 170 and Spurrett 1999a: 49).

At this point a committed Cartwrightian might object that all this simply misses the point, since what is at stake is the notion that this work on unification is typically mistaken for something which has a certain relation to the truth, while there is an argument on the table, the simulacrum account of explanation, that it does not and cannot. What we seem to have here is a conflict which is sustained by loyalty to different broad principles not likely to be the objects of compelling arguments either way, and hence perhaps a dead end.

The proper conclusion to be drawn here, I submit, is that the Hacking-style meta-induction, and the appeal to future fundamental laws are both under-motivated.[20] What is needed is a treatment of the status of *current* fundamental laws – if Cartwright's criticisms turn out to hold the field, then the day goes to the defenders of disorder. If, on the other hand, a principled defence of fundamental laws is possible, then there is no need to stake the physicalist programme on some ideal set of future super-laws. In the context of the present enquiry this is perhaps even more urgently so, since the arguments against both emergentism and promiscuous realism relied in significant ways on present fundamental physical laws. The remainder of the present section of

---

19 I have argued at length against a strikingly similar view, Lyotard's assertion that postmodern science is legitimated 'solely on paralogy' (*The Postmodern Condition*) in my 'Lyotard and the Post-modern Misunderstanding of Physics' (1999a).

20 Davies (1996) makes a similar point in connection with Dupré: that direct inferences from epistemology to ontology are under-motivated and typically unsupportable.

this chapter, then, is devoted to the attempt to defend the credentials of some of the fundamental laws which we already have, from the criticisms which Cartwright levels at them.

5.2. Stating the Facts

The brief overview of Cartwright's position presented above (section 3) should make clear that a crucial element in Cartwright's claim that theoretical laws should be regarded as lies depends on the way she understands the question whether they 'state the facts'. In due course I argue that it is on precisely this point that her position can be defeated, but for now need to make clear just how her argument relies on a particular notion of 'facticity'.

The 'facticity' view of laws is the view that what a scientific law does is describe the actual behaviour of some or other kind of physical system or process. Cartwright describes this as though the issue at stake is one about facts in general, glossing the facticity view as one to the effect that 'laws of nature describe facts about reality' (1980: 75). This is not quite correct, though. Cartwright has in mind a specific kind of fact, which is to say that facticity is not about any facts at all, but specifically about the kind of facts which apply to whatever actually happens, or, in her words 'how bodies actually behave' (1980: 75).[21]

It is crucial to note that the gap between facts and facts-about-how-bodies-actually-behave is significant. Partly to pre-empt the argument of section (5.4) below, a typical fundamental law, for example Newton's Law of Universal Gravitation says of two bodies that there will be a force acting on each of them which is a factor of their masses and the distance between their centres of mass. Similarly Coulomb's law of electrostatic force says that a force will be exerted between two bodies which is a factor of charge and distance.[22] It should hardly need pointing out that neither of those laws says anything at all about what will actually happen to the bodies in question. Granted, Newton's second law holds that a given body will accelerate in the direction of the net force upon it, so there are some laws which do make claims about what will happen. The point I need to emphasise here is simply that there is no reason at all to set out by holding all laws to a standard assuming that their purpose is to describe the actual flow of events.

Indeed when Cartwright asks of both Newton's and Coulomb's laws whether they 'truly describe how bodies behave?' (1980: 77) the question to be asked in response is why at all a law about a force should be expected to tell us, by itself, what happens. The facticity or falsehood option is a false dichotomy, and what we should be asking rather is what kinds of fact, if any, fundamental laws state. Cartwright may well be right that

---

21 Creary (1981: 149) notes that on Cartwright's view to 'state a fact' about an object is not simply to 'attribute to it a property which it in fact has' but that the property in question has to be a '*behavioral*' one.

22 Both of these laws, and others, are discussed shortly in section (5.4).

they are false, but if they are it will be because they fail to say anything true at all, not because the only alternative to accurate descriptions of behaviour is outright falsehood.

5.3. Bhaskar and Cartwright[23]

It will not have escaped attention that there are striking similarities between aspects of Cartwright's treatment of the status of scientific laws and that of Bhaskar, discussed in Chapter Two above in connection with emergentism.[24] Both raise the possibility that scientific laws may well be false for related reasons. What Bhaskar calls 'actualism' is very close to what Cartwright calls 'facticity'.

In *A Realist Theory of Science* (1977) Bhaskar identified and engaged with many of the same basic problems as Cartwright does in her work, and came to conclusions both similar in some respects and strikingly different in others. I want briefly to describe the core of his argument, and oppose some of its conclusions to those of promiscuous realism and nomological pluralism, in order to argue myself for the possibility of generally valid scientific knowledge in a way which is not, as Cartwright suggests, question begging. Bhaskar's focus in his defence of what he calls transcendental realism is, to a significant degree, the scientific experiment, precisely the context in which Cartwright allows that theoretical or fundamental laws may, sometimes, state the facts. Bhaskar distinguishes between open and closed systems, where closed systems alone manifest consistent and law-like relations between events,[25] and observes that 'an experiment is necessary precisely to the extent that the pattern of events forthcoming under experimental conditions would not be forthcoming without it' (1977: 33). This seems difficult to argue with, and it also makes sense to think of many experiments as procedures for bringing about closure in Bhaskar's sense. He then argues that if 'experimental activity is to be rendered intelligible' then we must allow that 'natural mechanisms endure and act outside the conditions that enable us to identify them [in order] that the applicability of known laws in open systems, i.e. in systems where no constant conjunctions of events prevail, can be sustained' (1977: 13).

What Bhaskar means by a natural mechanism, or as he more often puts it a *generative* mechanism, is some real and causally active feature of the world. Such generative mechanisms need not generally produce consistent outcomes, though, since they may be present in various combinations and conditions. So for

---

23 Bhaskar's *A Realist Theory of Science* was first published in 1975, Cartwright's 'Do The Laws of Physics State the Facts?' in 1980, and *How the Laws of Physics Lie* in 1983. Except for one reference to Cartwright in the postscript to Bhaskar (1989) the two have said nothing which I can find about one another's work. I conjecture that a common source for both of them may be Harré's work on causal powers, e.g. Harré and Madden (1973), or possibly Ryan's work on Mill (1970).

24 One of the few to have explicitly made the connection between Bhaskar and Cartwright is Chalmers (e.g. 1987), although see also Humphries (1995).

25 Bhaskar's specific critical target is positivism, so his typical closed system, by and large, is one in which a relationship of constant conjunction obtains between events. It does no significant violence to the structure of his arguments to deploy them more generally in the way I do here.

Bhaskar it is not generally the case that the typical activity of a generative mechanism stands in any stable relationship with the manifest flow of events, a point he marks by saying that the events are often 'out of phase' (1977: 13) with the mechanisms. When they are in phase (for whatever reason) we can say that the relevant system is closed. Bhaskar opts to call our knowledge of these generative mechanisms 'laws', but argues that laws as he sees them are generally explanatory tools, and only contingently descriptive or predictive. But then an important consequence for an influential traditional conception of laws, which sees laws as relating events, comes into the foreground. In open systems stable relations between events do *not* generally obtain, but 'clearly if the law-like statements whose antecedents are instantiated in open systems are interpreted as invariant empirical regularities they must be regarded as false' (1977: 132).

This is pretty much the point at which Cartwright would say that such laws are, in general, false, but also that no important consequences for their general applicability follow from the fact of their being true in the unusual conditions of an experiment. Bhaskar prefers to see the situation as a dilemma for anyone wedded to the notion that laws relate events:

> ...for to the extent that the antecedents of law-like statements are instantiated in open systems, he must sacrifice either the universal character or the empirical status of laws. If, on the other hand, he attempts to avoid this dilemma by restricting the application of laws to closed systems (e.g. by making the satisfaction of a ceteris paribus clause a condition of their applicability), he is faced with the embarrassing question of what governs phenomena in open systems (1977: 65).

Bhaskar, though, opts to sustain what he calls the 'universal character' of laws, i.e. their general applicability, by conceding that they are not primarily empirical. This position is in some respects similar to Cartwright's if we set aside the specific issue of which items of our knowledge we should call 'laws', since Bhaskar's contentions are that we should be realists about the tendencies, or ways of acting, of things (1977: 46) just as we should be realists about the things to which we attribute causal powers, while Cartwright defends realism about capacities and entities. Although there are differences of vocabulary and detail here, the agreement is striking. What separates the two philosophers on this point is rather the issue of the generality of our knowledge. Bhaskar's argument for generality is Kantian in form: 'experimental activity can only be given a satisfactory rationale if the causal law it enables us to identify is held to prevail outside the contexts under which the sequence of events is generated' (1977: 33).

Cartwright herself puts a Kantian gloss on her argument for the patchwork of laws, where the activities of planning, predicting, manipulating and so on would be impossible without the objectivity of local knowledge: 'Unless our plans about the expected consequences of our actions are reliable, our plans are for nought. Hence knowledge is possible' (1994: 277). But, as we have seen, she is also careful to point out that for her part the appropriate ontological caution requires that we restrict our confidence in the applicability of our

knowledge to cases where that knowledge can deliver accurate descriptions, i.e. that we say only '*when* we can spray them, they exist'.

Neither Cartwright's position nor Bhaskar's refutes the other. Both lay claim to the same evidence, but disagree over its interpretation, or over the appropriate metaphysical commitments which do best justice to the evidence. We do not generally, at least any more, expect metaphysics to tell us what the world must be like in the same ways that Descartes, for example, argued that the most basic laws of nature, including the principle of inertia, could be deduced from reflection on the nature of God, who *qua* perfect being would always act consistently (1991: 57-61). Rather the job we typically demand from metaphysics these days is that it make some kind of sense of our opinions about, and knowledge of, the world, and that it do no major violence to common sense.

We have a choice here between saying, with Cartwright, that when in experimental situation(s) *X* we have reason to suppose a capacity or entity *Y* to be real, we should conclude that *Y* is indeed real, but only in *X* situations, or saying, with Bhaskar, that in such a case we could properly conclude that *Y* is generally real, but only manifest in *X* situations. A Bhaskarian, or what I am here calling a moderate fundamentalist, would say that the fact that we can make a variety of different capacities manifest in this way makes sense of the fact that systems in general are not closed: if mechanisms did not interfere, compete and reinforce in a variety of ways experiments would not be necessary, and we could read the laws of nature directly off the normal course of events.

So to say that what we do in experiments is enable pre-existing and enduring tendencies to become manifest because of the simplicity or restrictedness of the experimental situation, then there is nothing unreasonable or dogmatic about taking those tendencies to be generally active, even when the flow of events is out of phase with them. To conclude this is, in general, no less reasonable than to suppose that the light in my fridge retains the capacity to come on when the door is opened, during intervals when the door is in fact closed.

Not only that, we can see why Bhaskar describes the problem of explaining what happens in open systems as 'embarrassing' for one who thinks that laws hold only where descriptively accurate, and how it might also be awkward for one holding the simulacrum account of explanation. A moderate fundamentalist confronted with an uncrumpled $1000 bill blowing in the wind can say that we know from experiments in a vacuum that falling bodies accelerate at the same rates, from other experiments with fluids and gasses that the air can exert a force, from other still that air is composed of molecules which can be turbulent, and so on. It is fair to say that given such knowledge we can *predict* that the path of the note will be practically impossible to plot in advance, and explain why this is so. And our explanation will not be a high-level simulacrum which organises otherwise disunified local knowledge, but rather a true statement of the actions of real tendencies in

the world which produce the forces acting on the note. There is nothing incautious about this, and no questions are being begged.[26]

I suggest that the burden of argument has already been shifted onto the Cartwrightian here: an approach to laws along the lines defended by Bhaskar makes sense of the failings of actualism/facticity, but also does justice to the ways we tend to think about at least some scientific achievements. Bhaskar show how we can take scientific laws to be true, and make sense of the fact that this only sometimes enables us to make predictions. To press the point home, though, it will be worth looking in more detail at one of Cartwright's own arguments and seeing how much of it is defensible along the lines of the moderate fundamentalism advanced here.

5.4. Force Laws and Composition: Mill and Bhaskar versus Cartwright

A crucial argument in Cartwright (1980) proceeds from an analysis of the composition of forces to the thesis that, except in specific and unusual circumstances, fundamental laws have to be regarded as false. Her argument is criticised by Creary (1981) and Cartwright's presentation of the same basic material in her (1983) is updated to include a response to Creary. A major foil for Cartwright's argument in both its incarnations is Mill's (1843)[27] discussion of the composition of causes, whose views are also criticised by Bhaskar (1975). Furthermore Nagel (1963) takes Russell (1903) to task for making a claim in some significant respects similar to that advanced by Cartwright while Gibbins (1984) discusses Cartwright (1983) in the light of Creary (1981).

Cartwright's argument is an especially important one for my purposes. Recall her admission that a fundamental law can be regarded true when it does 'state the facts'. It seems more than plausible (although this is exactly what Cartwright denies) to regard at least some situations, traditionally dealt with by means of composition of forces, as being characterised by more than one law being true of them. So, if the conclusion to her treatment of composition can be blocked, it points the way to a relatively general rejection of aspects of her anti-fundamentalist stance, in defence of moderate fundamentalism. What is at stake in the debate over composition is the question of what we are and are not entitled to say we know is the case where more than one causal law, tendency or capacity is at work. An argument which takes her admission that fundamental laws can sometimes be true, and develops a suitable truth-preserving account of composition, can significantly extend the scope of fundamental physical laws in the face of Cartwright's patchwork realism. I argue that just such an argument can be constructed, blocking Cartwright's anti-fundamentalist conclusion, and drawing for the most

---

[26] It might be objected at this point that much of this sounds reminiscent of things Cartwright herself says. Quite so – after all she allows that forces can be regarded as a kind of 'trying' even where their effects are not manifest (1994: 285-286). But what is at issue here are her claims to the effect that '*when* we can spray them, they exist' is more cautious a conclusion that the 'if' formulation, and that the view that the wind exerts a force on the banknote at any instant 'begs the question.'

[27] Cartwright refers to Mill (1856), i.e. to the fourth edition of Mill's *System of Logic*. My references are to the eighth (1872) edition, although there are no significant differences between the various editions for the purposes of the discussions which occupy the present sub-section.

part of Bhaskar and Mill,[28] although also making use of Nagel's criticism of Russell, and Creary's and Gibbins' criticisms of Cartwright. The laws which I will discuss are examples used by Cartwright herself: especially Newton's and Coulomb's laws. It is not important for my argument that Newton's, for example, is regarded as an approximation from the point of view of special and general relativity. The reason for this is that such arguments concern the potentially superior credentials of even more general fundamental laws, rather than the propriety of reasoning with fundamental laws at all. If Newton's laws are to be criticised on relativistic grounds, they need to be defended from Cartwright first. I begin with Cartwright's (1980) account of the issues.

Cartwright's argument, briefly, is (a) that laws must either satisfy the facticity requirement (i.e. 'state the facts') or be false, (b) that where more than one law combines neither of them satisfies this requirement, and hence that (c) in such cases both laws are false. The key premise[29] is (a), the 'facticity or falsehood' dilemma. As noted already, Cartwright's foil is, to some extent, Mill's account of the composition of causes. This was discussed above in Chapter Two (section 2.1) but the key point to be noted here is Mill's notion that in cases where composition takes a mechanical[30] form, both (or all) of the laws in play have their *full* effect:

> In this important class of cases of causation, one cause never, properly speaking, defeats or frustrates another, both have their full effect. If a physical body is propelled in two directions by two forces, one tending to drive it to the north and the other to the east, it is caused to move in a given time exactly as far in both directions as the two forces would have separately carried it; and is left precisely where it would have arrived if it had been acted upon first by one of the two forces, and afterwards by the other (1843: 428).[31]

Mill also makes the same point by saying that 'even when the concurrent causes annihilate each other's effects, each exerts its full efficacy according to its own law, its unique law as a separate agent', and that a key

---

28 I came upon the idea of opposing Mill to Cartwright on this point in a number of footnotes in McLaughlin (1992: 59-60, 70, 71) where McLaughlin refers to a paper entitled 'Mill's Response to Cartwright' as being in preparation. In correspondence Professor McLaughlin explained to me that the period of preparation has expanded in the usual manner, and that his paper remains uncompleted. McLaughlin (1992) refers sympathetically to Creary (1981).

29 Creary (1981) gives a similar reconstruction, but focuses his criticism on premise (b). Since the facts which he takes the laws as stating are not of the kind falling under the facticity requirement, though, his line of criticism is closer to mine than this apparent difference in target suggests.

30 Cartwright's discussion does not note that Mill held that there were two modes of composition, although Gibbins (1984: 394) suggests that one way of reading her position is as holding that all cases of composition are what Mill called chemical, i.e. that no physical science is deductive. Gibbins seems to have misconstrued Mill, though, since although it is true that Mill rejected the thesis that chemistry (or anything where composition occurs in the chemical mode) could be deductive on the basis of data concerning the constituents of chemical processes, he also held that once the chemical data or other was in, any science could become properly deductive. See Chapter Two above, especially section (2.1), and the comments on Oppenheim and Putnam's (1956) remarks on Mill in Chapter Three section (2.2).

31 Mill's treatment here combines aspects of the composition of displacement with that of force. The two issues are separate but related, and typical physics textbooks regard the latter as an extension of the former. (See, e.g. Sears, Zemansky and Young 1987: 10-14, 51-56, 71-75).

characteristic of conformity to the principle of the composition of causes is that some laws (one of Mill's examples is the conservation of 'weight') are 'rigidly fulfilled in every combination into which the objects enter' (1973: 373). Mill regards the final effect of these cases as being the sum of the effects of the laws taken separately, and takes the parallelogram law in mechanics as the model for addition in the mechanical case. (The question of how it is possible to regard the diagonal of some figure as the sum of its sides is one to which I return shortly.) There is a risk of ambiguity in Mill's description here, since at times he speaks of causes 'annihilating' one another's effects while still having full 'efficacy', and at others he maintains that both laws have their full effects.[32] Neither of these possibilities, though, is the same as what Cartwright seems to be taking Mill as saying, which is that the actual outcomes which would have obtained were each law to operate in isolation must actually take place for the law to be considered as having its full effect. We will see shortly how this interpretation affects both Cartwright's reading of Mill, and of the wider issues at stake here.

A paradigmatic example of composition along the lines presently under discussion is that given by Newton. The famous parallelogram rule of composition is presented as a corollary to his three laws of motion, to the effect that:

> A body, acted on by two forces simultaneously, will describe the diagonal of a parallelogram in the same time as it would describe the sides by those forces separately (1934: 14).

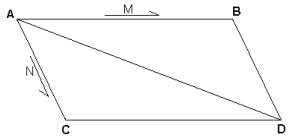
This is supposed to follow for the most part from Newton's second law, the law that changes of motion (accelerations) occur in the direction of the impressed force, and the first law, that motions persist with unchanging direction and speed unless forces are applied. Newton's geometrical reasoning is straightforward enough. Given a body starting at point A, and two forces where the one, M, would drive the body along a line AB to point B in some time, while the second, N, would drive it along the line AC to C in the same time:[33]

> … let the parallelogram ABCD be completed, and, by both forces acting together, [the body] will in the same time be carried in the diagonal from A to D. For since the force N acts in the direction of the line AC, parallel to BD, this force (by the second Law) will not at all alter the velocity generated by the other force M, by which the body is carried towards the line BD. The body, therefore, will arrive at the line BD at the same time, whether the force N be impressed or not; and therefore at the end of that time it will be found somewhere in the line BD. By the same argument, at the end of the same time it will be found somewhere

---

32 Geach (in Anscombe and Geach 1973: 102-3) has noted that Mill's thinking on this question combines aspects of a realist reading of causes as tendencies (a word which Mill uses) and a Humean notion of causal laws as regular relations between outcomes.

33 Pictorially:



> in the line CD. Therefore it will be found in the point D, where both lines meet. But it will move in a right line from A to D, by Law I (1934: 14).[34]

Newton's position, then, is that each force individually determines a range of possible outcomes, and that the combined affect of several forces will be that outcome where all the ranges intersect. It should be clear that the case of two forces in equilibrium is a limiting case where the body acquires no change in motion, since there is no net force, and that the case where two or more forces act in the same direction is another limiting case, where the sole effect of combination is a change in the magnitude of the force.[35]

Cartwright herself takes a Newtonian law, the universal law of gravitation as her starting point, noting that Feynman endorses the evaluation of it as 'the greatest generalization achieved by the human mind' (Feynman 1965: 14, in Cartwright 1980: 76).[36] It is precisely this notion of the generality of Newton's or any other law which Cartwright rejects. As we would express the law now, it reads:

$$F = Gmm'/r^2$$

Alternatively, in Feynman's words, also quoted by Cartwright:

> The Law of Gravitation is that two bodies exert a force between each other which varies inversely as the square of the distance between them, and directly as the product of their masses (Feynman 1965: 14, Cartwright 1980: 77).

Cartwright then asks whether this law truly describes 'how bodies behave' and answers that it does not, noting that Feynman himself 'gives us one reason why' in his description of Coulomb's law, stating the force due to electrical charge (another inverse square law).[37] This law relates the magnitude of the force in question to the charges of the two bodies, and reads:

$$F = qq'/r^2$$

---

34 The way Newton describes matters it is clear that he is thinking of forces impressed at an instant, resulting in the body travelling with a constant velocity. This is a semantic matter, though, and the account need not rely on the notion that the body acquire a constant velocity, or that forces be exerted instantaneously.

35 The notion of composing motions and tendencies to motion is older than Newton. Perhaps the classical modern treatment is Galileo's account of the path of a projectile, combining a inertial tendency perpendicular to a constant downwards acceleration, and resulting in a parabolic trajectory (see Galileo 1914: Fourth Day). Dugas (1955: 21) attributes to Aristotle a parallelogram rule for composing motions, although the text to which he refers, the *Mechanica*, is not generally regarded as the work of Aristotle. (See Aristotle 1936: 337-9, 848b 10-30).

36 It is worth noting that Cartwright has made a somewhat bizarre choice of text to use as her target here. Feynman (1965) is a text based on a series of popular public lectures, rather than an attempt either to offer a philosophical account of physical laws, or to speak to a scientifically sophisticated audience. Cartwright's references are to the (1967) MIT edition, mine to the (1965) BBC edition.

37 Feynman himself does not present Coulomb's law as an example of something which shows that the universal law of gravitation is ever false, simply as another example of a law which he regards as having great generality.

The 'facticity or falsehood' dichotomy, which I read as premise (a) in Cartwright's argument, begins to show itself when she considers the case of bodies which are both massive and charged:

> It's not true that for *any* two bodies the force between them is given by the law of gravitation. Some bodies are charged bodies, and the force between them is not Gmm'/r^2. Rather it is some resultant of this force with the electric force which Feynman refers to (1980: 77).

So far so good. Cartwright, though, describes this by saying that massive bodies are 'counterexamples' to Coulomb's law, and that charged ones are 'counterexamples' to Newton's, concluding that 'the two laws are not true; worse, they are not even approximately true' (1980: 77). I argue shortly that Cartwright's case has gone entirely off the rails by this early point, but not for the reasons she proposes when she notes that there is a possible rejoinder which could be put here, to the effect that she has not formulated the laws correctly, and that, allowing for the existence of 'implicit *ceteris paribus*' modifiers Newton's law should be taken to read something more like:

> *If* there are no forces other than gravitational forces at work, *then* two bodies exert a force between each other which varies inversely as the square of the distance between them, and varies directly as the product of their masses (1980: 77).

Cartwright is prepared to allow that *this* law is a true law 'or at least one held true within a given theory', but objects that it is not 'very useful' because it can explain only 'very simple ideal circumstances', that is, cases where only gravity is 'at work' (1980: 77). Cartwright notes at once, though, that real explanations, directed at complex circumstances, often break the phenomenon in question down into simple parts, and one version of this type of explanation is what Mill called explanation by composition of causes (Cartwright 1980: 78). A paradigm of this type of explanation is vector addition, and Cartwright asks whether vector addition provides a 'simple and obvious answer' (1980: 78) to her concerns, and replies that it does not. On the vector addition view the forces due to gravity and to charge would both be present, and the net force or resultant determine what actually takes place. Cartwright responds as follows:

> The vector addition story is, I admit, a nice one. But it is just a metaphor. *We* add forces (or the numbers that represent forces) when we do calculations. Nature does not "add" forces. For the "component" forces are not there, in any but a metaphorical sense, to be added; and the laws which say they are must also be given a metaphorical reading (1980: 78).

In explanation for this somewhat startling assertion Cartwright maintains again that in cases where both charge and gravity are in play the forces due to gravity and charge are 'not real, occurent forces', that in interaction 'a single force occurs – the force we call the "resultant" – and this force is neither the force due to gravity nor the electric force' (1980: 79). Noting that Mill would reject this claim, and quoting his description of the body simultaneously propelled with equal force in two directions, north and east (quoted above, Mill

1843: 428) Cartwright objects that while events may have temporal parts, they cannot have parts of the sorts described by Mill, that 'no pure north motion can be part of a motion which always heads northeast' (1980: 79). This is not an argument but an assertion.[38] Pushing the 'facticity or falsehood' dichotomy again, Cartwright urges:

> In the simplest case, the consequences that the laws prescribe must be exactly the same in interaction, as the consequences that would obtain if the law were operating alone. But then, what the law states cannot literally be true, for the consequences that would occur if it acted alone are not the consequences that actually occur when it acts in combination (1980: 83, 1983: 72).

Saying again that it is 'implausible' to regard the separate forces as parts of the net or resultant force, Cartwright does say that we can preserve the truth of the laws 'by making them about something other than the facts', that is 'the causal powers that bodies have' (1980: 79). This view sounds close to Bhaskar's (a point noted in Chalmers 1987: 91) although having raised it Cartwright immediately abandons it, saying that the task of giving an account of causal powers seems very difficult, and that 'no story I know about causal powers makes a very good start' (1980: 79, 1983: 61).[39]

There are a number of things which need to be separated carefully here. It is beyond dispute that Cartwright is correct in maintaining that except in unusual and often controlled conditions the laws of physics do not satisfy the facticity requirement. Whether that makes those laws false (or 'lies'), and especially whether it does so in the case of force laws, is another matter. Whether Cartwright's proposed solution, i.e. making the laws be about causal powers can be made to work without either some kind of realism about the component forces or the retention of other factors in the traditional account another still.

I propose that rather than asking whether the laws of physics state the facts, we should ask what kind of facts it is that they do state. Even though the laws only rarely, or in particular conditions, enable us to accurately describe the actual course of events, I maintain that they can be regarded as stating truths about what is happening. Furthermore I will argue that Cartwright's talk of causal powers, and her more recent talk of capacities, is best cashed out in terms of just the notion of laws, especially fundamental laws, which she wants to deny.

Creary (1981) has offered a criticism of Cartwright (1980) which attempts to salvage some facticity for fundamental laws. He regards the key step in Cartwright's argument as the premise that the 'laws used in

---

38 Cartwright says, without explanation, that 'We learn this from Judy Jarvis Thomson's *Acts and Other Events*' (Cartwright 1980: 79, 1983: 61) but gives no specific reference. I turn to the parts and sums question shortly.

39 Chalmers (1987: 91) suggests that 'Bhaskar has made a good start'. Making the laws be about powers adds to the explanatory task rather than solving it, since it leaves us having to explain how it comes to be that the powers are regularly exercised. Making the laws about regularly exercised powers, though, is to take the Bhaskarian step of making them attributions of tendencies, which is to keep them as laws.

explanation by composition of causes do not describe truly how the relevant things behave or what they do' (1981: 148) a premise which he argues is false. Joining the debate with Cartwright's example of the two massive and charged bodies, he contends that the account of the explanatory force of fundamental laws Cartwright entertains briefly following her rejection of their claim to truth and of the traditional account of composition, is 'woefully inadequate', quoting the following important passage:

> ... the law of gravitation claims that two bodies have the *power* to produce a force of size Gmm'/r^2. But they don't always succeed in the *exercise* of it. [...] ... the laws we use talk not about what bodies do, but about what powers they possess (Cartwright 1980: 79, in Creary 1981: 150).

Creary argues that this approach is a 'non-starter' for the reason that what we need in an explanation of the motions of the particles under consideration is a view about powers which *are* exercised, rather than simply possessed, if we are to be able to have any significant grip on what the particles do in fact do. In Bhaskar's terms we could say that what is needed is a way of thinking of things which can cope with the idea of a power exercised without being actualised. Unexercised powers don't explain much. He proposes that a 'facticity-preserving solution' is possible, though, which relies on a distinction between laws of causal *influence* and laws of causal *action* (Creary 1981: 150-151). The distinction works by making the force laws into laws of causal influence, which are to be read factually as saying that certain forces are exerted in certain conditions, but where what actually happens in a given situation depends on the laws of causal action, determining how the various causal influences will work together. There is, *prima facie*, a lot to recommend this view, since as noted already (5.2 above) force laws do not by themselves make any claims at all about what will happen, stating only that under certain conditions certain forces will be exerted. Creary wants to take these laws at face value, and to keep separate the question what will happen when forces are exerted, especially in combination. So, on Creary's view, Newton's first, second and third laws should be seen as laws of causal action, while his law of gravitation is a law of causal influence.[40]

On the basis of this distinction Creary opposes Cartwright's anti-realism about component forces in favour of the reality of the resultant. Creary inverts the priority, arguing that it is preferable to regard the component forces as real or 'natural', since they are related to our well-justified knowledge of the laws of causal influence, while the resultant is the fiction, the 'mathematical' component force (1981: 151-2). Creary also notes that Cartwright offers no direct argument in favour of her anti-realism about the component forces, but rather offers an attempted refutation of a single alternative view, based on considerations of the part-whole relation.

---

[40] Creary's view is closer to Mill's than he seems aware, a point made in McLaughlin (1992: 60), and is also similar in key respects to Chalmers (1987: 88f).

Creary makes clear that on his own view, where the component forces are real because they are the influences of real causes, 'causal influences are neither causes nor effects, but constitute a distinct third kind of entity in the causal ontology, that mediates between causes and their effects' and argues that this approach to causal influences is essential 'to their role in grounding the facticity of the laws of influence that figure so prominently in [...] explanations by composition of causes' (1981: 152). This is in direct opposition to Cartwright's view, which is that effects cannot typically be decomposed into parts, or seem as the sums of various partial causes. Creary maintains that his view, where the causal influences are an intermediate step, solves the problem of getting from cause to effect. It does this, he claims, by making forces independent of both the conditions under which they arise, and the outcomes to which they give rise, so that we can reason from situations to forces and thence to outcomes, although the final step will only be possible where we have suitable laws of causal action.

One final point in connection with Creary's proposal is that although he takes vector addition as a paradigm case of a law of action, he is also careful to note that not all composition or combination may take that form, and argues that the key requirement from the notion of laws of action is that 'it makes sense in general to speak of *reinforcement*, *interference*, and *predomination* occurring among the causal influences governed' by the laws of action (1981: 153). The key point, though, is that we do not need to know these laws in advance in order to salvage the laws of causal influence.[41]

Unsurprisingly Cartwright is having very little of this. In the section added to the (1983) version of 'Do the Laws of Physics State the Facts?' she raises the following two objections to Creary's argument.

The first is that there are rarely any *general* laws of interaction which might do the work done for Creary by vector addition in his account of the gravity-charge example, and that without such general laws 'the collection of fundamental laws loses the generality of application which Creary's proposal hoped to secure' (1983: 63). Cartwright goes on to discuss the use of models and approximations in the study of flow processes, illustrating how a variety of different models are used in a pragmatic manner, taking this to show further the irrelevance of fundamental laws. This line of objection to Creary does not do any significant damage,[42] for two reasons. Firstly Creary's view in no way depends upon the availability of laws of causal influence, and as noted he himself recognises that such laws may be unknown, difficult to apply and so forth. Secondly the 'generality of application' Creary was after has to do with the laws being true as laws of causal influence, which by itself carries no direct implication that they could play any role in prediction or modelling. It is entirely consistent

---

41 A point also made, although not specifically in terms of the distinction between laws of action and laws of influence, by Chalmers (1987: 88, 90).

42 Gibbins is not impressed either: 'Cartwright objects, rather lamely I think, that one cannot always *find* laws of causal action for combining casual influences.' (1984: 395)

with Creary's view that phenomenological laws could be more accurate and useful in most situations. Cartwright's first objection simply fails to hit the mark.

Cartwright's second objection is directly aimed at Creary's talk of causal influences. Cartwright points out that she and Creary disagree about the reality of the component forces: she wants the resultant only, and he the components. She notes that in the case of composing forces Creary's approach of placing the causal influence 'between the cause and what initially looked to be the effect' is 'plausible' (1983: 66), but argues that this way of doing things will not make for a reasonable general strategy for explanation:

> Take any arbitrary example of the composition of causes: two laws, where each accurately dictates what will happen when it operates in isolation, say 'C causes E' and 'C' causes E'', but where C and C' in combination produce some different effect E''. If we do not want to assume that all three effects — E, E', E'' — occur (as we would if we thought that E and E', were parts of E''), then on Creary's proposal we must postulate some further occurrences F and F', as the *proper* effects of our two laws, effects that get combined by a law to yield E'' at the end (1980: 66).

This does not seem to be getting anywhere, since Cartwright is simply repeating her position rather than offering any new arguments in defence of it. What is at stake here is Creary's proposal concerning a way of thinking about composition of forces where the causal influences are not identified with their effects, but with their production of certain tendencies. Cartwright's response, just quoted, is framed entirely in terms of a way of seeing things where laws are identified with their effects, which is how it is that she ends up with the worry about all three effects being real. It does not help that Cartwright aims to reject the facticity view, the point is that Creary's proposal is *neither* the facticity view nor Cartwright's own, which means that simply repeating anti-facticity arguments is not a proper response to Creary at all. In order to make any progress we need to get away from the current deadlock. In order to do this it is not necessary directly to refute Cartwright's argument; developing a satisfactory account of composition which saves fundamental laws and is immune to her lines of criticism will be more than sufficient. Bhaskar and Mill show the way to just such an account. After looking at their views I say something more about the question of parts and wholes, Cartwright's remarks on which are generating more heat than illumination here, and finally at some arguments Mill offers in favour of laws of maximal generality.

How are we best to think of cases where more than one causal influence is at work, or where it seems as though there is good reason to think that there are two at work? I propose to take it as a premise that given some grounds to think that a causal law is true based on the examination of cases where that law acts in isolation we have a good *prima facie* reason to think that the same law obtains in relevantly similar cases where either there are other known causally relevant factors or the actual outcomes do not conform to those predicted by the original law taken alone. This is to say, for example, that an experiment on the behaviour of massive bodies is something I will interpret as an experiment which tells, to the extent that it is successful, us about

mass in general, rather than the masses of just those bodies which are involved, or about what happens to masses when just that happens to them. The test here is to see if this premise stands up in the face of the composition problem. Taking, to stay with Cartwright's example, Newton's law of gravitation as a truth about bodies with masses in general and Coulomb's law as a truth about bodies with charges in general, how do we deal with the case where we are considering a pair of bodies which are both massive and charged?

Bhaskar thinks that there are two mistakes which need to be avoided here. As noted his view is that laws of nature should be regarded as typically being not nomic, that is not statements about what happens, but rather 'normic', that is statements of tendencies or the ways of acting of things. So he rejects Mill's contention that where two or more laws interact in a way which leads to a result unlike that predicted by either, that 'all are fulfilled' (Mill: 1973: 443) saying that 'Mill's mistake here is to suppose that whenever a tendency is set in motion the effect must be in some sense … occurring…' (1975: 99).[43] But Bhaskar also rejects the view, which he attributes to Geach,[44] of 'supposing that whenever no effect (of a given type) occurs, nothing can be in motion or really going on' (1975: 99). Quite correctly Bhaskar points out that the view he attributes to Geach is simply wrong, that 'Balaam's ass is pulled two ways; we do just manage to keep our tempers' (1975: 99-100).[45]

A crucial aspect of Bhaskar's treatment of tendencies is that they may be exercised unfulfilled. When this happens, he maintains that there are two things which need not be doubted:

> […] (a) that something actually happens, towards explaining which the exercise of the tendency goes some way; and (b) that something is really going on, i.e. there is a real generative mechanism at work, which accounts for the influence of the factor the tendency represents in the generation of the event (175: 99).

From Bhaskar's point of view both mistakes are versions of the same fundamental error: that of 'seeing the fulfilment of a tendency [as] a condition of its exercise' (1975: 100) which in Mill's case results in thinking of both tendencies as being actualised, and in Geach's in regarding neither of them as in play at all. The

---

43 There is an interpretive difficulty with Mill here, since as already noted his account combines elements of a Humean view of causality (which is subject to Bhaskar's criticisms of actualism and Cartwright's of facticity) with elements of a non-Humean view of causality in terms of tendencies, which is closer in spirit to Bhaskar (who uses the same term) and Cartwright, whose more recent work concerns causal capacities (e.g. Cartwright 1989). It is not important for my purposes to resolve the exegetical problems posed by Mill's hybrid, and possibly contradictory treatment.

44 Geach (1973: 103). It appears to me that Bhaskar is being unfair to Geach here, since Geach's view is actually rather close to Bhaskar's own. Nonetheless, whether or not Geach endorsed the thesis in question it is mistaken. Bhaskar also refers to Ryan (1970) which follows Geach on this point.

45 Bhaskar may have had Buridan's rational ass (equidistant between two equally attractive heaps of food, and hence subject to two equal and opposite tendencies) in mind here. Then again, Balaam's ass, in the account in the book of Numbers (Chapter 22), was driven forward by Balaam, but refused to budge since unlike its owner it could see the angel in the road ahead.

missing distinction is that which Bhaskar draws between the exercise and the fulfilment of a tendency. Just as a tendency or power may be present unexercised,[46] a power may be exercised un-actualised: the aspirin may do whatever it does, and yet my headache persist.[47]

It is instructive to consider cases of equilibrium, when in a clear sense the outcome is that nothing happens. With a nil resultant, and given Cartwright's anti-realism about component forces we would be forced to say that no forces at all are in play. This, though, would make it very difficult for us to say what is going on, especially if we consider related but only slightly changed situations. To keep consistency with the above example let us imagine a massive and charged body, which is suspended above the ground. That is to say, on the view I am attempting to save, that it is subject to two forces which balance out. Our understanding of Newton's law gives us to expect that it will be subject to a downwards force of a given magnitude. Our understanding of Coulomb's law gives us to expect an equal force in the opposite direction. The body, as noted, 'does nothing', so the 'resultant' on Cartwright's view and on views closer to Creary's or Mill's is nil. But for Creary, Mill and any realist about components the fact that the resultant has zero magnitude explains the fact that nothing happens: the forces balance out. For Cartwright there are no forces here at all, and no explanation seems possible. Were there, instead, two bodies with equal charges and unequal masses we would be unable to explain why the one moved as it would partly in terms of what it has in common with the other. Or, returning to the single body, were the electric field to be switched off, we would be unable to say that a gravitational force which had been there all along was now acting in isolation.

Viewing these laws (Newton's and Coulomb's) as force laws, which is what they seem to be anyway, supplemented by the notion of composition enables us to take the laws as true without committing ourselves to the excesses of the facticity view. We can say what the difference is between cases where the laws do enable prediction and those where they do not, and make sense of reinforcement, counteraction, equilibrium and other consequences of multiple tendencies being in play. The difference, briefly, is that it is only in some cases that we know how to work with a multitude of influences which may be reinforcing and counteracting one another. Contra Cartwright the outcomes of situations where more than one law are is play is not a counterexample (Cartwright 1980: 77) to either law, but an instance of its operation, which need not be identified with its effects. The burden of argument here is now even more clearly with the Cartwrightian, who has no independent argument for the unreality of the component forces. There is, though, the matter of what Cartwright says about sums and parts.

---

46 It 'remains true to say of a Boeing 727 that it … has the power to … fly at 600 m.p.h. even if it is safely locked up in its hangar' (Bhaskar 1975: 87)

47 This is one of Cartwright's examples, although I am using it in a way she would not. See Cartwright (1989: 3, 136, 141).

Cartwright insists repeatedly, although without any real justification, that vector components cannot be parts of vector sums, at least where the vectors represent forces, motions and perhaps other physical factors. On her view, as noted, 'no pure north motion can be part of a motion which always heads northeast' (1980: 79). The closest she gets to actually defending this view is her claim that 'we learn this from Judy Jarvis Thomson's *Acts and Other Events*' (Cartwright 1980: 79). This is not an argument, and in any event does not seem to be true. Thomson (1977) is a work about actions and events, and, in large measure, about the question of how best to think of parts of events. Thomson's main concern is with events which have temporal parts: consider a shooting which brings about a death after some delay. Is the whole process a killing? Is the shooting then identical with the killing itself, or a part of the killing? Thomson defends the thesis that events can have temporal parts, as long as the part 'plays a role' in the event of which it is a part (1977: 60). Thomson just does not argue that events cannot have spatial parts.[48]

Cartwright's concern here, though, is similar to one which vexed Russell (1903: 477) who considered a case in which there are three particles A, B and C, with B and C both exerting a force on A, and we would normally expect the resultant force to be the vector sum of each separate force. Russell objects as follows:

> But this composition is not truly addition, for the components are not *parts* of the resultant. The resultant is a new term, as simple as their components, and not by any means their sum. Thus the effects attributed to B and C are never produced, but a third term different from either is produced. […] Thus we reach an antinomy: the whole has no effect except what results from the effects of the parts, but the effects of the parts are non-existent.

This is close to Cartwright's view, although she does not express it as an antinomy. Nagel (1963)[49] objects against Russell that he is operating with a needlessly restrictive notion of 'sum', for which, furthermore, he has given no justification. Nagel contends that 'no antinomy arises from the supposition that, on the one hand, the effect which each component force would produce were it to act alone does not exist, while on the other hand the actual effects produced by the joint action of the components is the resultant of their partial effects' (1963: 142). From Nagel's point of view the 'sum' of more than one thing of some kind will not necessarily have the constituents as literal parts, as is the case with length, but will more generally be determined by some function.[50] What function exactly applies in any type of case will be an empirical question

---

48 There is much more that could be said here about Thomson's analysis, which not only does not do what Cartwright says, but is also framed in terms of a notion of causality which is far closer to the facticity view Cartwright is rejecting (or to Bhaskar's actualism) than to what either is defending. Rather than get side-tracked, though, I pursue the sums and parts question by reference to treatments within the philosophy of science.

49 See also Nagel (1961: 380-396). Most of this material appears in Nagel (1963).

50 In fact it can be the case that working out the function is itself a significant scientific achievement: Consider the formulae for adding velocities and other quantities which are central to the special theory of relativity. Cartwright's remarks about vector addition suggest that in cases where velocities in the same direction are

on Nagel's view, and he makes clear that the history of science teaches us to expect some variety: not all addition is linear, not all retains the added elements as proper parts, etc. Cartwright's demand that the parts be present in the sum is like the demand of a customer that a withdrawal from her bank account comprise the actual notes and coins that she herself deposited.

Not only that, forces have direction and magnitude, and can thus unproblematically be represented by vectors. The standard operations for adding vectors work perfectly well in at least some cases where the vectors are forces. Millikan's experiments to determine the charge on a single electron, the calculation of where a satellite or spacecraft will go under the influence of both its own engines and gravity, Coulomb's own experiments on the force due to charge, the paths of projectiles, powered flight, and so-on. But all of these are cases which fall in areas where Cartwright urges us to regard the laws as false, since they do not meet the requirements of the facticity view. That is to say, already, that there is more than good enough reason to regard the scope of laws as extending some of the way beyond what Cartwright allows. Furthermore, it would be unacceptably *ad hoc* to limit the validity of the laws at this point. Newton's law is a law about masses – we know that it holds more than well enough when we have isolated gravity from other influences, and we also know that in some other cases where other forces are in play, it holds too. We have no reason to believe that masses other than those we have examined, or masses in configurations other than those we can compute, are different *qua* masses. So the responsible conclusion is that the law holds, but we are uncertain how to apply it, or how to factor in the other influences. To say that the law cannot apply because we cannot do the sums, or because other factors will make the outcome different to what it would be if the law acted alone is to get the epistemological cart before the ontological horse.

Cartwright says she does not know why there are tracks in a bubble chamber unless there are electrons in the chamber (1983: 92). In the light of what has been said here I can reply that I don't know how we know the charge on electrons (so that we can say what they would do in a bubble chamber) at all unless we can compose forces.

There is one further argument to be made in the present sub-section, and it has already been made by Mill. Fundamental laws have been salvaged at the expense of arguing that the kind of facts which they state are ones which only rarely permit predictions. This may make it seem as though they are, even if possible candidates for truth, a poor substitute for empirically accurate phenomenological laws. Granting the predictive power and usefulness of phenomenological laws, though, does not mean relinquishing any sense of the value of the fundamental variety.

'added' that no addition takes place unless it does so in terms of Galilean relativity rather than the special theory. See Einstein (1905).

Mill argues, in fact, that the most general formulation of a law is, from one point of view, always the best. If his arguments are correct then we can also head off the simulacrum account of explanation, with a defence of the possibility of fundamental laws explaining in virtue of being true. He points out that all 'laws of causation are liable to be counteracted or frustrated' and hence that a law which 'embraces a greater number of instances' will be more likely to hold in any cases at all than one which does not (Mill 1973: 466).[51] In a passage almost designed for the present discussion Mill presses the point by arguing that:

> The same causes, acting according to the same laws, and differing only in the proportions in which they are combined, often produce effects which differ not merely in quantity but in kind (Mill 1973: 468)

What he has in mind here is the application of Newton's laws to celestial motions. Kepler's laws are fairly good for planetary motions, and Newton's laws approximate their predictions, while Newton himself used Kepler's in the course of developing his own. But Newton's also apply to bodies which do not travel in elliptical paths, from comets to cricket balls. What is different in these other cases is not the nature of motion, or the manner of acting of masses, but the particular configurations and conditions:

> The law, in short, of each of the concurrent causes remains the same, however their collocations may vary; but the law of their joint effect varies with every difference in the collocations (Mill 1973: 469).

That Newton's laws explain more than Keplers' is, then, not a matter of Newton's being some high level simulacrum which we have no reason to take seriously beyond the limits within which those laws satisfy the facticity requirement. We know already that the laws are valid beyond the facticity requirement by examining how they hold true in cases where forces are composed. Newton's laws explain with this range and generality because they are true, or at least far more nearly true than Kepler's less fundamental laws of celestial motion, Galileo's less fundamental laws of terrestrial motion, and so forth. That Newton's laws can be placed into a different context as a result of considerations arising from the general theory of relativity is hardly an argument against the position I am defending here, which favours of laws of maximal generality.

The treatment developed here applies in cases other than those of composing forces. This is so despite the fact that in Cartwright (1980) and again in Cartwright (1983) following her discussion of Creary (1981) Cartwright discusses the energy levels in the ground state of the carbon atom with a view to showing how 'real' composition of causes operates in a way quite unlike that imagined by those who add vectors. Her discussion (Cartwright 1983: 67-69) is not really a separate argument at all, though, since it crucially relies on her approach to the parts and sums question as she has already deployed it with reference to the question of vector addition.

---

[51] Mill gives as an example the law 'that the contact of an object causes a change in the state of the nerve, is more general than the law that contact with an object causes sensation, since, for aught we know, the change in the nerve may equally take place when, from a counteracting cause, as, for instance, strong mental excitement, the sensation does not follow...'(Mill 1973: 466-7).

There are three different approaches to the problem of modelling the carbon atom (see figure 1). One (central field approximation) predicts a single energy level (a). Another counteracts the effects of an averaging step taken in the first by 'considering the effects of a term which is equal to the difference between the exact Coulomb interaction and the average potential used in stage one' (Cartwright 1983: 68). This results in the three levels shown at (b). The third approach takes account of the fact, ignored in the second, that the 'internal angular momentum of the electron couples with its orbital angular momentum to create an additional potential' (1983: 68) with the result that one of the three levels predicted by the second technique is replaced by three different levels, giving a total of five.

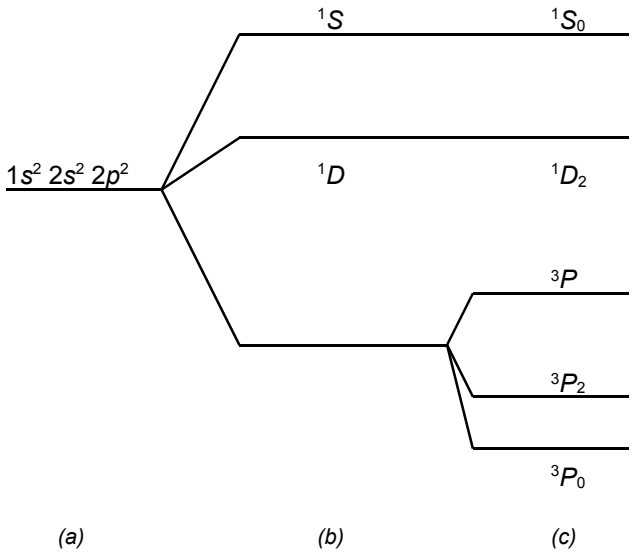


Figure 1: Energy levels in the ground state of carbon, after Cartwright (1983: 67).[52]

It is clear enough where this is going: the five-level model (which Cartwright is prepared to grant is true) is a combination of different potentials. This explains the final five levels, but Cartwright argues that we cannot regard the explanation as making use of laws, because the laws aren't true:

> But (as with the forces 'produced by gravity' in our earlier example) the levels that are supposed to be produced by the Coulomb potential are levels that do not occur. In actuality five levels occur, and they do not include the three levels of (b). In particular, as we can see from [the] diagram, the lowest of the three levels – the ^3P – is not identical with any of the five. In the case of the composition of motions, Mill tried to see the 'component' effects as parts of the actual effect. But that certainly will not work here.

[52] Cartwright follows Messiah (1961) from where she gets her version of this figure.

> The ^3P level in (b) may be 'split' and hence 'give rise to', the $^3P_0$, $^3P_1$, and $^3P_2$ levels in (c); but it is certainly not a part of any of these levels (1983: 69).

There is nothing to worry about here, though. The five-level (c) model is unproblematically the sum of the lower level and the additional consideration of spin-orbit coupling. Start with (c) and assume that no coupling takes place (i.e. set its effect to zero) and you'll get (b). Take (b) and add in the coupling effects and you'll get (c). The same goes for the relationship between (a) and (b). As with vector addition it is possible to recover the parts by subtraction. The kind of addition involved is perfectly precise and rigorous, and presents difficulties only to those who want to see the literal parts present in the sum.

This last point is important because it also shows that fundamental laws other than force laws can be defended against Cartwright's line of attack. Taking the Bhaskarian line and viewing laws as tendencies in no way means restricting tendencies to the exertion of forces: physical tendencies can relate to energy, energy levels, wavelengths, entropy or any other physical quantity, as long as there is good reason to think that the tendency expresses a law.[53]

I think that the arguments developed against Cartwright here are sufficient for my purposes. If fundamental laws are candidates for truth, against her claims both that they are in fact false, and have to be false to play their characteristic role in explanation, then the question whether physics is complete can properly be posed. Not only that, it can provisionally be answered on the basis of the available evidence, which means for my purposes that the arguments of Chapters Two and Three above can be allowed to stand.

6. Conclusion

As noted above, Hacking sums up his brand of causal-interactive realism with the slogan 'If we can spray them, they're real' (1983: 22). What gets sprayed in the particular instance Hacking has in mind are sub-atomic particles, and his point is, briefly, that when things have progressed to the stage that working scientists can describe themselves as producing a stream of such particles for some purpose and spraying them where they want, that philosophical worries about the reality of the entities in question are unwarranted. It's not spraying that does the work here, but the level of causal comprehension and instrumental control which enables spraying to take place. As also noted, Cartwright urges us to be what she calls more 'cautious' and say rather '*When* we can spray them, they're real' (1994: 277).

---

[53] Although not central to my purposes here, it is worth pointing out that the notion of tendencies as laws is one which Cartwright rejects in her more recent work, but which many critics think would save her position from serious difficulties. To say that something has a power which it may or may not exercise does not explain unless we can also explain why the power was exercised in any given case. As long as we do not think of the exercise of a power in line with the facticity view, though, there is no problem with regarding fundamental powers as laws. This approach is in line with Bhaskar, and informs the criticisms of, e.g. Cartwright (1995a) by Morrison (1995), Eels (1995) and Glennan (1997).

What kind of 'caution' is this though? I am reminded of the fact that Victor Gollancz is said at one time to have been unusually anxious about the possibility of his penis having disappeared, and was prone to making regular examinations to establish its ongoing presence (Edwards 1987: 382). (The inspections often took place near his office window, in clear view of those in the building opposite.) Such apparently neurotic behaviour is not easily distinguished from Cartwrightian epistemological caution. In the absence of suitable causal contact, or failing the presence of conditions which could justify the claim to know that some capacity was at work, the Cartwrightian about scientific epistemology is in the same boat as Gollancz's anatomical epistemologist.

My conclusion, then, is that reasoning along the lines suggested by Bhaskar shows the way to reasonable confidence in the *generality* of our experimentally gained knowledge, and hence provides a serious alternative to patchwork realism. The point made earlier which I wish to return to now is that this moderate fundamentalism is enough to get the completeness thesis off the ground. The reason for this is that if we can indeed reasonably take some of our knowledge to be general, we can meaningfully ask whether only physical capacities (associated with mass, charge and the like) are relevant to physical outcomes, or whether non physical ones (for example wealth or war) sometimes play a role. We can also set about trying to establish what kinds of evidence might count either way.

The better part of caution here is not hostility to laws. Cartwright is entirely correct to warn us of the dangers of understanding laws along the lines of the facticity view, but that is not the only way of thinking of laws as being true. The approach defended here does full justice to the ways laws are tested, and also used in explanations, including cases where more than one law is in play. And if laws have been successfully defended, then the arguments developed in Chapter Two can be allowed to stand: as long as there is a case for thinking of causal closure in terms of laws, there is ample reason to think that physics is complete.

Conclusions

My conclusion, in a nutshell, is that the Completeness Thesis can be coherently formulated, and that on balance the arguments and evidence for it are considerably stronger than those against it. Physics is, it appears, complete.

In the course its defence over the preceding chapters the Completeness Thesis has been subjected to a range of vigorous and serious criticisms. Some it has survived more easily than others, and in various ways it has been necessary to modify features of the way it is expressed, and to abandon some traditional associations. In what follows I would like briefly to review the major conclusions of the preceding four chapters, and then to note an especially pressing way in which important work remains to be done, towards which work this thesis as it stands contributes very little.

6.1. Review of Conclusions

A. Formulation

Although there are critical various positions maintaining that neither physicalism nor the Completeness Thesis can get off the ground because of difficulties with specifying what is to count as 'physical' it is these positions themselves which are the non-starters.

As was seen in Chapter One, most of them fail because they give no compelling justification for the amazingly high standards they set for an acceptable account of the physical, whether the demand is for prescience about the future development of physics, the capacity to specify what would count as physical in any possible world, or some other apparently attractive but actually unnecessary criterion. Most of the work which needs to be done to defuse such criticisms is done by pointing out that these demands are unreasonable, at least in the event that the argument at stake concerns the actual world, and takes seriously the open ended state of physics.

If there are any debates which require the kind of prescience or counterfactual elasticity demanded by critics such as Daly (1998) then I have not joined them here. My concern throughout has been with the credentials of the Completeness Thesis at the actual world. As it turns out it is not particularly important to come up with a positive characterisation of the physical in order to pursue this question. Given the repeatedly noted fact that physicalism is typically a thesis about two sets of properties where the non-physical set is in some way determined by the physical one, the fixing of the *boundary* between the two sets is more important than the determination of the shared features, if any, of the members of either.

It is for this reason that a partly negative characterisation of the physical turned out to be suitable for my purposes. The solution to the formulation problem adopted in Chapter One was similar to Meehl and Sellars' (1956) *physical*$_2$ which makes the key characteristic of the physical 'inorganic', and was motivated by arguments similar to those to be found in Papineau (1993) and Spurrett and Papineau (1999). Taking some class of pre-theoretically given effects as physical, physics is first made trivially complete by stipulating that anything needed to fix the likelihoods of those and similar effects is physical, then an empirical edge is added by hypothesising that some kind of properties will not need to be used to fix the chances of physical effects. In the event that it turns out that this kind of physical is complete, it will be possible to argue for physicalism about the properties which are not needed.

A feature of any negative account of the physical, especially one formulated with a view to expressing the Completeness Thesis, is that, *ceteris paribus*, the more that is excluded, i.e. the less that is physical, the more interesting and significant the Completeness Thesis would be if true. Papineau (1993) opts for the non-mental as the key restricting factor. Meehl and Sellars (1956) chose the inorganic. For my purposes I pushed the exclusion slightly further, so as to exclude independent *chemical* causes from playing any role in bringing about physical effects.

The upshot of this is that the version of the Thesis which I set out to investigate was one which asserted that no physical effects would have their chances of occurring affected by any peculiarly chemical, biological or psychological factors. This approach both settled the formulation question, and also made clear what would have to be the case for the Completeness Thesis to be shown to be false: there would have to be good reason to think that the causal ancestry of some physical outcomes included one or more of the excluded factors.

B. Credentials

The first challenge to the Completeness Thesis which I considered was that represented by strong emergentism, where strong emergentism is that form of emergentism which holds that downwards causation from the non-physical (usually chemical or biological) to the physical takes place. If strong emergentism is true, the Thesis is false.

The prospects for a philosophical argument either for or against strong emergentism seem poor. The arguments against it which I found assumed the truth of the Completeness Thesis, which is exactly what is supposed to be at issue. The argument for strong emergentism which I considered, due to Bhaskar (1978) was deeply flawed, even though interesting from the point of view of the question of how to understand the laws of nature. With the traditional emergentists such as Mill (1843) I take the question of the truth of emergentism to be an empirical one.

In empirical terms the case for strong emergentism is very weak. A survey of the most relevant evidence in nineteenth century biology, energy physics, thermodynamics, in early twentieth century microphysics and a variety of other areas especially those relating to self-organisation and complex systems, showed how again and again hypothesised non-physical causes turned out to be unnecessary to account for the phenomena of chemical bonding and life.

For the most part the contemporary revival of interest in emergence brings very little which need worry the defender of the Completeness Thesis. Most of the work which goes on under the label emergence today concerns weak emergence, which is to say that there is no suggestion of a breakdown of causal closure. In fact, since the phenomena of weak emergence are typically higher level structures, complex processes and forms of organisation, this work actually supports the Completeness Thesis, by explaining these phenomena in a way entirely compatible with the truth of the Thesis.

In contrast with emergentists, who tend to see the world in a hierarchical fashion with the physical in a foundational position, the remaining two challenges to the Completeness Thesis (Dupré, Cartwright) which I considered urge us to see the world as comprehensively disordered, both for reasons related to their analyses of the state of our knowledge of the world.

Dupré's major argument concerned problems of taxonomy, and he argued, against an essentialist conception of natural kinds, that the world, especially the world of living things, does not uniquely and neatly divide into a collection of kinds and sub-kinds in a way related to the structural and ultimately physical features of things, but rather forms a heterogeneous mixture of overlapping kinds which do not typically have essences. Dupré urges realism about this mixture of kinds, and thence, via a claim on behalf of the causal capability of them, to the rejection of any kind of causal completeness at all.

His line of argument cannot get him to the rejection of causal completeness, and hence the Completeness Thesis, though. This is because the association he takes for granted between completeness and reductionism is not necessary. As it happens completeness is not only *compatible* with an almost arbitrary degree of taxonomic disorder at non-fundamental scales, it is also the case that well established work in dynamics and abstract models of physical processes show that the assumption of causal closure brings considerable capacity to *explain* such disorder. That this is so does not show that the Completeness Thesis is true, but it does show that Dupré's work does not count as a criticism of the Thesis, subject to the Thesis not being tied to any general commitment to reductionism.

In contrast to Dupré, who argues from large scale issues of taxonomy, essentialism and reductionism, to the rejection of causal completeness, Cartwright develops a line of criticism aimed directly at fundamental physical laws, which she argues are false. The reasons she gives for rejecting the 'facticity' view of laws are similar to Bhaskar's for rejecting what he calls actualism. Both agree that laws cannot be seen as true and also

taken to be statements of event regularities. Where Bhaskar argues that laws are true statements of something other than regularities, Cartwright concludes that fundamental laws should be held false, except in the unusual conditions in which they do 'state the facts'.

Broadly Bhaskarian considerations suggest that Cartwright's conclusions can be resisted. If laws are regarded as statements about tendencies then they can be held true, without falling into the excesses of the facticity view/actualist view of laws. An important test case for the view defended here is the composition of forces. Cartwright argues that where two tendencies are in play, both of the fundamental laws concerning the forces in question are shown to be false. Regarding laws as tendencies, though, enables all of the laws in play in a given case to be regarded as true, even though none individually describes what will happen.

Were Cartwright's criticisms of fundamental laws to go through, the arguments of Chapters Two and Three would have been in serious trouble, since both traded heavily on the notion of the general validity of such laws. Given that a truth preserving account of fundamental laws could be developed which was immune to Cartwright's criticisms, though, those arguments could be allowed to stand, with the result that the Completeness Thesis was left in a very strong position indeed.

6.2. What Next?

In various ways the argument developed here has inclined to some form of functionalism. This was especially so with reference to my responses to the criticisms of the Completeness Thesis contained in and implied by the work of Dupré. One contemporary line of criticism of functionalism, associated especially with Kim (e.g. 1998), alleges that functionalism, in common with supervenience theories, tends to collapse into some form of epiphenomenalism. As a result of that failing, and consistent with his noted commitment to the Completeness Thesis, Kim urges a return to a more reductionist approach to physicalism.

Given the ways in which I have argued that the Completeness Thesis needs to be modified and reformulated this poses a significant difficulty. If the only way to save the Thesis is, as I have argued, to abandon reductionism in favour of something close to functionalism, and functionalism collapses into epiphenomenalism then it seems that I have defended epiphenomenalism. If, on the other hand Kim is correct that we need to be more reductionist and I am also correct that a reductionist version of the Thesis is untenable, then Kim's proposal, whatever its other merits, can hardly be expected to point the way to a preferable physicalism.

This looks like a dilemma: to save the Thesis I went to some lengths to distance it from reductionism and associated it broadly with functionalism, but functionalism now threatens epiphenomenalism. So should I bite the epiphenomenalist bullet? That seems entirely unacceptable. On the other hand to move closer to reductionism would mean to undermine the case developed against Dupré. This is so because my argument

against Dupré conceded much of the force of his anti-reductionism by emphasising the fact, and a fact it is, that completeness need not entail reductionism at all. So to move towards reductionism would be to undermine my own case for completeness, which would hardly solve the present difficulty. I think that there is a serious challenge for the further development of the physicalist programme here, and want to conclude by explaining very briefly how I see things standing on this point, since this is both an area of active contemporary work, and also one which the arguments developed here throw into stark relief.

I should emphasise that were my interest here the defence of the Completeness Thesis I could simply re-emphasise the distinctness of the completeness and reduction questions. What concerns me at the moment is not the defence of the Thesis, though, but the question of identifying a particular philosophical challenge made pressing by what I take to be the successful defence of the Completeness Thesis developed above.

Why does Kim think that functionalism is faced with a slide into epiphenomenalism? Well, the functionalist wants to be able to say that instances of mental properties can cause instances of other mental properties. As token physicalists functionalists also think that each instance of a mental property is realised by some physical property. So if *M* causes *M*, M* is realised by *P, M** is realised by *P*,* and we accept the Completeness Thesis what follows? In the first place *P* causes *P**, which means that *M** is realised as a result of *P*. But then why have a separate functionalist causal claim to the effect that *M* cased *M**, when it seems as though once *P* then *M** was going to happen anyway? (following Kim 1998: 38-56).[1]

Kim's response to this and other difficulties with non-reductive physicalism is to defend a renewed commitment to reductionism, not on the Nagelian model, but by means of 'functionalisation', where the properties to be reduced (say mental ones) are described in relational and extrinsic terms, so as better to permit their identification with the physical properties which realise them, and the reduction is seen as taking place locally where properties which can be suitably functionalised are regarded as reduced to the physical realisers of those functions (Kim 1998: 99, 118-119).

Functionalised properties are nonrigid, since the causal connections in question will be metaphysically contingent, so that mind-body identity will not become a necessary truth along Kripkean (e.g. 1980) lines, and hence objectionable for making what seems like it should be a contingent matter a necessary one. The kind of reduction in question is very localist – given the force of the multiple realisation argument which gets functionalism off the ground, there is no reason to expect that general reductions will be possible for properties which could be realised in a variety of ways. Kim himself seems reasonably comfortable with the consequences of this, describing the upshot of his arguments as follows:

---

[1] There is a lot in Kim's wider argument which I am eliding here, so as to focus on the central question of reductionism as it relates to the concerns of the present work. I do not pretend to offer a general engagement with Kim's recent work here. See also Kim (1989).

> In this way multiply realized properties are sundered into their diverse realizers in different species and structures, and in different possible worlds (Kim 1998: 111).

This approach saves some of functionalism by making it more explicitly physicalist, but at the expense of losing the capacity to say what it is that makes some apparently similar functional properties related in cases where the realisations are significantly different. My pain may not be the same kind of thing as yours, and mine *now* need not even be the same kind of thing as the pain I had a few minutes ago, even functionally speaking.[2] Kim notes that to those who might want to 'hang on to' these properties as 'unified and robust … in their own right' this will be a 'disappointment' but also maintains that the conclusion in question is 'inescapable' (1998: 111).

The kind of reductionism Kim endorses is significantly different from the megalomaniac version which Dupré, or for that matter Bhaskar, set out to reject. This means that we are not faced with a situation where in order to save functionalism we need to come up with a new way of rejecting Dupré's line of criticism of the Completeness Thesis, those arguments can be allowed to stand. (Since the type of reduction proposed by Kim is much more modest and local that that attacked by Dupré and Bhaskar, it is immune to most of their criticisms which presuppose a more extreme form of reductionism.) Nonetheless I think that the price which Kim is willing to pay may well be too high. Let us leave mental properties for a moment, and with Fodor consider the case of money, and also a possibly true law of economics: Gresham's law that bad money drives out good. Recall Fodor's appeal for the implausibility of general reduction along Nagelian lines about money:

> But banal consideration suggest that a description which covers all such events must be wildly disjunctive. Some monetary exchanges involve strings of wumpum. Some involve dollar bills. And some exchanges involve signing one's name to a check. What are the chances that a disjunction of physical predicates which covers all these events (i.e. a disjunctive predicate which can form the right side of a bridge law of the form "x is a monetary exchange [□][3] ...") expresses a physical natural kind? In particular what are the chances that such a predicate forms the antecedent or consequent of some proper law of physics? (Fodor 1974: 56)

It is not important for my purposes whether or not Gresham's law is true, or if it is whether it is best thought of as a causal law.[4] What is important is the question of multiple realisation. If we restrict ourselves to those economic exchanges naturally described in terms of 'money', it is clear enough that at least two things are generally true. The first is that there are many ways in which money can be realised. The second is that money

---

2 I am not especially concerned about properties which resist functionalisation, such as qualia, but rather with the apparent unity of some kinds of phenomena even though they may be variably realised.

3 Recall that in Fodor's original this symbol appears as a pair of horizontal arrows facing in opposite directions. I have used the more conventional symbol for a biconditional. (See Chapter Three section 7.2).

4 If it is both true and causal, then it would be possible to run the identity argument discussed in the Introduction above, section (1.1).

is translatable or transferable between them. (This second point is as good as definitive of money – if some putative 'dollar' cannot be converted into cash, say, then it is not a dollar at all.[5]) We have cash, cheques, bank balances, electronic forms of payment and more besides.

It is one thing to grant that the physical realisers in each case are different – that point is not up for dispute here – but another entirely to say that money is not in some significant sense *one* kind of thing in all these cases. Kim is, it seems, willing to say that there is not one kind of thing which money is, because multiple realisation without functionalisation leads to epiphenomenalism. On Kim's view we should accept the alternative locally reductive notion of functionalism which he supports, and swallow our 'disappointment' over the loss of true general claims about functional kinds.

I think that the price Kim is willing to pay to is too high: it cuts at the foundations of explanatory capabilities and forms of description which we actually have, and which have considerable merit. When I am owed a definite sum of money I really am properly considered paid in the same sense whether I am given cash or a cheque, or am the recipient of a direct electronic transfer. Looked at another way, Kim claims to take multiple realisation seriously, but it would be more correct to say that he abandons it. To take it seriously is to hold onto the notion that there is some one thing which is realised in many ways, not to say that there are in fact many different things which are each realised in different ways depending on how they are functionalised. We should not accept a solution to a problem which does violence to our manifest explanatory resources without attempting to find some way of accommodating and doing justice to them first.

So I want to say that we need something which stands in a sense *between* standard Fodorian functionalism and the more reductive version which Kim supports. I think that we need it to do justice to the fact that there are many kinds of thing, for example money, which seem to have something important in common no matter how they are realised. I also do not know how to start thinking about this challenge, since as things stand to move closer to standard functionalism means to face the danger of epiphenomenalism, and I think Kim is correct about how that follows from standard functionalism. Nonetheless, assuming that the arguments of the preceding chapters are broadly supportable, then it seems to me that the challenge I have just described is one worth taking on.

Physics is complete, but that's not the end of it.

---

[5] It Deepak Mistrey who suggested to me that the case of money was of particular interest with respect to these questions.

References

Albert, D. Z. 1992. *Quantum Mechanics and Experience*, Cambridge, Mass.: Harvard University Press

Alexander, S. 1920. *Space, Time, and Deity*. 2 Vols, London: Macmillan.

Anderson, P. W. 1972. 'More is Different: Broken symmetry and the hierarchical structure of nature', *Science* 177, number 4047, pp. 393-396.

Anscombe, G. E. M. and Geach, P. T. 1973. *Three Philosophers: Aristotle, Aquinas, Frege*, Oxford: Blackwell.

Aristotle. 1936. *Minor Works,* (translated by W. S. Hett.), London, Heinemann.

Armstrong, D. 1978. *Universals and Scientific Realism, Vol. II. A Theory of Universals*, Cambridge: Cambridge University Press.

Armstrong, D. 1980. *The Nature of Mind*, St Lucia, Queensland: University of Queensland Press.

Auyang, S. 1998. *Foundations of Complex-System Theories in Economics, Evolutionary Biology and Statistical Physics*, Cambridge: Cambridge University Press.

Bain, A. 1870. *Logic,* London: Longmans, Green, Reader and Dyer.

Barrow, J. D. 1992. *Pi in the Sky: Counting, Thinking and Being*, London: Penguin.

Bechtel, W. and Abrahamsen, A. (1991) *Connectionism and the Mind*, Oxford: Blackwell.

Bechtel, W. and Richardson, R. C. 1992. 'Emergent Phenomena and Complex Systems', in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 257-288.

Beck, W. S. 1961. *Modern Science and the Nature of Life*, London: Penguin.

Beckermann, A. 1992a. 'Introduction – Reductive and Nonreductive Physicalism' in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 1-21.

Beckermann, A. 1992b. 'Supervenience, Emergence, and Reduction' in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* (1992) Berlin: de Gruyter, pp. 94-118.

Beckermann, A., Flohr, H. and Kim, J. (eds.) 1992. *Emergence or Reduction: Essays on the Prospects of Nonreductive Physicalism*, New York: de Gruyter.

Bedau, M. A. 1997. 'Weak Emergence', *Philosophical Perspectives,* 11, *Mind, Causation, and World*, Oxford: Blackwell, pp375-399.

Bell, J. S. 1964. 'On the Einstein Podolsky Rosen Paradox', *Physics*, 1, pp. 195-200.

Bernal, J. D. 1939. *The Social Function of Science*, London: George Routledge & Sons.

Bhaskar, R. 1978. *A Realist Theory of Science*, (Second Edition) Sussex: The Harvester Press.

Bhaskar, R. 1989a. *The Possibility of Naturalism*, (Second Edition) London: Harvester Wheatsheaf.

Bhaskar, R. 1989b. *Reclaiming Reality*, London: Verso.

Bhaskar, R. 1991. *Philosophy and the Idea of Freedom*, Oxford: Blackwell.

Blackburn, S. 1993. *Essays in Quasi–Realism*, Oxford: Oxford University Press.

Block, N. 1980. 'Troubles with Functionalism', in Block, N. (ed.) *Readings in the Philosophy of*

Psychology, Volume 1, London: Methuen, pp. 268-305.

Bohm, D. 1952. 'A Suggested Interpretation of the Quantum Theory in terms of "Hidden" Variables' (Parts I and II), in J. A. Wheeler and W. H. Zurek, (eds.) (1983) *Quantum Theory and Measurement*, Princeton: Princeton University Press, pp. 369-396.

Bohm, D. 1980. *Wholeness and the Implicate Order*, London: Routledge and Kegan Paul.

Bohm, D. and Hiley. B. J. 1993. *The Undivided Universe*, London: Routledge.

Boyd, R. 1980. 'Materialism without Reductionism: What Physicalism Does not Entail', in Block, N. (ed.) *Readings in the Philosophy of Psychology, Volume 1*, London: Methuen, pp. 67-106.

Broad, C. D. 1925. *The Mind and Its Place in Nature,* London: Routledge and Kegan Paul.

Broad, C. D. 1927. *Sir Isaac Newton: Annual Lecture on a Master Mind*, London: Oxford University Press for the British Academy.

Broad, C. D. 1934. *Inaugural Lecture: Determinism, Indeterminism, and Libertarianism*, Cambridge: Cambridge University Press.

Brooke, J. H. 1991. *Science and Religion: Some Historical Perspectives*, Cambridge: Cambridge University Press.

Campbell, D. T. 1974. ''Downward Causation' in Hierarchically Organized Biological Systems', in: Ayala, F. J. and Dobzhansky, T. (eds.) *Studies in the Philosophy of Biology*, Los Angeles: University of California Press, pp. 179-186.

Cartwright, N. 1980. 'Do the Laws of Physics State the Facts?', *Pacific Philosophical Quarterly* 61, pp. 75-84.

Cartwright, N. 1983. *How the Laws of Physics Lie*, Oxford: Clarendon Press.

Cartwright, N. 1989. *Nature's Capacities and their Measurement*, Oxford: Clarendon Press.

Cartwright, N. 1993. 'Is Natural Science "Natural" Enough? A Reply to Phillip Allport', *Synthese* 94, pp. 291-301.

Cartwright, N. 1994. 'Fundamentalism vs. the Patchwork of Laws', *Proceedings of the Aristotelean Society*, XCIV (1994), pp. 279-292.

Cartwright, N. 1995a. 'Précis of *Nature's Capacities and Their Measurement*', *Philosophy and Phenomenological Research*, 55, pp. 153-156.

Cartwright, N. 1995b. 'Reply to Eells, Humphreys and Morrison', *Philosophy and Phenomenological Research*, 55, pp. 177-187.

Causey, R. L. 1987. *Unity of Science*, Dordrecht: D. Reidel.

Chalmers, A. 1987. 'Bhaskar, Cartwright and Realism in Physics', *Methodology and Science,* 20, pp. 77-96.

Chalmers, A. 1993. 'So the Laws of Physics Needn't Lie', *Australasian Journal of Philosophy,* 71, pp. 196-205.

Chalmers, A. 1996. 'Cartwright on Fundamental Laws: A Response to Clarke', *Australasian Journal of Philosophy,* 74, pp. 150-152.

Chomsky, N. 1968. *Language and Mind*, New York: Harcourt Brace Jovanovich, Inc.

Chomsky, N. 1972. *Language and Mind*,(enlarged edition) New York: Harcourt Brace Jovanovich, Inc.

Chomsky, N. 1980. *Rules and Representation,* Oxford: Basil Blackwell.

Churchland, P. M. 1981. 'Eliminative Materialism and the Propositional Attitudes', *Journal of Philosophy*, 78, pp. 67-90.

Churchland, P. M. 1985. 'Reduction, Qualia, and the Direct Introspection of Brain States', *Journal of Philosophy*, 82, pp. 8-28.

Churchland, P. S. 1986. *Neurophilosophy*, Cambridge Mass.: Bradford/MIT Press.

Cilliers, P. 1998. *Complexity and Postmodernism*, London: Routledge.

Clarke, S. 1995. 'The Lies Remain the Same: A Reply to Chalmers', *Australasian Journal of Philosophy,* 73, pp. 152-155.

Coleman, W. 1977. *Biology in the Nineteenth Century*, Cambridge: Cambridge University Press.

Collier, A. 1999. *Being and Worth*, London: Routledge.

Craig, E. 1987. *The Mind of God and the Works of Man*, Oxford: Oxford University Press.

Crane, T. 1991. 'Why indeed? Papineau on supervenience', *Analysis*, 51, pp. 32-37.

Crane, T. and Mellor, H. 1990. 'There is no Question of Physicalism', *Mind*, 99, pp. 185-206.

Creary, L. G. 1981. 'Causal Explanation and the Reality of Natural Component Forces', *Pacific Philosophical Quarterly,* 62, 148-157.

Cummins, R. 1989. *Meaning and Mental Representation*, Cambridge, Mass: MIT Press.

Daly, C. 1995. 'Does physicalism need fixing?', *Analysis*, 55, pp. 135-41.

Daly, C. 1998. 'What are Physical Properties?', *Pacific Philosophical Quarterly,* 79, pp. 196–217.

Davidson, D. 1970. 'Mental Events', in L. Foster and J. Swanson (eds.) *Experience and Theory*, Amherst, Mass.: University of Massachusetts Press, pp. 291-303, reprinted in Moser and Trout (1995) pp. 107-121.

Davies, D. 1996. 'Explanatory disunities and the unity of science', *International Studies in the Philosophy of Science,* vol. 10, number 1, pp. 5-21.

Dear, P. (ed.) 1997. *The Scientific Enterprise in Early Modern Europe: Readings from Isis*, Chicago: University of Chicago Press.

Dear, P. 1995. *Discipline and Experience: The Mathematical Way in the Scientific Revolution*, Chicago: University of Chicago Press.

Descartes, R. 1985. *The Philosophical Writings of Descartes*, 2 vols., (translated and edited by J. G. Cottingham, R. Stoothoff, and D. Murdoch), Cambridge: Cambridge University Press.

Descartes, R. 1989. *The Passions of the Soul*, (translated by S. H. Voss.), Indianapolis: Hackett.

Descartes, R. 1991. *Principles of Philosophy,* (translated and edited by V. R. Miller and R. P. Miller), Dordrecht: Kluwer Academic Publishers.

Descartes, R. 1991. *The Philosophical Writings of Descartes, volume 3*, (translated and edited by J. G. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny), Cambridge: Cambridge University Press.

Descartes, R. 1998. *The World and Other Writings*, (translated and edited by Gaukroger, S.), Cambridge: Cambridge University Press.

Dretske, F. 1995. *Naturalizing the Mind*, Cambridge, Mass., MIT Press.

Driesch, H. 1908. *The Science and Philosophy of the Organism: The Gifford Lectures Delivered Before the University of Aberdeen in the Year 1907 (2 Vols.)*, London: Adam and Charles Black.

Dugas, R. 1954. *A History of Mechanics*, Neuchâtel: Éditions du Griffon. (Reprinted 1988, New York: Dover.)

Duhem, P.-M.-M. 1954. *The Aim and Structure of Physical Theory*, (translated by P.P. Wiener), New Jersey: Princeton University Press.

Duhem, P.-M.-M. 1980. *The Evolution of Mechanics*, (translated by M. Cole and edited by G.AE. Oravas), Alphen aan den Rijn: Sijthoff and Noordhoff.

Dupré, J. 1983. 'The Disunity of Science', *Mind* vol. XCII, 321-346.

Dupré, J. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Cambridge Mass: Harvard University Press.

Dupré, J. 1996a. 'Promiscuous Realism: Reply to Wilson', *British Journal for the Philosophy of Science*, 47, pp. 441-444.

Dupré, J. 1996b. 'The Solution to the Problem of the Freedom of the Will'. In J. E. Tomberlin (ed.) *Philosophical Perspectives*, 10, *Metaphysics*, Oxford: Blackwell, 385-402.

Dupré, J. and Cartwright, N. 1998. 'Probability and Causality: Why Hume and Indeterminism Don't Mix', *Noûs* 22, pp 690-698.

Durkheim, E. 1964. *The Rules of Sociological Method*, (translated by George Catlin), New York: Free Press.

Earman, J. 1986. *A Primer on Determinism*, Dordrecht: D. Reidel.

Eccles, J. 1980. *The Human Psyche: The Gifford Lectures, University of Edinburgh 1978-1979*, New York: Springer.

Edwards, R. D. 1987. *Victor Gollancz: A Biography*, London: Victor Gollancz.

Eells, E. 1995. 'Cartwright on Probabilistic Causality: Types, Tokens, and Capacities', *Philosophy and Phenomenological Research*, 55, pp. 169-175.

Einstein, A. 1905. 'On the Electrodynamics of Moving Bodies' in Einstein, A., Lorenz, H. A., Weyl, H. and Minkowski, H. *The Principle of Relativity*, (1923) translated by W. Perrett and G. B. Jeffery, New York: Dover, pp. 37-71.

Einstein, A., Lorenz, H. A., Weyl, H. and Minkowski, H. 1923. *The Principle of Relativity*, translated by W. Perrett and G. B. Jeffery, New York: Dover

Einstein, A., Podolsky, B. and Rosen, N. 1935. 'Can Quantum-Mechanical Description of Physical Reality be Considered Complete?', in J. A. Wheeler and W. H. Zurek (eds.) (1983) *Quantum Theory and Measurement*, Princeton: Princeton University Press, pp. 138-141.

Elsasser, W. M. 1958. *The Physical Foundation of Biology: An Analytical Study*, London: Pergamon.

Ereshefsky, M. 1995. Critical Notice of Dupré (1993), *Canadian Journal of Philosophy*, 25, number 1, pp. 143-158.

Evnine, S. 1991. *Donald Davidson*, Cambridge: Cambridge University Press.

Faraday, M. 1861. *A Course of Six Lectures on the Chemical History of a Candle*, London: Chatto and Windus.

Faraday, M. 1952. *Experimental Researches in Electricity*, in *Great Books of the Western World volume 45*, Chicago: Encyclopaedia Britannica.

Feigl, H. 1953. 'Unity of Science and Unitary Science' in H. Feigl and M. Brodbeck (eds.) *Readings in the*

Philosophy of Science, New York: Appleton-Century-Crofts, Inc., pp. 382-384.

Feigl, H. 1958. 'The "Mental" and the "Physical"', in Feigl, H. et al (1958) 370-497.

Feigl, H. 1967. *The "Mental" and the "Physical": The Essay and a Postscript,* Minneapolis: University of Minnesota Press.

Feigl, H., Scriven, M. and Maxwell, G. (eds.) 1958. *Concepts, Theories, and the Mind-Body Problem: Minnesota Studies in the Philosophy of Science Volume II*, Minneapolis: University of Minnesota Press.

Feynman, R. 1965. *The Character of Physical Law*, London: British Broadcasting Corporation. [Also published in 1967 by MIT Press, Cambridge, Massachusetts.]

Fiegl, H. and Scriven, M. (eds.) 1956. *The Foundations of Science and the Concepts of Psychology and Psychoanalysis: Minnesota Studies in the Philosophy of Science Volume I*, Minneapolis: University of Minnesota Press.

Fleetwood, S. (ed.) 1998. *Critical Realism in Economics*, London: Routledge.

Fodor, J. 1974. 'Special Sciences, or the Disunity of Science as a Working Hypothesis', *Synthese* 28, 97-115. Reprinted in Moser, P. K. and Trout, J. D. (1995) 51-64.

Fodor, J. 1990. *A Theory of Content and Other Essays*, Cambridge Mass: MIT Press.

Forrest, S. (ed.) 1991. *Emergent Computation*, Cambridge Mass: MIT Press.

Forrest, S. 1991. 'Emergent Computation: Self-organizing, Collective and Cooperative Phenomena in Natural and Artificial Computing Networks - Introduction to the Proceedings of the Ninth Annual CNLS Conference', in Forrest (1991) 1-11.

Galileo, G. 1914. *Dialogues Concerning Two New Sciences*, (translated by Henry Crew and Alfonso de Salvio), New York: Macmillan

Garber, D. 1992. *Descartes' Metaphysical Physics*, Chicago: University of Chicago Press.

Gardner, M. 1970. 'Mathematical Games', *Scientific American*, 223 No. 4, pp120-123.

Ghirardi, G. C., Rimini, A. and Weber, T. 1986. 'Unified Dynamics for Microscopic and Macroscopic Systems', *Physical Review,* D34, pp. 470-491.

Gibbins, P. 1984. 'Nancy Cartwright's New Philosophy of Physics', *British Journal for the Philosophy of Science*, 35, 4, pp. 390-401.

Glennan, S. S. 1997. 'Capacities, Universality, and Singularity', *Philosophy of Science* 64, pp. 605-626.

Gross, P. R. and Levitt, N. 1997. *Higher Superstition: The Academic Left and Its Quarrels With Science*, Baltimore: Johns Hopkins University Press.

Guhl, A. M. 1956. 'The Social Order of Chickens', *Scientific American*, 194: 42-46.

Hacking, I. 1975. *The Emergence of Probability*, Cambridge: Cambridge University Press.

Hahn, R. 1967. *Laplace as a Newtonian Scientist: A Paper delivered at a seminar on the Newtonian Influence held at the Clark Library 8 April, 1967*, Los Angeles: William Andrews Clark Memorial Library.

Hale, B. 1993. 'Physicalism and Mathematics', in Robinson (1993), pp. 39-59.

Hall, A. R. 1962. *The Scientific Revolution 1500-1800: The Formation of the Modern Scientific Attitude*, Boston: Beacon.

Hall, A. R. 1981, [1963]. *From Galileo to Newton*, Dover: New York.

Hall, M. B. 1994. *The Scientific Renaissance 1450-1630*, Dover: New York.

Harman. P. M. 1982. *Energy, Force, and Matter: The Conceptual Development of Nineteenth-Century Physics*, Cambridge: Cambridge University Press.

Harré, R. and Madden, E. H. 1973. *Causal Powers*, Oxford: Blackwell.

Heidegger, M. 1993. *Basic Writings*, (revised and expanded edition, edited by Krell, D.), London: Routledge.

Hellman, G. P., and Thompson, F. W. 1975. 'Physicalism: Ontology, Determination, and Reduction', *Journal of Philosophy*, 72, pp. 551-64.

Hofstadter, D. R. (1982) *Artificial Intelligence: subcognition as computation,* Technical Report 132, Bloomington: Indiana University.

Holland, J. H. 1998. *Emergence: From Chaos to Order*, Oxford: Oxford University Press.

Horgan, T. 1993. 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World', *Mind*, Vol.102, 408, pp. 555-586.

Hume, D. 1975. *Enquiries concerning Human Understanding and concerning the Principles of Morals*, (third edition, revisions and notes by P. H. Nidditch), Oxford: Oxford University Press.

Hume, D. 1978. *A Treatise of Human Nature*, (second edition, revisions and notes by P. H. Nidditch), Oxford: Oxford University Press.

Humphries, P. 1995. 'Abstract and Concrete', *Philosophy and Phenomenological Research*, 55, pp. 157-161.

Humphries, P. 1997. 'How Properties Emerge', *Philosophy of Science*, 64, pp. 1-17.

James, W. 1907. *Pragmatism: A New Name For Some Old Ways of Thinking*, New York: Longmans, Green & Co.

Jolley, N. (ed.) 1995. *The Cambridge Companion to Leibniz*, Cambridge: Cambridge University Press.

Joseph, G. 1985. Review of Cartwright (1983), *Philosophical Review*, XCIV, 4, pp. 580-583.

Kant, I. 1929. *Critique of Pure Reason*, (translated by Norman Kemp Smith), London: MacMillan.

Kant, I. 1985. *Philosophy of Material Nature*, (translated by Ellington, J. W.), Indianapolis: Hackett.

Kant, I. 1987. *Critique of Judgement*, (translated by Werner Pluhar), Indianapolis: Hackett.

Kapp, R. O. 1940. *Science versus Materialism*, London: Methuen and Co.

Kapp, R. O. 1951. *Mind Life and Body*, London: Methuen and Co.

Karl, F. R. and Davies, L. 1983-. *The Collected Letters of Joseph Conrad* (8 volumes), Cambridge: Cambridge University Press.

Kellert, S. 1993. *In the Wake of Chaos*, Chicago: University of Chicago Press.

Kemeny, J. C. and Oppenheim, P. 1956. 'On Reduction', *Philosophical Studies*, 7, pp. 6-19.

Kim, J. 1989. 'The Myth of Nonreductive Physicalism', *Proceedings and Addresses of the American Philosophical Association*, 63, pp. 31-47.

Kim, J. 1992. '"Downward Causation" in Emergentism and Nonreductive Physicalism', in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 119-138.

Kim, J. 1993. *Supervenience and Mind: Selected Philosophical Essays*, Cambridge: Cambridge University Press.

Kim, J. 1998. *Mind in a Physical World*, Cambridge Mass.: MIT Press.

Kincaid, H. 1990. 'Molecular Biology and the Unity of Science', *Philosophy of Science*, 58, pp. 575-593.

Klee, R. 1984. 'Micro-Determinism and Concepts of Emergence', *Philosophy of Science* 51, 44-63.

Konner, M. 1982. *The Tangled Wing: Biological Constraints on the Human Spirit*, London: Penguin.

Koyré, A. 1978. *Galileo Studies*, (translated from the French by Mepham, J.), Sussex: Harvester Press.

Koza, J. R. 1992. 'The Genetic Programming Paradigm: Genetically Breeding Populations of Computer Programs to Solve Problems', in Soucek, B. et al. 1992. *Dynamic, Genetic, and Chaotic Programming: The Sixth Generation*. New York: John Wiley & Sons, pp. 203-321.

Kripke, S. 1972. 'Naming and Necessity', in Gilbert Harman and Donald Davidson (eds.) *Semantics of Natural Language*, Dordrecht: D. Reidel, pp. 253–355.

Kripke, S. 1980 [1972]. *Naming and Necessity,* Oxford: Basil Blackwell.

Kuhn, T. S. 1957. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*, Cambridge: Harvard University Press.

Kuhn, T. S. 1970. *The Structure of Scientific Revolutions,* (2nd edition), Chicago: University of Chicago Press.

Kuhn, T. S. 1977. *The Essential Tension*, Chicago: University of Chicago Press.

Langton, C. G. 1991. 'Computation at the Edge of Chaos: Phase Transitions and Emergent Computation', in Forrest (1991) 12-37.

Laplace, P. S. 1951. *A Philosophical Essay on Probabilities*, New York: Dover.

Lavoisier, A. L. 1952. *Elements of Chemistry*, (translated by Robert Kerr), in *Great Books of the Western World volume 45*, Chicago: Encyclopaedia Britannica.

Leibniz, G. W. 1989. *Philosophical Essays*, (translated by Ariew, R. and Garber, D.), Indianapolis: Hackett.

Levy, S. 1993. *Artificial Life: The Quest for a New Creation*, London: Penguin.

Lewes, G. H. 1874, 1875. *Problems of Life and Mind. (First series) 2 Vols,* London: Trübner, & Co.

Lewes, G. H. 1875. *Problems of Life and Mind. (Third series) 2 Vols,* London: Trübner, & Co.

Lewes, G. H. 1877. *Problems of Life and Mind. (Second series),* London: Trübner, & Co.

Lewis, D. K. 1966. 'An Argument for the Identity Theory', *Journal of Philosophy*, 63, pp. 17-25.

Lewis, D. K. 1983. 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61, pp. 343-377.

Lewis, D. K. 1986 [1973]. *Counterfactuals,* Oxford: Basil Blackwell.

Lloyd Morgan, C. 1923. *Emergent Evolution,* (The 1922-3 Gifford Lectures, volume 1), London: Williams and Norgate.

Lloyd Morgan, C. 1926. *Life, Mind, and Spirit,* (The 1922-3 Gifford Lectures, volume 2), London: Williams and Norgate.

Locke, J. 1975. *An Essay Concerning Human Understanding*, (Edited and with a foreword by P.H. Nidditch), Oxford: Clarendon.

Longworth, G. 1999. 'What Determines the Meaning of Meaning?', unpublished paper.

Lorenz, E. N. 1963. 'Deterministic Nonperiodic Flow', Journal of the Atmospheric Sciences (20) 1963, 130-

141.

Lovejoy, A. O. 1961. *The Reason the Understanding and Time*, Baltimore: Johns Hopkins.

Lowe, E. J. 1993. 'The Causal Autonomy of the Mental', *Mind*, 102, 408, pp. 629-644.

Lucas, J. R. 1961. 'Mind's, Machines and Gödel', *Philosophy*, 36, pp. 112-27.

Lyotard, J-F. L. 1984. *The Postmodern Condition: A Report on Knowledge*, (translated by Bennington, G and Massumi, B.), Manchester: Manchester University Press.

Lyotard, J-F. L. 1988. *The Differend: Phrases in Dispute*, (translated by Georges Van Den Abbeele), Minneapolis: University of Minnesota Press.

Mandelbrot, B. 1977. *The Fractal Geometry of Nature*, New York: Freeman.

Margolis, J. 1987. *Science Without Unity: Reconciling the Human and Social Sciences*, Oxford: Blackwell.

Marquis, J-P. 1998. 'Epistemological Aspects of the Application of Mathematics to Itself', in Anapolitanos, D., Baltas, A. and Tsinorema, S. (editors) *Philosophy and the Many Faces of Science*, New York: Rowan and Littlefield, pp. 183-194.

Marquis, J-P. 1999. 'Mathematical engineering and mathematical change', , *International Studies in the Philosophy of Science*, Vol. 13, No.3, pp. 245-259.

Marx, K. 1972. *Capital 3,* London: Lawrence & Wishart.

Maskill, H. 1985. *The Physical Basis of Organic Chemistry*, Oxford: Oxford University Press.

McGuiness, B. F. (editor) 1985. *Moritz Schlick*, Dordrecht: Reidel.

McLaughlin, B. 1992. 'The Rise and Fall of British Emergentism', in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 49-93.

McLaughlin, B. 1995. 'Varieties of Supervenience', in E, Savellos and Umit D. Yalcin (eds.) *Supervenience: New Essays*, Cambridge: Cambridge University Press.

Meehl, P. and Sellars, W. 1956. 'The Concept of Emergence'. In Fiegl, H. and Scriven, M. (eds.) *Minnesota Studies in the Philosophy of Science Volume 1: The Foundations of Science and the Concepts of Psychology and Psychoanalysis.,* pp. 239-252.

Messiah, A. 1961. *Quantum Mechanics*, Amsterdam: North Holland.

Mill, J. S. 1843. *System of Logic*, London: Longmans, Green, Reader, and Dyer. [8th edn. 1872.]

Mill, J. S. 1972. *A System of Logic Ratiocinative and Deductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, (2 vols.) J. M. Robson (ed.), Volumes VII and VIII of *The Collected Works of John Stuart Mill,* (XXI volumes.) J. M. Robson (General editor), Toronto: University of Toronto Press.

Mill, J. S. 1989. *Autobiography*, London: Penguin.

Morrison, M. 1995. 'Capacities, Tendencies and the Problem of Singular Causes', *Philosophy and Phenomenological Research*, 55, pp. 163-168.

Moser, P. K. and Trout, J. D. (eds.) 1995. *Contemporary Materialism: A Reader*, London: Routledge.

Mott, F. K. 1948. *The Universal Design of Birth: An Analysis of the Configurational Involvement of Birth and Its Relation to Emergence Generally*, Philadelphia: David McKay & Co.

Moyal, G. J. D. (ed.) 1991. *René Descartes: Critical Assessments,* (4 Volumes), London: Routledge.

Nagel, E. 1961. *The Structure of Science*, New York: Harcourt, Brace and World.

Nagel, E. 1963. 'Wholes, Sums, and Organic Unities', in: Lerner, D. (ed.), *Parts and Wholes,* New York: The Free Press.

Newman, D. V. 1996. 'Emergence and Strange Attractors', *Philosophy of Science*, 63, pp. 245-261.

Newton, I. 1696. *Philosophiae Naturalis Principia Mathematica*, London: S. Pepys.

Newton, I. 1934. *Principia,* (2 Volumes), translated by Andrew Motte, revised by Florian Cajori, Berkeley: University of California Press.

Newton, I. 1951. *Optics*, New York: Dover.

Nicolson, M. H. 1960. *The Breaking of the Circle*, (revised edition), Oxford: Oxford University Press.

Norris, C. 1997. *Against Relativism: Philosophy of Science, Deconstruction and Critical Theory*, Oxford: Blackwell.

O'Brien, F. 1964. *The Dalkey Archive*, London: Picador.

Oppenheim, P. and Putnam, H. 1958. 'Unity of Science as a Working Hypothesis', in Feigl, H. et al (1958) 3-36.

Pais, A. 1991. *Neils Bohr's Times, in Physics, Philosophy, and Polity*, Oxford: Clarendon.

Papineau, D. 1990. 'Why Supervenience?', *Analysis,* 50, pp. 66-70.

Papineau, D. 1993. *Philosophical Naturalism*, Oxford: Blackwell.

Peper, S. 1926. 'Emergence', *Journal of Philosophy*, 23, pp. 241-245.

Peters, S. L. 1995. *Emergent Materialism: A Proposed Solution to the Mind/Body Problem*, New York: University Press of America.

Poland, J. 1994. *Physicalism: The Philosophical Foundations*, Oxford: Clarendon.

Popper, K. R. and Eccles, J. C. 1977. *The Self and its Brain,* Berlin/Heidelberg/London/New York: Springer.

Putnam, H. 1975. *Mind, Language and Reality: Philosophical Papers, Volume 2*, Cambridge: Cambridge University Press.

Putnam, H. 1978. *Meaning and the Moral Sciences*, London: Routledge and Kegan Paul.

Putnam, H. 1983. *Realism and Reason: Philosophical Papers, Volume 3*, Cambridge: Cambridge University Press.

Putnam, H. 1988. *Representation and Reality*, Cambridge, Mass: MIT Press.

Quine, W. V. O. 1960. *Word and Object*, Cambridge Mass: MIT Press.

Quine, W. V. O. 1979. 'Facts of the Matter', *Southwestern Journal of Philosophy*, 9, pp. 155-69.

Reynolds, C. 1987. 'Flocks, Herds, and Schools: A Distributed Behavioral Model', *Computer Graphics*, 21, p25.

Robinson, H. (ed.) 1993. *Objections to Physicalism*, Oxford: Clarendon.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*, Princeton: Princeton University Press.

Rorty, R. 1980. *Philosophy and the Mirror of Nature*, Oxford: Blackwell.

Rueger, A. and Sharp, W. D. 1996. 'Simple Theories of a Messy World: Truth and Explanatory Power in Nonlinear Dynamics', *British Journal for the Philosophy of Science*, 47, pp. 93-112.

Russell, B. 1903. *The Principles of Mathematics*, Cambridge: Cambridge University Press.

Ryan, A. 1970. *The Philosophy of John Stuart Mill*, London: Macmillan.

Savellos, E. E. and Yalçin, Ü. D. (eds.) 1995. *Supervenience: New Essays*, Cambridge: Cambridge University Press.

Schlick, M. 1953. 'Philosophy of Organic Life', in H. Feigl and M. Brodbeck (eds.) *Readings in the Philosophy of Science*, New York: Appleton-Century-Crofts, Inc., pp. 523-536.

Schröder, J. 1998. 'Emergence: Non-Deducibility or Downwards Causation?', *The Philosophical Quarterly*, 48, pp. 433-452.

Schwabe, W. W. 1979. *Control of Plant Development*, Monograph 3, British Plant Growth Regulator Group.

Schwabe, W. W. and Clewer, A. G. 1984. 'Phyllotaxis – A Simple Computer Model Based on the Theory of a Polarly-Translocated Inhibitor', *Journal of Theoretical Biology*, 109, pp. 595-619.

Scott, A. 1999. *Nonlinear Science: Emergence and Dynamics of Coherent Structures*, Oxford: Oxford University Press.

Sears, F. W., Young, H. D., and Zemansky, M. W. 1987. *University Physics (seventh edition)*, New York: Addison Wesley Longman.

Serra, R. and Zanarini, G. 1990. *Complex Systems and Cognitive Processes*, Berlin: Spinger-Verlag.

Shamos, M. H. (editor) 1987 [1959]. *Great Experiments in Physics: Firsthand Accounts from Galileo to Einstein*, New York: Dover.

Silberstein, M. and McGeever, J. 1999. 'The Search for Ontological Emergence', *Philosophical Quarterly*, 49, 195, 00. 182-200.

Singer, C. 1959. *A History of Biology*, (revised edition), London: Abelard-Schuman.

Smart, J. J. C. 1989. *Our Place in the Universe: A Metaphysical Discussion*, Oxford: Basil Blackwell.

Smart, J.J.C. 1978. 'The Content of Physicalism', *Philosophical Quarterly,* 28, pp. 339–341.

Smith, A. D. 1993. 'Non-reductive physicalism', in Robinson (1993), pp. 225- 250.

Snowdon, P. 1989. 'On Formulating Materialism and Dualism', in Heil, J. (ed.) *Cause, Mind and Reality: Essays in Honour of C.B.Martin* (Dordrecht: Kluwer Press) pp. 137–158.

Sokal, A. and Bricmont, J. 1998. *Intellectual Impostures*, London: Profile Books.

Soucek, B. and the IRIS Group. 1992. *Dynamic, Genetic, and Chaotic Programming: The Sixth Generation*. New York: John Wiley & Sons. [BL there should be an accent (like this: _)above the 'c' - can't get it on this font, sort out for final version, and check references in Chapter 3.]

Sperry, R. W. 1966. 'Brain bisection and mechanisms of consciousness', In J. C. Eccles, ed., *Brain and Conscious Experience*, New York: Springer, pp. 298-313.

Sperry, R. W. 1976. 'Changing concepts of consciousness and free will', *Perspect. Biol. Med.* 20: 9-19.

Sperry, R. W. 1978. 'Mentalist monism: consciousness as a causal emergent of brain processes', *Behav, Brain Sci.* 3: 367.

Sperry, R. W. 1983. *Science and Moral Priority*, Oxford: Basil Blackwell.

Sperry, R. W. 1986. 'Discussion: Macro- Versus Micro-Determinism', *Philosophy of Science* 53, 265-270.

Spurrett, D. 1993. *A Philosophical Enquiry into the Freedom of the Will, and the Defensibility of the Doctrine of the Compatibility of Freedom and Determinism*, MA Thesis in Philosophy at the University of Natal, Durban.

Spurrett, D. 1997. 'Beyond Determinism', *South African Journal of Philosophy*, Volume 16 number 1, 14-22.

Spurrett, D. 1999a. 'Lyotard and the Postmodern Misunderstanding of Physics', *Theoria*, 93, Science and Civilisation, pp. 29-52.

Spurrett, D. 1999b. Review Article of Cilliers, P. (1998) in *South African Journal of Philosophy*, Volume 18, number 2, pp. 258-274.

Spurrett, D. 1999c. 'Fundamental Laws and the Completeness of Physics', *International Studies in the Philosophy of Science*, Vol. 13, No.3, pp. 261-274.

Spurrett, D. and Papineau, D. 1999. 'A Note on the Completeness of 'Physics'', *Analysis*, 59, 1, pp. 25-29.

Stephan, A. 1992. 'Emergence – A Systematic View on its Historical Facets', in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 25-48.

Stroud, B. 1987. 'The Physical World', *Proceedings of the Aristotelean Society*, LXXVII (1986/7), 263-277.

Swammerdam, J. 1758. *Book of Nature* (translated by T. Flloyd). London.

Swoyer, C. 1996. 'Theories of Properties: From Plenitude to Paucity', *Philosophical Perspectives*, 10, pp. 243-264.

Taylor, C. 1985. *Human Agency and Language: Philosophical Papers 1*, Cambridge: Cambridge University Press.

Teller, P. 1986. 'Relational Holism and Quantum Mechanics', *British Journal for the Philosophy of Science*, 37, pp. 81-81.

Teller, P. 1992. 'A Contemporary Look at Emergence', in A. Beckermann, H. Flohr, and J. Kim (eds.), *Emergence or Reduction?* Berlin: de Gruyter, pp. 139-153.

Thomson, J. J. 1977. *Acts and Other Events*, Ithaca: Cornell.

Tye, M. 1995. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, Mass., MIT Press.

Van der Waerden, B. L. (ed.) 1968. *Sources of Quantum Mechanics*, New York: Dover.

Van Fraassen, B. 1980. *The Scientific Image*, Oxford: Clarendon.

Von Neumann, J. 1966. *Theory of Self-Reproducing Automata*, edited and completed by A.W. Burks, Urbana: University of Illinois Press.

Wagner, S. J. 1993. 'Truth, Physicalism, and Ultimate Theory', in Robinson (1993), pp. 127-158.

Walker, R. 1993. 'Transcendental Arguments against Physicalism', in Robinson (1993), pp. 61-80.

Westfall, R. 1977. *The Construction of Modern Science: Mechanisms and Mechanics*, Cambridge: Cambridge University Press.

Westfall, R. 1980. *Never at Rest: A Biography of Isaac Newton*, Cambridge: Cambridge University Press.

Wheeler, J. A. and Zurek, W. H. (eds.) 1983. *Quantum Theory and Measurement*, Princeton: Princeton University Press.

Whitehead, A. N. 1978. *Process and Reality: An Essay on Cosmology (Gifford Lectures Delivered in the University of Edinburgh During the Session 1927-28)*, corrected edition, edited by D. R. Griffin and D. W. Sherburne, New York: Free Press.

Wigner, E. P. 1961. 'Remarks on the mind-body question'. In *The Scientist Speculates*, I. J. Good, (ed.) pp.

284-302. New York: Basic Books. (Reprinted in Wheeler and Zurek, 1983.)

Wilson, R. A. 1996. 'Promiscuous Realism', *British Journal for the Philosophy of Science*, 47, pp. 303-316.

Wittgenstein, L. 1958. *Philosophical Investigations*, (translated by G E M Anscombe), Oxford: Blackwell.

Wolfram, S. 1983. 'Statistical Mechanics of Cellular Automata'. *Rev. Mod. Phys.* (55) 601-644.

Wolfram, S. 1984. 'Universality and complexity in cellular automata'. *Physica D.* (10) 1-35.