ABSTRACT

OF

One World and the Many Sciences: A Defence of Physicalism

by Andrew Melnyk (Corpus Christi College, Oxford)

(Submitted for the degree of D.Phil. in Philosophy, Trinity 1990)

The subject of this thesis is physicalism, understood not as some particular doctrine pertaining narrowly to the philosophy of mind, but rather as a quite general metaphysical claim to the effect that everything is, or is fundamentally, physical. Thus physicalism explicates the thought that in some sense physics is the basic science. The aim of the thesis is to defend a particular brand of physicalism, which I call eliminative type physicalism. It claims, roughly, that every property is a physical property -- a property mentioned in the laws of physics, and hence that any putative property not identifiable with a physical property must be eliminated from our ontology.

Eliminative type physicalism is apt to face three objections, and so my thesis, like Caesar's Gaul, falls into three parts. In the first, I argue against the idea that there are tenable positions, both physicalist and non-physicalist, alternative to eliminative type physicalism. I argue that each of these positions -- token physicalism (Fodor, middle Putnam), supervenience physicalism (Lewis, Horgan) and and a non-physicalist view I call pluralism (Goodman, late Putnam) -- is defective. In the second part, responding to the objection that there is just no reason to be a physicalist, I develop a positive argument for eliminative type physicalism, an argument resting upon a strong version of the explanatory test for reality according to which only explanatorily indispensable properties can justifiably be said to exist. In the third and final part, I argue, against the charge that eliminative type physicalism cannot accommodate what I call phenomenal properties (qualia, raw feels etc.), that there is no good reason to deny, and one good reason to affirm, that phenomenal properties just are physical properties.

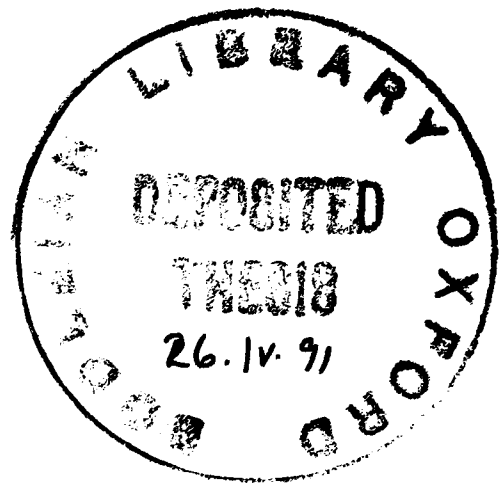
ONE WORLD AND THE MANY SCIENCES:

A DEFENCE OF PHYSICALISM

(Submitted for the degree of D.Phil. in Philosophy, Trinity 1990.)

Andrew Melnyk

Corpus Christi College, Oxford



Contents

Introduction 1

Part 1

One: What's Wrong With Token Physicalism 16

1. The Thesis Of Token Physicalism 17

2. The Nature of Explanation 20

3. The Argument From Overdetermination 63

4. An Unsound Motivation For Token Physicalism 72

5. The Epistemological Argument 76

6. The Generalised Davidsonian Argument 89

Two: Physicalism: From Supervenience To Elimination 96

1. Supervenience And Supervenience Physicalism 98

2. Supervenience Physicalism And Possible Worlds 105

3. The Grand-Property Hypothesis 112

4. The Best Explanation Of Claims Of Supervenience 119

5. Conclusion 123

Three: Problems For Pluralism 125

1. An Initial Problem For Pluralism 141

2. Two Pluralist Tendencies 143

3. Cookie Cutting Criticised 149

4. The Goodmanian Case Examined 155

5. Putnamian Internal Realism 173

6. Conclusion 192

Four: Eliminative Type Physicalism 193

1. "Physical" 194

2. The Fate Of The Eliminated Sciences 196

3. The Positive Argument Outlined 201

Part 2

Five: The Explanatory Test Defended: I 215

1. Scepticism About The External World 216

2. Inference To The Best Explanation 226

3. The Data For The Inference 244

Six: The Explanatory Test Defended: II 257

1. Davidson 259

2. Putnam 277

3. Constructionism 286

Seven: The Best Explanation: PHYSICS 307

1. PHYSICS As A Candidate Explanation 309

2. PHYSICS As The Best Explanation 312

Part 3

Eight: Eliminative Type Physicalism And Phenomenal Properties 318

1. Solution By Elimination? 320

2. Two Kripkean Arguments 325

3. The Knowledge Argument 334

4. Phenomenal Properties As Physical Properties 343

5. Phenomenal Properties And Multiple Realisability 350

6. Conclusion 353

Conclusion 355

Bibliography 361

INTRODUCTION

> The practical scientist does the business but the philosopher keeps the books. (Goodman 1972, p.168)

My central concern in this thesis is with the nature of reality and, specifically, with one view about it: the doctrine of physicalism. But the physicalism which interests me is not to be understood in some narrow way as pertaining especially to the philosophy of mind. Rather, my concern is with a quite general metaphysical claim, to the effect that everything is, or is at bottom, physical, from which someone might wish to deduce as a special case their physicalism in the philosophy of mind. Such a general metaphysical claim expresses the thought that in some sense physics is the basic science.

My reason for wanting to consider a general physicalism is that in my view the varieties of physicalism (and anti-physicalism), understood as general metaphysical claims, are best seen as responses to a completely general philosophical problem. This problem I shall refer to throughout the thesis as the problem of the many sciences, and it arises in the following way.

In describing and explaining the world, we make use of many, apparently distinct, systems of discourse. Each of the

sciences, such as physics, chemistry and sociology, constitutes such a system of discourse, but science is not the only source of these systems, since in common sense we can find a folk physics and a folk psychology, both of which qualify as systems of discourse in my sense. Let us decide to call these systems of discourse sciences. Now we usually take it, at least before we begin philosophy, that the properties mentioned in the many sciences genuinely exist, so that if we speak of gravity, or hydrogen, or revolutions, or tables, or beliefs, we are speaking of things that are real. Let us decide to call the properties mentioned in a science the domain of the science. The problem of the many sciences arises because, alongside our acknowledgement of the many sciences and their corresponding domains, we feel the force of a deep metaphysical intuition to the effect that there is only one world. The problem is that of supplying some plausible general account of the relations between the many sciences and between their corresponding domains: how do the many sciences and their domains fit together so as to constitute one world? If we follow Wilfrid Sellars in thinking that "The aim of philosophy ...is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term", then the problem of the many sciences must rate as the very model of a philosophical problem (see Sellars 1963, p.1).

It is important to see that a certain naive response to the problem of the many sciences is inadequate. According to it, the many sciences deal with different spatio-temporal parts of a single world, rather in the way in which the many Ordnance Survey maps of Great Britain deal with different parts of that country; and just as we could create a single map of Great Britain by arranging together a complete set of Ordnance Survey maps, so the many sciences could be stitched together to form a single, unified account of the one world.

But this naive response is inadequate, because the stitching together whose possibility it entails seems to be impossible. There simply do not appear in general to exist lawlike connections, such as one might on the naive view expect, between the domains of the many sciences; for instance, between biological and micro-physical things. (This is the phenomenon called by Jaegwon Kim nomological incommensurability (see, for example, Kim 1978).) And, indeed, there are not even always clear spatio-temporal relations between domains; for example, between mental events and neurological events. It is, in fact, due to a recognition of these truths that there has been in recent years a tendency for philosophers to be suspicious of physicalism, at least in its reductionist form (see, for instance, Fodor 1980).

The problem of the many sciences has a long history, and aspects of it have repeatedly been attended to by philosophers -- and non-philosophers -- of the past. Let me sketch a few historical examples of treatments of the problem.

Aristotle's metaphysics of matter and form, and his related doctrine of the four causes, can be interpreted as an effort to reconcile physico-chemical accounts of the world, such as the Pre-Socratics sketched, with biological and zoological accounts of living things, such as Aristotle himself pioneered; the idea being that physico-chemical accounts deal with matter, and biological or zoological ones with form. And it may be correct to suppose, with Martha Nussbaum, that in attempting this reconciliation Aristotle was responding to the proto-physicalism of Democritus (see Nussbaum, 1978).

Much later, in the seventeenth century, philosophers and scientists alike -- men such as Galileo and Descartes, Boyle and Locke -- were troubled by the question whether commonsense ascriptions to external objects of the so-called secondary qualities of colour, taste, smell and so on should be revised in the light of the developing physical science of the time (see, for instance, Hacker 1987, Ch. 1).

In the nineteenth century, a question of great concern to

ordinary men and women who were not philosophers was that of whether it is possible to reconcile a traditional Judaeo-Christian account of the origin of the world and of animal species in terms of divine creation with the emerging contemporary accounts of that origin in terms of evolution. I need hardly add that this question is disputed to this day.

In our own century, two particular disputes stand out as exemplifying discussion of the problem of the many sciences. The first concerns whether the picture of the world embodied in our commonsense beliefs about it, the picture according to which it contains such items as solid, brown desks, is compatible with the picture of it supplied by modern physics, according to which it contains swarms of microparticles interacting in the incredibly complex ways described by microphysics (see, for instance, Eddington 1929, pp.ix-xvii; Stebbing 1944, pp.40-55; and Mellor 1969). The second dispute has been over the traditional mind-body problem, in which, as I prefer to see it, we attempt to come to an understanding of the relations between, on the one hand, talk of persons' minds and, on the other, talk of their bodies against the background assumption that persons are in some sense unified things.

What these assorted intellectual disputes have in common, I believe, is that they are all concerned with the question of

whether, and if so, how, one system of discourse about the world can be made to fit together with another system of discourse about the world. As such, these disputes are all aspects of what I have called the problem of the many sciences. But although the problem is an old one, it is a particularly pressing problem today. The reasons are, first, that in this century we have so many more sciences than were known even to Locke -- today we have, for example, meteorology and neurophysiology; and, secondly, that the sciences, especially physics, are so much better-developed than they were in any previous century.

My aim in this thesis is to confront this ancient problem of the many sciences in its full generality, and hence at a fairly high level of abstraction. This abstract approach doubtless has its costs, but if its results are found interesting, then it will have been sufficiently vindicated. This thesis has been written in the conviction that there <u>are</u> interesting things to say at this abstract level, even if there are also interesting things to say at levels of greater particularity (see, for instance, the survey in Hooker 1981).

My claim a few paragraphs back was that the varieties of physicalism (and anti-physicalism) are best seen as responses to the problem of the many sciences. These responses may be

classified, at a high level of abstraction, as either monistic or pluralistic, according to whether they retain or reject the metaphysical intuition to the effect that there is only one world. Monistic theories attempt to maintain the unity of the world by abandoning descriptive egalitarianism and instead elevating one of the many sciences to an elite position, so that one of the sciences, with its domain, is treated as in some sense basic. The science which is almost bound to be selected for this elevation is, of course, physics. Non-physicalist monisms are conceivable, but I know of none, except idealism, which has been defended, and I regard them as a possibility which may safely be ignored. So the only monistic theories which I shall consider are brands of physicalism, each of which, in its different way, assigns to physics and the physical a certain metaphysical primacy. Pluralistic theories, by contrast, break with the intuition that there is only one world, and instead countenance a genuine plurality of worlds, each of which can serve as the domain for one science. Since in fact I know of only one pluralistic theory, I shall call it pluralism. It is, of course, a form of anti-physicalism.

So my concern is with physicalism. But it is no neutral concern: I come to praise physicalism, not to bury it. And the

brand of physicalism that I come to praise is perhaps physicalism in its least attractive form. It is, in effect, an updated form of Democritan atomism. I call it eliminative type physicalism. It is a form of type physicalism, because it claims that every type (or property) is a physical type (or property), where a type (or property) is physical iff it is mentioned in the laws of physics. (This account of the physical is a crude approximation; it is refined, when this becomes necessary, at later points, and especially in Chapter Four.) And it is eliminative (as opposed to reductionist), because the proposal is to eliminate -- that is, to cease to believe in the reality of -- any type (or property) which cannot be type-identified with, and hence reduced to, some physical type (or property). Since such type-identification is, I take it, impossible to achieve with most of the types (or properties) mentioned in the non-physical sciences (henceforth, the special sciences), eliminative type physicalism amounts, in effect, to a proposal to eliminate the special sciences. The special sciences can perhaps live on, subject to the condition that they are interpreted in some non-realistic way, for example, instrumentally; but we will not be able to accept as real their corresponding domains (on this see further Chapter Four, Section Two).

No one is more keenly aware than I am of the radical nature of eliminative type physicalism. I must appear, indeed, to be some kind of bizarre intellectual vandal, hacking whole branches off the tree of knowledge. But the notion of elimination at the heart of the doctrine is a familiar one. Let me illustrate it by reference to a central issue in metaethics. Suppose, as some philosophers would insist, that we allow our practice of moral assessment to count as a science, on the grounds that it is a system of discourse we employ in describing the world. Then a respectable number of philosophers, those who are not moral realists, will for a variety of reasons refuse to believe in the properties mentioned in those descriptions, that is, in moral properties. Accordingly, they will hold either that our practice of moral assessment is systematically erroneous, since it purports to ascribe to the world non-existent properties, or that it must be interpreted in some non-realist fashion, perhaps prescriptively. But, expressed in my terms, these opponents of moral realism can be said to have proposed the elimination of our practice of moral assessment. Most philosophers, I should say, propose the elimination of some kind of discourse; I happen to propose the elimination of many such kinds.

Let me now indicate the overall organisation of what is to

follow. Eliminative type physicalism is apt to face three kinds of objection, and my defence of it falls, correspondingly, into three parts. In Part 1, which comprises Chapters One to Four, I argue against the objection that there are alternative positions available, both physicalist and anti-physicalist, which attain the same level of generality as eliminative type physicalism, and hence which constitute responses to the problem of the many sciences, but which may be adopted instead it. My aim in Part 1, of course, is to knock out positions alternative to my own, so that mine alone is left standing. So I examine each such alternative position known to me, and try to show that it is seriously defective. Accordingly, Part 1 contains chapters criticising token physicalism (Chapter One), supervenience physicalism (Chapter Two) and anti-physicalist pluralism (Chapter Three). Argument for the crucial claim that I do in fact consider all the genuine alternative positions is supplied in the first few pages of Chapter Three. Part 1 ends (Chapter Four) with a clarification of the nature of eliminative type physicalism, and an outline of the positive argument for it to be developed throughout Part 2.

Part 2, which involves Chapters Five to Seven, is taken up with responding at length to the charge that there is simply no

positive reason to be an eliminative type physicalist. This charge might be made by one who favoured a possible response to the problem of the many sciences that I have not yet mentioned, a response neither physicalist nor pluralist, which we can all scepticism. The sceptic (in this sense) is impressed by the sheer multiplicity and complexity of the many sciences, and indeed of the theories within sciences, and judges that we just do not yet know, and perhaps never will, how all the sciences, and their parts, fit together. (There is a hint of such scepticism in Cartwright 1983, p.12 and Joseph 1980.) My response to this sceptical challenge that we have no reason to be a physicalist consists in the development and defence, in Part 2, of a positive argument in favour of eliminative type physicalism. The crucial philosophical claim I there defend is that we are justified in believing in external properties if, but only if, those properties are mentioned in the best explanation of the regularities we find in our sensory experience (Chapters Five and Six). Since, as I conjecture in Chapter Seven, it is physics which constitutes that explanation, it follows from the philosophical claim that we are justified in believing in a property if, but only if, it is a physical property; which conclusion is, roughly, eliminative type physicalism. My aim in Part 2 is thus to set on a somewhat firmer argumentative foundation the statements of

physicalist faith one sometimes encounters (see, for example, Field 1978, p.9), and, more particularly, to make more precise the idea expressed by some authors that physics is by itself explanatorily adequate (see, for example, Levin 1979a, p.87; Lewis 1983b, p.105).

In Part 3, which contains only Chapter Eight, I try to meet the anti-physicalist challenge of those who feel that there are undeniable aspects of reality -- namely those constituting the so-called subjective character of experience -- which the austere ontology of eliminative type physicalism cannot accommodate. My response is to argue that there is no good reason for denying, and one good reason for asserting, that, in a sense to be explained, phenomenal properties (as I shall call them) <u>are</u> physical properties.

Naturally, the topic of this thesis -- the problem of the many sciences and the various possible responses to it -- is one which I regard as extremely interesting. But I should like to indicate briefly why I hope the thesis may be held to be of interest to philosophers who are <u>not</u> especially interested in the varieties of physicalism and anti-physicalism.

The reason is the bearing it has on a number of other issues of wider interest to philosophy. For instance, my argument

throughout Part 2, which links explanation to justification, aims to supply a philosophical motivation for the explanatory test for reality discussed in full generality, but apparently rejected, by David Wiggins (see Wiggins 1976, p.362-3). Something very like this explanatory test is used by Gilbert Harman in his criticism of moral realism (Harman 1977, pp.6-9), and by John Mackie in his case against the reality of colours (Mackie 1976, pp.17-21). My argument similarly helps to motivate the causal account of knowledge assumed by Paul Benacerraf in his well-known discussion of mathematical truth (Benacerraf 1983, pp.412-415).

Another issue which I hope I may have illuminated is, as I have already hinted, that of whether common sense and scientific accounts of the world are or are not compatible with one another, my answer being, of course, that they are not. Again, Part 2 involves a fairly full response, based on inference to the best explanation, to the traditional problem of scepticism about the external world, a response which leads to a defence of realism about so-called unobservables. Finally, since the argument for my favoured brand of physicalism is eliminative in character, my exposition of it is relevant to related, though importantly different, applications of eliminative arguments in contemporary philosophy of mind

(see, for example, Churchland 1981).

Before closing this Introduction, one final matter needs attention. There is an objection to eliminative type physicalism -- I shall call it, unsurprisingly, the obvious objection -- which I have not so far mentioned but which deserves a response right at the start. According to the obvious objection, eliminative type physicalism must be false, because it straighforwardly entails the non-existence of such things as chairs and elephants which obviously do exist. My response to the objection falls into two parts.

First, I deny the obviousness, in a philosophical context, of the existence of chairs and elephants. The issue of physicalism is fundamentally an ontological issue -- it concerns what kinds of things exist -- and the issue is trivialised if we are allowed simply to resort to our intuitions about what exists. So some <u>account</u> must be given of how it is that we know, if we do, that chairs and elephants exist. But if such an account is offered, then it must, I believe, confront squarely the traditional challenge posed by the sceptic about the external world who denies that we know that any external objects at all exist. But now, as I argue in Part 2, the only satisfactory response to this traditional scepticism plays straight into my hands, by supplying the key

premiss in a positive argument for eliminative type physicalism. There is an important sense, which will become clear in Part 2, in which, although eliminative type physicalism seems so counter-intuitive, it nonetheless saves the appearances.

Secondly, no matter on what grounds the obvious objector insists upon the existence of chairs and elephants, we need to know to what kind of general metaphysical view, adequate to constitute a response to the problem of the many sciences, this insistence leads. For I feel confident that when the obvious objection is developed into such a view, the view will turn out to be one of those I discuss and criticise in Part 1. So the obvious objector might, for instance, be forced by the extravagance of his ontological intuitions to embrace pluralism; and in that case he would need to face the problems for that view which I raise in Chapter 3. The general point here is that, if my arguments in Part 1 are correct, then no view, of a generality comparable to that of eliminative type physicalism, is philosophically costless. So it will not do to reject eliminative type physicalism merely on the grounds that it comes at a price.

Part 1

CHAPTER ONE

What's Wrong with Token Physicalism

In this chapter I shall begin the task of trying to show that all responses to the problem of the many sciences other than the eliminative type physicalist response that I favour are prone to serious objections. Specifically, I shall in this chapter be discussing and criticising the doctrine of token physicalism.

Let me indicate the overall plan of the chapter. In Section One, I describe the thesis of token physicalism. Then, in Section Two, I outline and defend some views about the nature of explanation on which some of my subsequent arguments depend. In Section Three, I present my first major argument against token physicalism, which I call the Argument From Overdetermination. The aim of this argument is to show that, if token physicalism is true, then the world is massively overdetermined. Section Four seeks to undermine a commonly-encountered motivation for token physicalism. In Section Five, I present my second major argument against token physicalism, which I call the Epistemological Argument. The aim

of this argument is to show that, embarrassingly, the token physicalist can find no way in which he could come justifiably to believe the claims of token (or event) identity to which he is committed, given the rest of what he believes. (Thus this argument makes no challenge to the very idea of event identities.) Finally, in Section Six, I consider and then reject an argument for token physicalism based upon Donald Davidson's celebrated argument for Anomalous Monism.

1. The Thesis of Token Physicalism

What is the thesis of token physicalism, to which I am opposed? The thesis of token physicalism has its origin in discussions about the traditional mind-body problem (see Fodor 1980 and Putnam 1980). But it is in fact a quite general account of the relations between the many sciences, and is not peculiarly or exclusively concerned with the relation between (say) neurophysiology and folk psychology. The core of token physicalism is expressed by these three claims:

(a) Every event, whether psychological, sociological, meteorological or whatever, is identical with some physical event or other.

(b) Special scientific properties (types) cannot be identified with physical properties (types).

(c) But special scientific properties (types) do exist -- they

are genuine.

Accordingly, a token physicalist might say that the (folk psychological) property of believing that grass is green is a genuine property, even though it is not identical with any physical property, but that every instance of believing that grass is green -- every event of this folk psychological type -- is identical with an instance of some physical property, is identical, that is, with some physical event. Similarly, a token physicalist might say that the (economic) property of being a money is a genuine property, even though it is not identical with any physical property, but that every economic event of a monetary type is identical with some (presumably complex) event of a physical type.

Token physicalists are apt to speak of levels of description, and of explanation. Thus, suppose that I raise my hand at an auction; then this event may have a description and an explanation at many levels -- as an economic event (a bid) at the economic level, as a physiological event (a muscle-contraction) at the physiological level, as a chemical event (of a certain complex type) at the chemical level, and as a microphysical event (of a certain very complex type) at the microphysical level. These levels seem often to be thought of as constituting some sort of hierarchy. The most

fundamental level, at least as far as current knowledge reveals, is that of microphysics, but there are also chemical, physiological, psychological and sociological levels (among others). At each level, the appropriate science will mention autonomous properties which cannot (usually) be identified with properties mentioned in any lower level science. But it remains true that the level of physics is basic, for every special scientific, i.e. every higher-level, event is identical with some physical event. No other science than physics enjoys this privileged status -- which is why token physicalism qualifies as a version of physicalism, i.e. as a view of the world which assigns <u>some</u> kind of primacy to physics.

It is clear from this how it is that token physicalism constitutes a response to the problem of the many sciences: although the many sciences mention mutually irreducible properties, they are still united, according to token physicalism, inasmuch as they describe and explain the very same events (albeit at different levels).

Finally, how is token physicalism related to functionalism? As it happens, token physicalists are usually functionalists, as Jerry Fodor and Hilary Putnam both were at the time when they wrote the papers mentioned above. But functionalism does

not entail token physicalism. One reason for this is that functionalism, with its emphasis on multiple realisability, does not, strictly speaking, entail any ontological doctrine. But another reason is that a functionalist could forswear an ontology of events and adopt supervenience physicalism instead of token physicalism (see Chapter Two). So functionalism is distinct from token physicalism.

2. The Nature of Explanation

The goal of this section is to expound and, to a limited extent, to defend my views concerning the nature of explanation, especially scientific explanation. The defence of my views will be more detailed in some places than in others -- more detailed in the places where my views are more controversial -- and this simply reflects considerations of space: a full defence of my views would no doubt require a thesis-length treatment in its own right. I hope, however, that where my defence is thinner, the authors whom I cite will be consulted by the interested reader.

The point of an exposition and defence of my views concerning the nature of explanation is two-fold. First, and more narrowly, the argument against token physicalism which I present in the next section -- what I call the argument from

overdetermination -- relies upon the truth of certain aspects of these views; so I want to prepare the ground for that argument now. But secondly, and more broadly, many of the arguments and considerations adduced in the remainder of this thesis are informed to some degree by my views about explanation; so it will, I hope, be useful if I lay my cards on the table right now by clearly setting out those views.

(i) The fundamental conception of explanation that I favour is taken from David Lewis (Lewis 1986a) and is this: to explain an event -- to explain why it occurred -- is to give information about its causal ancestry. An event is conceived of as the product of a causal history, a history which may stretch back in time as far as you like. This history will typically be very rich, so that it will usually be impossible, or at least humanly impossible, to give information about the whole of an event's causal history. But one need not do this in order to explain an event: one need only give some information about the causal history of the event.

How might one do this? One might do it in a variety of ways. One might simply cite one or more of the causes of the event; or one might cite one or more what are sometimes referred to as the background conditions against which a cause operates; or one might cite some law in virtue of which a

familiar event was a cause of the event to be explained; or one might point out that a certain kind of cause did not on this occasion operate; or one might do it in other ways (see Lewis 1986a).

Two points about this conception of explanation may be noted at once. First, this conception may be described as one according to which explanation is causal explanation -- but only if "causal explanation" is not understood so narrowly as to include only those explanations that explicitly involve the citation of causes. For the Lewisian conception, as I have indicated, allows more than the explicit citation of causes to be explanatory. Secondly, this conception of explanation implies that it is possible to explain an event without necessarily having to cite some causal law under which the event falls; for instance, we can explain an event just by citing one of its causes. Thus this conception is not liable to the objection, made by Nancy Cartwright against the deductive-nomological account of explanation, that we seem perfectly able to explain events even when no appropriate law is to hand (see Cartwright 1983, pp.44-53).

I shall now present some examples meant to illustrate this conception of explanation. (In presenting them, I am not, of course, committed to accepting them as correct.) I include a

couple of everyday, non-scientific examples so that even readers who lack my considerable sympathy for science can see the intuitive appeal of the conception. The fact that the Lewisian conception of explanation gives a satisfying account of examples such as these constitutes important support for the correctness of the conception. Support for it also comes from the fact that many examples which prove difficult for the traditional deductive-nomological conception of explanation to cope with are easily handled by the Lewisian conception (or, more precisely, by any conception which invokes causation in the way Lewis' view does). This is nicely shown by Paul Humphreys (see Humphreys in Kitcher and Salmon 1989, pp. 300-301).

1. "Why is there an acrid smell in the kitchen? Because at the back of a cupboard some onions are rotting." In this example, one of the causes in the causal chain which culminated in the presence of an acrid odour is cited. The odour was generated by the chemical processes involved in the onions' rotting, and was then transmitted to the whole kitchen area by ordinary currents of air.

2. "Why does Barringer Crater in Arizona exhibit shock-metamorphic effects, such as rocks shattered in a conical and striated pattern, and the presence of stishovite? Because

of the impact of an iron meteorite with a mass of about one million tonnes -- an impact producing enough energy to shatter rocks in the distinctive way, and to turn quartz into stishovite." Here, the explanation cites an earlier cause and alludes to the lawful consequences of such an event.

3. "What was responsible for my dizziness then after using my bronchial inhaler? Because in some 5% of the population albuterol, which your inhaler delivers, causes dizziness." Here, the explanation involves redescribing an event (that of my using my inhaler) and then citing a law in virtue of falling under which the event caused the effect to be explained.

4. "Why has this paint gone hard? Because you left the lid off it after you last used it." The explanation here gives information about something which did not happen in the causal history of the event to be explained: the sealing up of the paint. Because the paint was not sealed up, the air got to it, and so it went hard.

5. "Suppose that a neutron belonging to an atom of a radioactive isotope has decayed, producing an electron and a proton. Why did the proton and the electron move off at an angle to one another as they did? Because the neutron decay produced a third particle, a neutrino, as well, and the

neutrino is responsible for the observed paths of the electron and proton." Here, the explanation merely cites an earlier cause of the event to be explained.

6. "Suppose I have shone some ultra-violet light onto a negatively-charged piece of metal, and that the metal has lost its charge. Why did the metal lose its charge? Because the photons making up the UV light are highly energetic, and when they strike the metal they dislodge electrons from it, thereby causing it to lose its charge."

To explain an event, then, is to give information about its causal history. But what is an explanation? The word, "explanation", has, I think, at least these two uses. In its first use, "explanation" refers to a linguistic item -- a set of statements, or propositions, which describe some aspect of the causal history of an event to be explained. When "explanation" is used in this way, explanation is clearly distinct from causation: explanations are linguistic things, whereas causation is a real relation in the world. But in its second use, "explanation" refers to actual events and processes in the world, those events and processes which, on the conception I support, constitute or partly constitute the causal history of the event to be explained. We are using "explanation" in this second way when we speak of discovering

or seeking the explanation for some event; or when we say that the disaster must have an explanation, even if it currently eludes us; or when we ask, "What was the explanation for the earthquake?" and say in reply, "An unusually high level of activity in the San Andreas fault (was the explanation)". In these cases, we use "explanation" to refer to events and processes in the world. When we use "explanation" in this, the second way, explanation and causation (or at least causal processes) are scarcely distinguishable. When these two uses of "explanation" are borne in mind, I think that it becomes much clearer how exactly I am understanding the relation between explanation and causation.

The Lewisian conception of explanation that I am defending is not eccentric or peculiar. Other philosophers who conceive of explanation in fundamentally the same way, and who therefore connect explanation and causation in the same way, are Wesley Salmon (see Salmon 1984) and Paul Humphreys (see Humphreys 1989). Each of these authors is also keenly aware that the causation which they connect with explanation need not be deterministic causation, and Humphreys in particular tries to develop the notion of probabilistic causation. (Perhaps I should add that I cannot be assumed to accept the theories of causation, whether deterministic or probabilistic, endorsed by

these philosophers, despite my general sympathy with their views on explanation.)

(ii) The view of explanation that I favour is a version of what Jaegwon Kim calls explanatory realism (see Kim 1989b). That is to say, whether or not some claim really does constitute an explanation of why some event occurred depends upon the way the world is, causally; and the way the world is, causally, is in general a thoroughly objective and mind-independent affair. (See Chapter Three, Sections Four and Five for remarks in defence of the idea that the way the world is is objective and mind-independent.) So what explains what is a thoroughly objective and mind-independent affair.

The view of explanation that I am defending thus contrasts sharply with any view according to which explanation has essentially to do with the removal of puzzlement or wonder. Indeed, I do not believe that explanation does have essentially to do with the removal of puzzlement or wonder. For although explanations often do in fact remove puzzlement or wonder, by fitting events to be explained into the causal network of the world and hence making them more predictable, the removal of puzzlement or wonder is neither a necessary nor a sufficient condition of being an explanation. It is not a necessary condition, since an excellent explanation may fail to reduce

puzzlement or wonder if the audience simply does not understand it. Also, there are specific counterexamples such as the following. You come upon a deck of cards in your new hotel room, and they are arranged in their proper order, ace to king, by suit. What is the explanation for this? In fact, the room's previous occupant shuffled them thoroughly, and they just came out arranged in proper order. This is the explanation, but if anything it increases, not reduces, your puzzlement or wonder (see Lewis 1986b, p.133). The reduction of puzzlement or wonder is not a sufficient condition of being an explanation, since one may reduce the puzzlement or wonder of a gullible person by communicating truths which seem, intuitively, to be explanatorily irrelevant.

So the conception of explanation that I favour is a thoroughly realist one. Now, it might be thought that on such a strongly realist view of explanation there was no scope for attaching any importance to pragmatic factors in explanation. But that thought would be an error. On the conception of explanation that I favour, there is plenty of scope for the influence of pragmatic factors.

The fundamental reason for this is that, as already noted, to explain an event is to give some, but not necessarily all, of the information about the causal history of the event. Which

items of information you give may permissibly be -- and usually will be -- affected by pragmatic factors, such as the likely interest and relevance to your audience of the information. But although this is so, explanation remains realist and objective, since whether the causal information you cite is correct, and whether it truly is information about the causal history of the event to be explained, are both determined by the way the world is, causally, and not by anything else.

Here is an illustrative example. The family car is slowing down for no apparent reason. So I ask my wife, "Why is the car stopping?", and she replies, "Because the axle bearings are not perfectly friction-free." By this reply she has certainly given me information about the causal history of the car's stopping, so that her reply certainly constitutes an explanation. But, given the context, the information she has given me is still pragmatically inappropriate. For I knew, and she knew that I knew, that axle bearings are never perfectly friction-free. What I was requesting, given the context, was a piece of causal information that I did not already possess -- such as the negative information that we had run out of (causally necessary) petrol.

This example illustrates the way in which pragmatic considerations can affect explanation, or at least acts of

explaining, even on the Lewisian conception I am defending. But examples of this sort do not show that there is some special relevance of pragmatic considerations to explanation. For if to explain an event is to give information about the causal history of the event, then acts of explanation are acts of information-giving, and are as such liable to pragmatic considerations (see Lewis 1986a, pp. 226-228).

But although this is so, there are some devices commonly used in requesting explanatory information to which it is worth drawing explicit attention. These devices help the requester to indicate which aspect or aspects of the causal history of the event to be explained he would like to be informed about. Thus, these devices help to determine what will be (for instance) a relevant explanation of some event.

One such device is the use of the contrastive form of the request for explanation, i.e. the form of the request on this pattern: "Why p, rather than q?" (where "q" can perfectly well be "not-p"). Some philosophers think that all requests for explanation must be regarded as having this form, even if they do not appear to have it; but I shall not comment on that claim (though compare Lewis 1986a, pp.229-231). It is worth giving a brief account of contrastive requests for explanation simply because such requests do sometimes occur.

What I claim is happening when explanations are requested by using the constrastive form is this: the requester indicates that he does not want just any old information about the causal history of the event to be explained -- he wants those items of information which would not <u>also</u> apply to the event described by g. Here is an example. Suppose the bank-robber Willy Sutton is asked: "Why do you rob banks, rather than, for instance, department stores?" It would be pragmatically inappropriate for Sutton to reply: "Because I had a deprived childhood". This reply certainly gives information about the causal history of his bank-robbing, and certainly helps to explain it. But this reply is pragmatically inappropriate to the explanatory request as formulated, since this reply describes a causal factor -- the deprived childhood -- which would have operated even if Sutton's life of crime had taken a different turn and he had gone into shop-lifting rather than bank-robbing. A pragmatically appropriate bit of information about the causal history of Sutton's bank-robbing might be this: "Because I realised that banks had so much more that was worth stealing -- <u>that's</u> why I rob banks rather than department stores". Sutton's bank-robbing had a long and complex causal history, involving <u>both</u> a deprived childhood <u>and</u> a realisation that there are rich pickings for robbers in banks. But the

contrastive form of the request for explanation of the bank-robbing indicates that it is the latter aspect of the causal history about which information would be appropriate.

The contrastive form of requests for explanation is one device by which someone who seeks an explanation can indicate what would and what would not be a relevant and genuinely informative reply. Another such device is the use of emphasis. To continue the example of Willy Sutton: instead of asking him, "Why do you rob banks, rather than department stores?", we could have asked, "Why do you rob banks?", thus emphasising the word "banks". But I believe that the same analysis applies here too: the effect of the emphasis is to indicate which items of information about the causal history of the bank-robbing would be relevant and genuinely informative.

(iii) In the last sub-section I tried to sketch how pragmatic considerations may enter into the Lewisian conception of explanation that I favour. Thus, I allow that there may be a pragmatics of explanation. However, I deny that we should accept a pragmatic account of explanation. Such a pragmatic account of explanation would make explanation entirely a pragmatic affair, which would conflict with the conception I favor. In this sub-section, therefore, I shall present a reason for rejecting as unsatisfactory the best-known pragmatic

account of explanation, that owed to Bas Van Fraassen (see Van Fraassen 1980, pp. 134-153). My reason for rejecting Van Fraassen's account is taken from a paper by Philip Kitcher and Wesley Salmon (see Kitcher and Salmon 1987); so I shall simply summarise their criticism. For a full account, of course, their original paper must be consulted.

I begin by sketching Van Fraassen's pragmatic account of explanation. The root idea behind it is that an explanation is an answer to a why-question; Van Fraassen's account is an elaboration of this root idea. According to him, the standard form of a why-question is this:

Why (is it the case that) Pk?

where Pk describes the fact to be explained. This question may be identified with an ordered triple:

Q= <Pk, X, R>

where Pk is the topic of the question, X={P1, P2,..., Pk,...} is the contrast class, and R is the relevance relation. The contrast class contains, in addition to Pk, propositions stating possible but non-actual facts other than the one reported by Pk, so that the why-question asks why Pk, rather

than P1, or P2, or... .

How should such a why-question be answered? The canonical form of a direct answer to Q is this:

Pk in contrast to the rest of X because A.

This direct answer, if it is to be one, must satisfy the following four conditions:

(i) A is true (ii) Pk is true (iii) No member of X other than Pk is true (iv) A bears relation R -- the relevance relation --to <Pk, X>.

The difficulty for Van Fraassen's account which Kitcher and Salmon expose emerges even from this concise summary of the account, and concerns the relevance relation, R. The difficulty is that Van Fraassen lays no restriction whatever on what the relevance relation can be. Because of this, we can claim, on Van Fraassen's account, that any fact explains any other -- which seems absurd. We can claim this because we can always fabricate a relation between A -- the core of a direct answer to Q -- and the ordered pair <Pk, X>, thus satisfying condition (iv) above. We can always fabricate the relation between A and <Pk, X> because a relation is, from a formal point of view, a set of ordered pairs; so we can make relation R be the

following unit set of ordered pairs: {<A, <Pk, X>>}. If relation R can be any relation, then this is enough to ensure that A stands in relation R to <Pk, X>, and hence enough to ensure the satisfaction of condition (iv). Since we can easily arrange for conditions (i) to (iii) to be satisfied, we have to conclude that, on Van Fraassen's account, so long as A is true, A can figure in a direct answer to any why-question, "Why Pk?", i.e. figure in an explanation of Pk, whatever fact it reports. Thus, Van Fraassen's pragmatic account account of explanation can be trivialised. Nor can this result be avoided by appealing to Van Fraassen's account of how answers to why-questions should be assessed (see Kitcher and Salmon 1987, pp. 319-322).

Van Fraassen therefore needs to provide some restrictive account of what the relevance relation R may be: he must say that it is logical deducibility, or statistical relevance, or causation, or whatever. But to do this would involve his entering into the traditional disputes about the character of explanation (e.g. about the problem of asymmetries of explanation; see Kitcher and Salmon 1987, pp. 326-8), and whatever he settled on as the properly restricted relevance relation, or relations, it would no longer be true to say that his account was a pragmatic account of explanation, i.e. one

according to which explanation was entirely a pragmatic affair: it would be a non-pragmatic account of explanation which nevertheless included an account of the pragmatics of explanation.

(iv) So far I have set out a conception of what is involved in the explanation of events, events such as my dizziness on a particular occasion, the acrid odour in the kitchen this week, the car's coming to a standstill on a particular occasion, and so forth. But I need to show how this conception of the explanation of particular events can be extended so that it applies also to explanation of uniformities or regularities. This is an especially pressing need for someone who, like me, takes scientific explanation very seriously, since scientific explanations are very commonly explanations of uniformities or regularities.

While particular events are, I hold, the basic explananda, it is also legitimate to seek and to provide explanations for uniformities or regularities. But uniformities and regularities are explananda only in a derivative sense -- a sense I hope to make clear by developing some of Lewis' remarks (see Lewis 1986a, pp. 225-6).

Let me begin with the case of uniformities. An example of a

uniformity would be the fact that, in nearly all human males, the left testicle hangs lower than the right testicle. How, on the conception of explanation that I favour, is this uniformity to be explained? To answer this question, I need first to make a general point which another simple example will illustrate. Suppose that, within a single week, all ten of the 747 Jumbo jets manufactured by Boeing during 1988 crash; although this would strike us as indicating some systematic defect in the 747s manufactured in that year, it might just turn out to be the case -- as investigation could reveal -- that the cause of each crash was quite unconnected to the cause of any other: one crash was due to pilot error, a second due to sabotage, a third due to an engine fire, and so forth. If this was so, how would we explain the uniformity? The answer is that we would have to explain it simply by explaining each of the component events -- the individual crashes -- which made up the regularity. From this example I want to draw a general moral: it is at least possible to explain a uniformity simply by explaining each of the component events which make up the uniformity.

Let me now return to the question of how to explain the uniformity concerning human testicles. My suggestion is this: we should explain this regularity by explaining, or by giving a recipe for explaining, each instance which makes up the

uniformity, i.e. by explaining, or by giving a recipe for explaining, why each male who exhibits the relevant phenomenon has his left testicle hanging lower than his right. In other words, I am suggesting that the same approach which we would use in explaining the uniformity involving the 747 crashes should be applied also to the uniformity in male anatomy.

However, there is one crucial difference between the two cases, which I must take account of. This difference can be described intuitively by saying that the ten 747 crashes in a single week were just a coincidence, whereas the uniformity in male anatomy is, presumably, not. Because of this difference, it might be objected that whereas explaining each event that makes up the uniformity is acceptable in the 747 case, which is just a coincidence, it is not acceptable in the case of the male anatomy, where something more is required. Now, I accept that something more is required in the case of uniformities which we think of as non-coincidental, but I think the extra can be supplied -- without departing from the conception of explanation I favour -- as follows. When we explain a uniformity which is not just a coincidence, we must identify something which the causal histories of the events constituting the uniformity have in common. I do not mean that we must find that the events spring from a common cause (though that could

happen); rather, we must identify some type of causal factor such that a token of that type is operative in the causal history of each event making up the uniformity. This can be illustrated with the example of the uniformity in male anatomy, which should be explained like this. We explain, or, better, give a recipe for explaining, why each man who exhibits the relevant phenomenon does so by citing the influence of his genes in determining the specific form of his anatomy, including, of course, genital anatomy, and, because we do not have here a coincidence, we point out that the genes of all men who exhibit the relevant feature are alike in a key respect. And it is in virtue of citing something in common between the causal histories leading up to each event making up the uniformity that we may conclude that the uniformity is not a coincidence.

So far I have been trying to show that the Lewisian conception of explanation that I favour, which takes particular events as the basic explananda, can in a natural way be extended so as to yield an account of the explanation of uniformities. Let me turn now to the case of regularities. Can the Lewisian conception handle regularities also? I shall argue that indeed it can.

I shall say that there is a regularity just in case an event

of one type is often, or mostly, or always, followed (or accompanied) by an event of a second type. A regularity will often be lawlike, in the sense that it sustains counterfactuals of a certain kind: the regularity that Fs are always followed by Gs is lawlike if it is true that, if X had been F, it would have been G. What I say about regularities is meant to apply whether or not they are lawlike.

How, then, are regularities to be explained? The strategy of explanation, I suggest, should mimic that employed in explaining uniformities, but must be more complex. Here is how it should go. Suppose that we wish to explain the regularity that all Fs are followed by Gs. We begin by decomposing this regularity into its constituents, each of which is the fact that one particular F-event is followed by a particular G-event. What we do next will depend on what kind of case we have. In one kind of case, what we do next is to explain, or give a recipe for explaining, each particular G-event, and then notice that in the causal history of each G-event there is an F-event. So if we want to explain why I often have an asthma attack after drinking wine, we will explain each particular (wine-following) asthma attack -- as the result of an allergic reaction to sulfites -- and then point out that wine often contains sulfites. Here, we explain the regularity that Fs are

often followed by Gs by citing something common to the causal histories of all the Gs, namely, the causal influence of Fs.

In another kind of case, what we must do, once we have broken up the regularity, is to explain, or give a recipe for explaining, both the G-event and the F-event. Then -- assuming we do not have a coincidence -- we must identify some type of common cause a token of which, in each case of a G-event following an F-event, brings about both an F-event and a G-event. So, for example, we might want to explain why, in certain patients, Kaposi's sarcoma and Pneumocystis carinii pneumonia go together (see Humphreys 1989). As we now know, in each of these patients there is a common cause operating to produce both the types of effect, and that common cause is the presence of a certain group of retroviruses, those that cause AIDS. Citing this type of common cause, a token of which operates in the case of each instance which constitutes the regularity, serves to explain the regularity; and it does so, in line with the Lewisian conception of explanation, by providing information about the regularity's causal history.

(v) The Lewisian conception of explanation which I have been expounding and defending could be expressed compactly like this: explanation is causal explanation. Even though we are understanding "causal explanation" generously, so that it may

include as explanatory more than the mere citation of causes, this is a strong claim, since it implies that there is no such thing as non-causal explanation. Lewis, of course, is well aware of this, and defends the implication explicitly (see Lewis 1986a, pp. 221-224). He may, in the end, be right. But I suspect, in fact, that he is wrong: that is, I suspect that there are other types of explanation than causal explanation. Here are two examples. First, it seems possible to explain by using what Peter Forrest calls the Limited Possibility Model of explanation (see Forrest 1988b, p.96). On this model, we explain something by showing that, because of (e.g.) a conservation law, there were fewer possibilities alternative to the something than we previously supposed. Secondly, it may be possible to explain an empirical correlation between tokens of one type and tokens of a second type by affirming an <u>identity</u> between the types. So, for instance, if mental states of a given type are perfectly correlated with brain states of a certain type, we might explain this by supposing that the mental type is <u>identical</u> with the brain-state type (see Hill 1984, pp. 297-300).

So I think it is probably not the case that all explanation is causal explanation. Nevertheless, when I speak of explanation I will, unless explicitly indicated, be referring

to causal explanation only. My excuse for this is that even though I deny that all explanation is causal explanation, it still seems to me that a very great deal of explanation, including, of course, a very great deal of scientific explanation, is in fact causal explanation. I do not believe that any of the arguments in the rest of this thesis will be significantly impaired by my reliance on the admittedly approximate claim that all explanation is causal explanation.

I have a reason for claiming that a very great deal of explanation is causal explanation. My reason is simply that, aside from the two possible types of non-causal explanation noted above, each kind of putative non-causal explanation seems either not to be a kind of explanation at all or to turn out upon analysis, perhaps despite first impressions, to be an instance of causal explanation. I cannot substantiate this, as ideally I should, by showing how every supposed example of non-causal explanation is either non-explanatory or causal after all, but I can attempt to substantiate it for a few such examples. This I shall now do.

Take, first, common-sense explanation of behaviour by reference to common-sense mental states such as belief and desire. This has often seemed to philosophers not to be causal explanation. Some of the bad reasons for saying this are

criticised in a famous paper by Donald Davidson (see Davidson 1980, pp.3-19); and other reasons, which rely on the idea that common-sense psychological explanation is "rationalising" explanation, are well-criticised by Janet Levin (see Levin 1988; and compare my remarks on charity in Chapter Six, Section One). I have nothing to add to such negative criticisms. What I do want to suggest positively, however, is that an example devised by Christopher Peacocke (for a slightly different purpose) strongly encourages us to think that, in offering common-sense explanations of behaviour, we think of ourselves as citing (some of) the causes of that behaviour, and hence as offering causal explanations of that behaviour (see Peacocke 1983, pp. 204-6).

Here is (my version of) Peacocke's example. Suppose that we make the acquaintance of a creature with what appears to be a perfectly normal human body, which behaves in a perfectly normal human fashion; we will no doubt take this creature to be a normal person, and, consequently, part of our treatment of it will be the ascription to it, for the purpose of explaining its behaviour, of such propositional attitudes as beliefs and desires. So far, so good. But now let us suppose that we discover -- it does not matter why -- that this creature has no human brain, but instead contains a radio receiver which is

controlled by a cleverly-programmed computer on Mars. Signals picked up by the receiver cause the body of the creature to behave in the human-like ways we observe. Now, I think that upon making this discovery we would want to cease ascribing propositional attitudes to the creature, and withdraw those ascriptions made in the past. But why? The reason seems to be that what we discover about the creature is that the causation of its behaviour is quite different from what we imagined it to be, and that in particular the causation of its behavior does not involve beliefs and desires. But if this explanation of our withdrawal of attitude ascriptions is right, then we must have been taking ourselves, in ascribing beliefs and desires to the creature, to be citing (some of) the causes of its behaviour, i.e. to be offering causal explanations. And there seems no reason not to generalise from this example to conclude that, quite generally, common-sense psychological explanation is a species of causal explanation.

Let us turn now to another kind of explanation that might be thought to be non-causal, that which alludes to things' purposes, to what the things are for. In response to the question, "Why does the horseshoe crab have eyes?", it may seem appropriate to say, "In order to detect mates". Or, to give a sociological example, in response to the question, "Why does

the state exist?", it is sometimes replied, "To protect the interests of the ruling class". Are such response as these explanatory? And if so, are they causally explanatory, as I claim?

I believe that responses such as these are explanatory only if they are understood as causal explanations, that is, as giving information about causal histories. One way in which saying what a thing, or feature of a thing, is for can be explanatory is if the feature or the thing is an artefact, designed and built for a certain purpose; in that case, the reference to the purpose of the feature or thing can be understood as a highly compressed reference to the intentions of the agents who designed and created the artefact. This would make the explanation a common-sense belief/desire explanation and hence, if the argument of the last few paragraphs is right, a causal explanation. But this way of interpreting putative explanations that state the purposes of a thing will only apply, obviously, to artefacts, and since neither horseshoe crabs nor states (unless you are a conspiracy theorist) are artefacts, it will not apply to those cases. So can stating the purpose of a horseshoe crab's eyes, or the purpose of the state, be truly explanatory, given that neither the eyes nor the state are artefacts?

Perhaps they can. For it may be that we can understand such putative explanations as supplying information about the evolutionary history which led up to the crab's eyes or the state's existence. Consider first the horseshoe crab. Of any particular crab it is true to say that it has eyes because it is genetically programmed to have eyes. But if we trace back the history of the species, we will find that certain ancestors of the current generation of crabs were relatively advantaged by having eyes one of whose consequences was that the detection of mates was easier. That ancestors of the current generation of crabs had eyes, and that those eyes made the detection of mates easier, are aspects of the causal history of all current crabs. And this is so despite the fact that our knowledge of the detailed evolutionary development of the crab may be very sketchy. So if the statement that the crabs have eyes in order to detect mates is understood as alluding to these aspects of the causal history culminating in today's crabs, then the statement can be allowed to be explanatory; and of course it will be a Lewis-type causal explanation.

Consider now the state. On the assumption that something like evolutionary mechanisms, however scant our knowledge of them, operate in the social realm, a similar account to that just given may also be given to show that the response,"In

order to protect the interests of the ruling class", could be explanatory -- and in a causal way. The idea would have to be that if we traced back the (evolutionary) history of the currently existing state, we would find that its tendency to protect the interests of the ruling class contributed to its preservation and non-elimination, presumably in some pretty significant way. So if we take the response, "In order to protect the interests of the ruling class", as drawing our attention to this feature of the history of the state, then we can take this response also as a Lewis-type causal explanation. (My remarks here are influenced, and further supported, by Jon Elster's critique of G.A. Cohen's defence of functional explanation. See Cohen 1978 and 1982; and Elster 1980, 1982 and 1986.)

What the considerations of the last few paragraphs are supposed to suggest -- the considerations, I know, are far from conclusive -- is that commonsense belief/desire explanations, and teleological or functional explanations, if they are to be any good, must be understood, as they can be, as causal explanations, i.e. as attempts to give information about causal histories. This tends to support my claim that a very great deal of explanation is in fact causal explanation, understood along Lewisian lines (even if it is possible for

there to be non-causal explanation).

(vi) In this sub-section I wish to turn to a new topic, the connection between explanation and truth. It seems to me that explanations have to be true: a piece of information cannot be an explanation unless it is true. So if I explain your rash by saying that you touched some poison ivy, then it must be true that you touched the poison ivy; if I explain the shock-metamorphic effects observed in a given crater by saying that a large meteorite struck the Earth there some 150 million years ago, then it must be true that a large meteorite struck the Earth there some 150 million years ago. You cannot explain something by citing an untruth: if some proffered information is not true, then that is sufficient reason for denying that the information is genuinely explanatory. And if you offer some information with the intention of its constituting an explanation, then you must regard the information as true, i.e. believe (as opposed to accept) it. (I discuss the relation between explanation and truth further in Chapter Five (Section Two).)

Furthermore, given that explanations must be true, it follows that the entities and properties mentioned in explanations must truly exist; for the entities and properties are precisely those to which one is ontologically committed in accepting the

explanation as true. So it makes no sense for me to claim that the explanation for your rash was touching the poison ivy but then deny (or shrink from affirming) that there was any poison ivy, or that poison ivy exists; it makes no sense for me to claim that the explanation for the shock-metamorphic effects was the impact of a meteorite but then deny (or shrink from affirming) that such a meteorite existed, or that any meteorites exist. You cannot explain something by talking about non-entities: if an entity postulated in a putative explanation turns out not to exist, then that is sufficient reason for rejecting the putative explanation as not genuinely explanatory. And if you offer a putative explanation in which an entity is postulated, then you must suppose the entity to exist. (Cf. Cartwright 1983, pp.87-99; and Grimes 1987, pp.88-89.) The strongly realist way in which I understand entities to exist is clarified and defended in Chapter Three (Sections Three, Four and Five).

The claim that it is a conceptual requirement on explanations that they be true is often endorsed (see, for example, Salmon 1984, pp.107-8); but it is contested by Van Fraassen (see his 1980, pp.97-101). He contests the claim on the grounds that it seems perfectly acceptable to say, for instance, that Newton's theory explained the tides, even while allowing that Newton's

theory is not (and never has been) correct. But I doubt that, strictly speaking, this is a perfectly acceptable thing to say. If we really believe that Newton's theory is false, then we will realise that it does not, because it cannot, explain anything; a fortiori, it does not, and never did, explain the tides. What is true is that Newton's theory purported to explain the tides: it was such that, had it been true, it would have explained the tides. Newton's theory was, if you like, a candidate or potential explanation. But if it is false, it cannot have been an actual explanation.

I suspect that "explanation" is often elliptical for "candidate explanation", and that "explains" is often elliptical for "would explain, if true". These ellipses are harmless enough in ordinary contexts, because the more pedantic expansions of them spring naturally to mind. But that they are ellipses should not be permitted to blind us to what the consideration of any number of examples makes clear: if a purported explanation turns out not to be true, then that is sufficient reason for rejecting the purported explanation as not constituting a genuine explanation.

Another possible objection to my claim that explanations, if they are to be such, must be true might be made by someone sympathetic to the notion of verisimilitude, or approximation

to the truth. Such a critic might object that an explanation need not be true simpliciter, but only true enough, or sufficiently true. But even if we leave aside the controversy surrounding the very idea of verisimilitude, this objection may have little force. For I would suggest that if a theory is to be sufficiently true to be genuinely explanatory, then it is sufficiently true for acceptance of it to commit one to belief in the reality of the entities postulated by it (even if it is denied that our description of those entities is perfectly correct). And it is belief in the entities postulated which is what I really want from acceptance of an explanation. (This, by the way, is why Nancy Cartwright's view that theories do not have to be true to be explanatory is no threat to my position. For she also holds that causal explanations, if accepted, do commit one to belief in the reality of the entities postulated. See her 1983, pp. 87-99.)

(vii) Finally, I want in this last sub-section to defend a claim upon which I shall be relying in my next section where I expound my first argument against token physicalism, the argument from overdetermination. The claim to be defended here and now is this: explanatory contexts are extensional -- permit the truth-preserving substitution of co-referring singular terms -- in respect of the explanandum place. So if "e" refers

to an event (an explanandum), and if "f" refers to an event (an explanandum), and if e=f, then if e has a so-and-so explanation it follows that f has a so-and-so explanation also; if e occurred because such-and-such it follows that f occurred because such-and-such also.

Notice how modest is my claim: I do not say that explanatory contexts are extensional in the sense of permitting the truth-preserving substitution of coextensive predicates; nor do I say that explanatory contexts are extensional in respect of anything but the explanandum place. My claim concerns extensionality only in the sense of the truth-preserving substitutability of co-referring singular terms; and it concerns extensionality only in respect of the explanandum place.

Although my claim is modest, it goes against a conventional view. The conventional view holds that whereas causal contexts are extensional, explanatory (and hence causal explanatory) contexts are not. In making the claim that explanatory contexts are, in the specified respect, extensional, I am rejecting the conventional view. Why do I do so? There are two considerations, I believe, which support my claim that explanatory contexts are extensional in respect of the explanandum place. Of these considerations, the first is by

far the less important one.

The first consideration is simply that there is a class of what appear to be perfectly good scientific explanations which must be rejected as inadequate by anyone who denies my claim that explanation is extensional in respect of the explanandum place. What all the explanations belonging to this class have in common is that they require the event to be explained to be identified with an event referred to by a different expression before the explanation can get under way. One who denies my claim of extensionality must say, implausibly, that the explanandum event has not really been explained. A pair of examples will illustrate these points.

1. Suppose that I have allowed a hydrogen flame to play on a cool surface and that upon that surface have formed droplets of a colourless, odourless liquid. What is the explanation for the formation of those droplets of colourless, odourless liquid? The explanation is this: the formation of those droplets of colourless, odourless fluid just is the formation of groups of molecules of H_2O in a liquid state; and the explanation for the formation under those circumstances of groups of molecules of H_2O in a liquid state is that the hydrogen combines with oxygen in the air to form H_2O in gaseous form which then condenses on the cool surface to form a

liquid.

This account seems to me very clearly to explain the formation, in the circumstances, of droplets of colourless, odourless liquid. But the explanation proceeded, not directly, but by the explanation of an event referred to by a singular term (a definite description) co-referential with the singular term (a different definite description) referring to the event whose explanation was originally sought. This strategy is possible, of course, only if explanatory contexts are extensional in respect of the explanandum place, so that an explanation of an event referred to in one way may be automatically transfered to an event identical with it but referred to in another way. And it would, I think, be thoroughly perverse to protest that the formation, in the circumstances, of droplets of a colourless, odourless liquid had <u>not</u> been explained, on the grounds that although an event identical with this event had been explained, it had been explained under a different, and hence irrelevant, description.

2. Why are most gases transparent at room temperature? The explanation is this: gases' being transparent <u>just is</u> gases' not absorbing (i.e. allowing to pass through) photons at a frequency within the visible range; most gases at room

temperature do not absorb photons at those frequencies because the energy of those photons is insufficient to excite atoms of the gases from their ground state to their first excited state (in which they would absorb photons). Again, this account seems to be a perfectly acceptable explanation of a certain regularity. And yet it depends entirely upon the propriety of redescribing the explananda, and then purporting to explain the explananda, as originally described, by explaining the explananda as redescribed. This explanatory strategy would be illegitimate if explanatory contexts were non-extensional.

These examples are perhaps not all that unusual. Indeed, the reader will easily be able to produce more such examples by reflecting upon theoretical identities of the sort beloved by identity theorists in the philosophy of mind: such identities as that between flashes of lightning and sudden electrical discharges, or heat and high average molecular kinetic energy, or colours of surfaces and triplets of reflectance efficiencies of those surfaces. Using such identities, we can come up with explanations of relatively familiar events which depend upon identifying events referred to in different ways, and hence which depend upon the extensionality of explanatory contexts.

The second consideration in favour of my claim that explanatory contexts are extensional in respect of the

explanandum place is that this conclusion follows from (i) the Lewisian conception of explanation that I have been defending and (ii) the widely accepted claim that causal contexts are extensional. The reason is as follows.

According to the Lewisian conception of explanation, to explain an event is to give information about its causal history, i.e. about how it was caused. Suppose, now, that we have explained an event, i.e. given information about how it was caused. If causal contexts are extensional, then, if it is true that e was caused in a certain way, and that e=f, it follows that f was caused in that same way. Explanation of an event referred to in one way is automatically transferable to an event referred to by a co-referential singular term, given the Lewisian conception and the extensionality of causal contexts.

I have defended the Lewisian conception of explanation at length already, and I will add nothing here. That causal contexts are extensional is something I am taking for granted. (For defence of this assumption, see Davidson 1980, pp.149-162; and Levin 1976.) I conclude, then, that explanatory contexts are extensional in respect of the explanandum place.

Let me bring this sub-section to an end by considering an

objection to my claim that explanatory contexts are extensional in respect of the explanandum place. Here is the objection: "Explanation is of events under descriptions. So one and the same event may be explained under one description but not under another."

What does it mean to say that events are explained under descriptions? And is there a sense in which this claim is both true and leads to the conclusion that explanatory contexts are not extensional? One thing that might be meant by saying that events are explained under descriptions is simply that in order to explain events we have to be able to talk about them, and in order to talk about them we have to have means for referring to them. Understood in this way, the claim that events are explained under descriptions is true, indeed a platitude. But, understood in this way, the claim clearly fails to entail that explanatory contexts are non-extensional. For, by parallel reasoning, we would have to conclude, since every kind of claim about things requires means of referring to those things, that every kind of claim gives rise to a non-extensional context; which is false.

Something more likely to be meant by the claim that events are explained under descriptions is this: events are explained in virtue of their falling under laws, and their falling under

laws is a matter of their satisfying one description rather than another. Understood in this way, the claim has a measure of truth (I do not say a full measure, because the claim perhaps implies acceptance of the deductive-nomological account of explanation, an account importantly different from the account I have defended). But even if the claim, thus interpreted, is unreservedly true, it still does not follow that explanatory contexts are non-extensional. The fundamental reason for this is that whether an event satisfies a given description is itself an entirely description-independent matter. For instance, whether last week's tragedy was a student's suicide or a car smash or an unexpected tornado in no way depends on how one refers to the event in question. Consequently, an event can satisfy a given description, and can even be <u>said</u> to satisfy it, despite its being referred to by the use of a <u>different</u> description. Hence (to consider the case of explanation) it may be true that an event is explained only in virtue of its falling under a law and thus of its satisfying a description, but the event will go on satisfying the description, and thus falling under the law, no matter how we refer to it. (See Levin 1976.)

One way of seeing this is to consider a plausible version of the canonical form of an explanation involving a law. It is, I

believe, obvious upon inspection that the occurrence of the explanandum term, "e", in the following is extensional:

> Why did e occur? Because c occurred, and c satisfies description D, and e satisfies description D′, and some law links events satisfying D with events satisfying D′.

It seems clear that if e satisfies description D, and e=f, then f satisfies description D also. So the observation that explanation is of events under descriptions does not support the idea that, contrary to me, explanatory contexts are non-extensional.

The second, and final, objection that I shall consider to my claim that explanatory contexts are extensional in respect of the explanandum place consists in an appeal to our intuitive judgements about certain mundane examples. I shall discuss one such example only. The response I make to it can, I hope, be applied to all other such examples, though I will not try to show this.

Here is the example. Let us allow (though it is rightly controversial) that the event of Brutus' stabbing Caesar is the very same event as the event of Brutus'killing Caesar, given that, as a matter of fact, Brutus killed Caesar by stabbing him. Now, although it seems right to say that Brutus' stabbing

Caesar occurred because Brutus judged that stabbing was the best way to kill Caesar (if this does not sound right, try stressing the word "stabbing"), it does not seem right to say that Brutus' killing Caesar occurred because Brutus judged that stabbing was the best way to kill Caesar. This seems to be a counterexample to my claim. How can I respond?

I respond by denying that the second explanation is false. That is, I think that it is true to say that Brutus' killing occurred because Brutus judged that stabbing was the best way to kill Caesar. For to explain an event is to give information about its causal history, and Brutus' deliberation about how best to kill Caesar is certainly part of the causal history that led up to Brutus' killing Caesar. So to give information about that deliberation is to explain Brutus' killing Caesar. If this still seems implausible, then consider the following, slightly amended, claim: Brutus' killing Caesar occurred in part because Brutus judged that stabbing was the best way to kill Caesar. This amended claim seems to me unexceptionable.

Still, it might be objected, does it not sound odd to say that Brutus' killing Caesar occurred because Brutus judged that stabbing was the best way to kill Caesar? Indeed, I agree that it sounds odd; but the oddity arises not because the claim is false, but because it is pragmatically inappropriate. The

claim is naturally construed as a response to the question "Why did Brutus kill Caesar?", which could be expanded to read "Why did Brutus kill Caesar, rather than not kill him?". In line with my earlier discussion of the pragmatics of explanation, I want to say that a contrastive demand for explanation of this sort is a demand, not just for any old information about the causal history of Brutus' killing Caesar, but for information about that part of the causal history which differs from the way the causal history of the world would have been if Brutus had not killed Caesar, i.e., in this case, information about Brutus' decision to kill Caesar, and the practical reasoning leading to it. Given our natural assumption about the contrastive form of the demand for explanation to which the odd-sounding claim is a response, information about Brutus' deliberation over the best way to kill Caesar seems irrelevant. This irrelevance, I suggest, is what accounts for the oddity of the odd-sounding explanatory claim. That claim is odd, but still true.

The example of Brutus and Caesar is not, of course, unique: there are others which aim to establish the same conclusion. But I believe that they could all be undermined by an appeal -- perhaps not by exactly the same appeal -- to the pragmatics of explanation. The Devil can quote scripture too.

3. The Argument From Overdetermination

In this section, I present my first major argument against token physicalism, which is intended to show that token physicalism leads to a consequence that is very hard to believe. Specifically, the aim of this section is to try to show that:

> (C) If token physicalism is true, then there is massive overdetermination in the world.

This consequence -- massive overdetermination -- is, as I have said, hard to believe, but it would be possible for a token physicalist to bite the bullet and simply accept it. At the end of this section, therefore, I will make make a few remarks meant to bring out the intuitive unacceptability of this consequence.

My argument for conclusion (C) has three premisses, which I now list:

> (P1) If token physicalism is true, then very many events have more than one explanation.
>
> (P2) If very many events have more than one explanation, then very many events -- those just mentioned -- are overdetermined.
>
> (P3) If very many events are overdetermined, then there is massive overdetermination in the world.

It is obvious that these premisses entail the conclusion (C). So the soundness of my argument hinges on whether these premisses are true. Premiss (P3) I take to be analytic, and so I shall give it no further defence. But (P1) and (P2) are very far from being obvious. So I shall defend each of these premisses in turn, assigning the defence of each premiss to a sub-section of its own.

(i) According to token physicalism (see Section One), very many events are such that they fall under descriptions proper to more than one of the sciences. So, for example, a given event may fall under both an economic, and a psychological, and a neurophysiological, and a microphysical description. And, at each of these levels of description, there will be an explanation of the event under the description at that level. So -- to continue the example -- our event will have an economic explanation at the economic level of description, a psychological explanation at the psychological level of description, a neurophysiological explanation at the neurophysiological level of description, and so forth.

This is the token physicalist picture of the world. (In presenting it, I am not, of course, committed to endorsing it.) My claim is that if this picture of the world is correct,

then very many events have more than one explanation. In order to argue for this claim, I shall now consider a highly simplified example. This example will clarify the logic of my argument. Since the argument based on the example will not depend on any special features of the example, it will be clear that the argument can be generalised to apply to <u>any</u> event which, according to token physicalism, has an explanation at more than one level of description. Here is the simplified example.

Let 'the event of microphysical type M which occurred at <u>t</u>' name an event which satisfies a microphysical description, and hence which has a microphysical explanation; and let 'the event of neurophysiological type N which occurred at <u>t</u>' name an event which satisfies a neurophysiological description, and hence which has a neurophysiological explanation. We can say, therefore, that:

> (a) The event of microphysical type M which occurred at <u>t</u> has a microphysical explanation.
>
> (b) The event of neurophysiological type N which occurred at <u>t</u> has a neurophysiological explanation.

Now, let us suppose that the microphysical event in question <u>just is</u> the neurophysiological event in question, i.e. let us suppose that we have here one of the countless token (event)

identity claims that the token physicalist, as such, is committed to making. We can, therefore, also say that:

> (c) The event of microphysical type M which occurred at t = the event of neurophysiological type N which occurred at t.

I claim that from (a), (b) and (c) it follows, by Leibniz Law, that:

> (d) The event of microphysical type M which occurred at t has a neurophysiological explanation.

From (d) and (a) we can infer, trivially, that:

> (e) The event of microphysical type M which occurred at t has two explanations, viz. a microphysical explanation and a neurophysiological explanation.

This argument, if sound, shows that, on token physicalist assumptions, a sample event has more than one explanation. But this argument has not traded on any special feature of the example. Therefore an analogous argument could be run for any of the very many events of which the token physicalist would wish to claim that it had explanations at different levels of description. So this argument, if sound, establishes premiss (P1) of the argument from overdetermination, viz. the claim

that if token physicalism is true, then very many events have more than one explanation.

But is the argument sound? I see only one line of possible objection to it. In inferring (d) from (a), (b) and (c), I appealed to Leibniz Law in a familiar way. But this use of Leibniz law is valid if, but only if, the explanatory context '... has an explanation of a given type' is extensionsal, i.e. permits the substitution salva veritate of co-referring singular terms, in respect of the place to be filled by a term referring to an event (the explanandum). (Notice that my argument does not require explanatory contexts to be extensional in any other respect.) So it might be objected that my inference of (d) from (a), (b) and (c) is not valid, on the grounds that explanatory contexts are not extensional in the way that I require. But my reasons for disregarding this objection have already been given. For, as I have already argued in Sub-section (vii) of the preceding Section Two of this chapter, explanatory contexts are extensional in the way that the valid use of Leibniz Law in the above inference requires, i.e. explanatory contexts do permit the substitution salva veritate of co-referring singular terms which occur in the event-place (explanandum-place) of those contexts. I refer the reader back to the cited sub-section where I argue for

this.

My defence of premiss (P1) of the argument from overdetermination is now complete. Using a highly simplified example, I have argued that, on token physicalist assumptions, the event in the example has more than one explanation. And I have pointed out that this argument can legitimately be generalised, so as to sustain the conclusion that, if token physicalism is true, then very many events have more than one explanation.

(ii) I turn now, therefore, to a defence of premiss (P2) of the argument from overdetermination. Premiss (2), of course, states that:

> (P2) If very many events have more than one explanation, then very many events -- those just mentioned -- are overdetermined.

In order to defend this premiss, I shall continue to use the highly simplified example introduced in the preceding sub-section. And, as in that sub-section, it will be clear that we can generalise from this single example, this time to establish the general claim of (P2).

The example I was working with concerned an event -- we may call it e -- which (as I argued) has two explanations, a

microphysical explanation and a neurophysiological explanation. I am now going to defend my suggestion that, since e has two explanations, it is for that reason overdetermined.

In the previous section, Section Two, I defended the claim that to explain an event is to give information about its causal ancestry, and in the light of that conception of explanation I think that my suggestion is plausible: since e has two explanations, there are two causal stories, one microphysical, the other neurophysiological, to be told about its ancestry; and since there are these two causal stories, each one of which (we may suppose) describes a set of causal factors sufficient for the occurrence of e, it appears to follow that e is overdetermined. In short, e is caused both by some set of preceding events inasmuch as they are microphysical and by some set of preceding events inasmuch as they are neurophysiological; and this seems to imply that e is overdetermined.

One might doubt this. For under certain conditions an event can have two explanations and still not be overdetermined. One such condition is this: if the two explanations of the one event describe different stages in the same causal chain which leads to the event, then it will not follow that the event is

overdetermined. So, for instance, from the fact that we can explain a house-fire by describing a gradual build-up of gas and by describing a short-circuit it does not follow that the fire was overdetermined; for the build-up of gas and the short-circuit are different stages in one causal chain which led to the fire. But this is not the situation in respect of e: the microphysical causal story and the neurophysiological causal story do not describe one single causal chain that leads to event e. And what goes for e will go also for the very many events of which e is an illustrative example.

Another condition under which an event can have two explanations, without thereby being overdetermined, is this: if the two explanations of the one event describe simultaneous but complementary causal factors, then there need not be overdetermination. So, for instance, if we explain the explosion at the petrol station by describing a smoker's lighting up and by describing the simultaneous presence of a heavy concentration of petrol vapour, then we will not be able to infer that the explosion was overdetermined, the reason being that the smoker's lighting up and the concentration of petrol vapour are complementary causal factors which interact with one another (and other factors) to produce the explosion. But, again, this is not the situation in respect of event e and

its two explanations; for it is not the case that the microphysical factors described by the microphysical explanation interact with, are complementary to, the neurophysiological factors described by the neurophysiological explanation -- each set of factors is presumably complete by itself. (And in any case we have no theory of such interaction.)

So, although there are conditions under which an event can have two explanations without thereby being overdetermined, event e does not satisfy those conditions. Therefore I believe that we should accept my claim that event e, since it has two explanations, is overdetermined. But what goes for e goes also, I think, for the very many plurally-explained events of which e is an illustrative example. So we can conclude that the general claim expressed by premiss (P2) is correct: if very many events have more than one explanation, then very many events -- those just mentioned -- are overdetermined.

(iii) I have now defended each of the three premisses of the argument from overdetermination. Together, these premisses entail that if token physicalism is true, then there is massive overdetermination in the world.

I find this consequence of token physicalism to be highly

counterintuitive, and so I regard it as counting heavily against token physicalism. But is not counterintuitive because there is something conceptually or logically improper about the very idea of overdetermination -- I do not think it is improper in these ways. Rather, it is the sheer scale of the overdetermination apparently implied by token physicalism which is alarming, a scale which is quite unexpected. Also, overdetermination is, no doubt, acceptable when it is unavoidable, but in the case of token physicalism the overdetermination seems to be entirely unnecessary. For if, as token physicalism holds, every event has a physical description and hence has a physical explanation, then every event can be explained without appealing to the special sciences at all; so it seems gratuitous to superimpose the special scientific levels of description and explanation. If it is a good principle that instances of overdetermination should not be multiplied beyond necessity, then token physicalism is a grave offence against this principle of parsimony.

4. An Unsound Motivation For Token Physicalism

What, in broad terms, is the philosophical motivation which gives rise to token physicalism? It is primarily, I think, a desire to be a physicalist without being a type physicalist. Type physicalism, it is thought, is a version of physicalism

implausibly strong and, in the light of the availability of token physicalism, unnecessarily strong. I shall now examine the objection most commonly offered against type physicalism (and hence, indirectly, in favour of token physicalism).

It is distinctive of the type physicalist to maintain, not only that every event is a physical event, but that every property is a physical property, so that every science can, in a very strong sense, be <u>reduced</u> to physics. And the standard token physicalist objection to this is to claim that there are very many properties, for instance, biological or economic properties, such as being a kidney or being a money, which are simply not identical with physical properties; which are, indeed, generally not even coextensive with physical properties (see, for example, Kim 1978).

It is, however, crucially important to see that this objection simply does not refute the claim that every property is a physical property. Let us distinguish two forms of type physicalism, which we can call <u>eliminative</u> and <u>non-eliminative</u>. Both forms involve the claim that every property is a physical property. But non-eliminative type physicalism adds to this claim the further claim that <u>there exist</u> special scientific properties. Hence non-eliminative type physicalism is committed to the view that special scientific properties are in

general identical with physical properties. And it is of course this last view which the standard token physicalist objection no doubt demolishes. By contrast, eliminative type physicalism denies the existence, proposes the elimination, of any special scientific properties which cannot be identified with physical properties. Hence the eliminative type physicalist is not committed to affirming the claims of special scientific/physical property identity which the token physicalist, quite correctly, points out to be false. Non-eliminative type physicalism is strongly reductionist, and is indeed liable to the token physicalist objection noted. But eliminative type physicalism, which I defend, is not reductionist, and hence escapes this objection.

The token physicalist has noticed the irreducibility -- the non-identifiability -- of most special scientific properties to physical properties, and has inferred from it the falsity of type physicalism. This inference, I have just argued, is a mistake. What I want to suggest now is that from this irreducibility (or non-identifiability) the token physicalist should have drawn a very different conclusion. For the situation may be like that in which we notice for the first time that the property of being phlogiston cannot be identified with any property of physical theory, and conclude from this

fact that there is probably no reason left to believe in the existence of phlogiston. In this situation it would surely be absurd to take the irreducibility of the property of being phlogiston as indubitable proof of the autonomy of the phlogiston theory of combustion! The irreducibility of the property of being phlogiston seems rather to lead to the conclusion that phlogiston theory, and with it phlogiston, should be eliminated; and this is because, but only because, with the acceptance of the oxidation theory, the phlogiston theory is left with no explanatory work to do.

In the light of the possibility illustrated by the case of phlogiston, let us now return to token physicalism. If it is really true, as the token physicalist insists, that every event is a physical event, and therefore has a physical explanation, then, since events are the only explananda, there is, after physics has done its stuff, no explanatory work left to be done by the special sciences. So the special sciences, with their attendant ontology, should be, like the phlogiston theory of combustion, eliminated on grounds of redundancy. And this will lead the would-be token physicalist straight to eliminative type physicalism and hence to the strong claim that every property is a physical property. And this, of course, is the very different conclusion I mentioned. (By the way, the

argument of this paragraph is intended only as a straw in the wind, not as a full-scale argument for eliminative type physicalism. The latter is supplied in Part 2.)

5. The Epistemological Argument

In the last section I considered a general philosophical motivation for token physicalism and suggested that, quite apart from failing to support token physicalism, the chief component of this motivation pointed, rather, in the direction of eliminative type physicalism. In this section I shall present my second major argument against token physicalism, the Epistemological Argument. This argument is also connected with the reasons we could have for endorsing token physicalism.

Specifically, my concern will be to press the question how a would-be token physicalist could come justifiably to believe the particular token identity claims which token physicalism entails must be true. I have in mind, of course, such claims as that a particular, dated, biological event is identical with some particular, dated, physical event. The token physicalist must hold that countless such claims are true. My argument against token physicalism will be that there is no way in which the token physicalist could, even in principle, acquire knowledge of such token identity claims, consistently with

maintaining reasonable belief in the other claims constitutive of the doctrine (i.e. claims (b) and(c)). (This project is inspired by Hornsby 1980-1.) There is, I shall argue, no good reason to accept specific token identity claims of the kind that a fleshed-out version of token physicalism must involve.

Perhaps I should emphasise that I shall not in this section be arguing that there is anything wrong with accepting an ontology of events (the tokens in question), or with asserting claims of token (i.e. event) identity, or even with asserting trans-scientific claims of token identity of the kind to which the token physicalist is committed. For all that I say in this section, all these things may be entirely free from impropriety. My argument, as its name suggests, is an epistemological one. The problem it raises for token physicalism is the problem of how the token physicalist, as such, can come justifiably to believe the required claims of token identity.

Here is an example of the sort of problem I mean. One obvious way of coming justifiably to believe token identity claims would be to exploit the fact that they are entailed by claims of type identity. So one could learn that a given special scientific event is identical to a given physical event by learning first of all that every special scientific property

is identical with some physical property. Then, having identified an event as being of a certain special scientific type, one could at once infer with what physical event it is identical, i.e. an event of the physical type known to be identical with the relevant special scientific type. But this way of coming to know special scientific/physical token identities has a very obvious drawback: it provides the token physicalist, by hypothesis, with excellent reasons for affirming precisely those type (i.e. property) identity claims which token physicalism, in claim (b), officially denies!

So how could the token physicalist come justifiably to believe the token identities? Since the tokens in question are events, we should consider some views concerning the identity conditions for events, in the hope that these views may shed light on the practical problem of identification faced by the token physicalist. I shall be considering these views, not in order to argue that they are wrong, that they mis-specify the identity conditions of events, but in order to show that they do not help the token physicalist to avoid the epistemological predicament he faces.

According to Jaegwon Kim, "an event (or state) is a structure consisting of a substance (an n-tuple of substances), a property (an n-adic relational attribute), and a time" (see Kim

1976, pp.159-177, esp. 160). Using Kim's operator, [, ,], which enables us to form expressions referring to events, we may state Kim's identity conditions for events as follows, where "x" and "y" denote substances, "P" and "Q" properties, and "t" and "t'" times:

> [x,P,t] = [y,Q,t'] just in case x=y, P=Q, and t=t'. (Kim 1976, p.161)

But this account of the identity conditions of events cannot possibly be used by the token physicalist. This is because, on this account, a special scientific event could only be identical with a physical event if the property exemplified in each event was the same; that is, if the special scientific property was identical with the physical property. But then, in asserting the event identity, the token physicalist would be obliged to accept a type identity -- contrary to claim (b) of token physicalism. The Kim criterion, as it stands, does not permit the statement of token physicalism (see Horgan 1981, pp.400-401).

Yet a modification to the Kim account has been proposed which is based on Goldman's notion of level-generation (see Swinburne 1982, pp.176-7 and Robinson 1982, pp.17-19). It differs from Goldman's original view inasmuch as Goldman was concerned only

with those events which are actions, and explicitly set out to define a relation holding between distinct actions (see Goldman 1970, p.20), whereas the modification concerns all kinds of events and is supposed to be part of an account of event identity. It runs as follows. First, the Kim account should be weakened so that it states only a sufficient, and not a necessary and sufficient, condition of event identity. Then, we add the claim that e=e' if either e level-generates e' or e' level-generates e. Finally, we specify the ways in which an event may level-generate another event:

> e level-generates e' iff EITHER e constitutes e', given the context, which may be conventional or causal, OR e is e' but under a more specific description, OR it follows, from the occurrence of e and the circumstances of its occurrence, that e' occurred.

Will any of this help the token physicalist? The first mode of level-generation appeals to contexts both conventional and causal. The appeal to conventional contexts looks to be unpromising. For, first of all, the token physicalist may not want to say that some physical event counts as a special scientific event in virtue of a convention, because that might seem to cast into doubt the reality of the special scientific event and the properties mentioned in its description: it might suggest a picture of the world in which the physical alone was

real, but onto which conventional properties were projected. But secondly, and much more importantly, the relevant conventions just do not seem to exist. It is a familiar convention that we recognise the utterance of certain words in certain contexts as constituting an act of promising. But no similarly familiar conventions spring to mind concerning, say, biological events. The appeal to causal contexts, however, seems on its face to hold out more hope. Causal contexts are introduced to enable us to say that, for example, Brutus' stabbing Caesar just is Brutus' killing Caesar, given that the stabbing caused Caesar's death. Someone tempted by functionalism in the philosophy of mind might think that there was an opportunity here. For suppose we give functionalist accounts of all special scientific properties in terms of their causal roles; then, thinking of any particular special scientific event as being any physical event caused and causing in one of the appropriate ways, we ought to be able to tell with what special scientific event a given physical event was identical by noticing the latter's causal role and then consulting the details of our functionalist theories. But this ambitious strategy would encounter grave problems. To begin with, the functionalist theories would have to be reductionist in this sense: ultimately, the causal roles of the functional properties would have to be specified by reference to physical

causes, i.e. causes physically described. For if this requirement were not met, and special scientific properties were specified only by reference to causes special-scientifically described, then there would be no deciding whether a given physical event played the relevant causal role to qualify as a special scientific event unless one already knew a number of event identities. But the stringency of this requirement throws into sharp relief the main problem with the strategy under consideration, which is as follows. The token physicalist who would employ it must not only know that functionalism, in respect of all the special sciences, is true, but know also the details of particular functionalist theories, functionalist theories satisfying the stringent requirement. I remark first that no one at present has this detailed knowledge. Unsurprisingly; for how could one acquire it? One could perhaps become convinced that <u>some</u> functionalist theory must be true, on the grounds that type identity claims were empirically falsified, but this knowledge is insufficient for the token physicalist purpose. Nor is it remotely plausible to suggest that the relevant functionalist theories were somehow analytically true, so that they could be learnt <u>a priori</u> by reflecting on ordinary language. My suspicion, though I do not see how to prove it, is that one could only establish the needed functionalist theories if one already had

a way of establishing claims of token identity.

The second mode of level-generation appeals to the idea that an event can be referred to under more or less full descriptions. An apple's turning coloured may just be its turning red, if red is the colour it turned; and of course "being red" is a determinate of the determinable "being coloured". Might this mode of level-generation enable the token physicalist to establish special scientific/physical identities? Presumably so, but only if it is plausible to think of physical properties as determinates of determinables which are special scientific properties. Is this plausible? Suppose we analyse the property of being coloured disjunctively; so to be coloured just is to be red, or to be blue, or ... . We might then think we could offer a parallel analysis of special scientific properties in terms of a disjunction of physical properties; so to be S just is to be P1, or to be P2, or ... . Knowing such analyses, we could, at last, inspect the physical state instantiated before us and tell with which special scientific state(s) it was identical. But the snag here is obvious. What I have been calling analyses are not analyses in the sense that entails the possibility of their being known a priori. And they are hard to distinguish from a detailed statement of token physicalism's

event identity claims. So if we assume that they can be known unproblematically, we have just assumed away the epistemological problem of token physicalism. This solution to the token physicalist's predicament presupposes that it has already been solved.

The final mode of level-generation depends on the idea that it sometimes follows from the fact that an event of one type has occurred in a certain set of circumstances that an event of a second type has occurred. So my asserting that p may just be my contradicting myself, if I have recently asserted that not-p. But I cannot see how this mode of level-generation could be of assistance to the token physicalist unless it was in one of the ways I have just discussed. For the required entailments could only be sustained by analytically, or in some other way necessarily, true definitions of special scientific properties; and I do not see how these definitions could fail to be of the functionalist or disjunctive sorts already considered.

An identity criterion for events which, in contrast to the unmodified Kim criterion, certainly does permit the formulation of token physicalism is that defended by Donald Davidson (Davidson 1980, pp.163-180). Two events (as it were) are the same iff they have exactly the same causes and effects:

x=y if and only if ((z)(z caused x ↔ z caused y) and (z)(x caused z ↔ y caused z)

But this criterion, though not entirely useless to the token physicalist, is of very sharply limited utility. The reason is that, as Davidson himself is keen to stress, causes and effects are themselves events, and this has the consequence that, although the criterion involves no formal circularity, a certain epistemic circularity afflicts the token physicalist's efforts. This will become clear from a simplified example.

Suppose we are interested to know whether a mental event, m, is or is not the same event as some physical event, p. Let us suppose that m has just one cause and one effect, c and e; and that p also has just one cause and one effect, c′ and e′. Then, according to the Davidsonian criterion, m=p iff c=c′ and e=e′. But these causes and effects are almost certain to be referred to by descriptions couched in different vocabularies; that is, c and e in mentalistic vocabulary, and c′ and e′ in physical vocabulary. So it will not be just obvious whether or not the required identities hold. But in fact, of course, we have made no progress at all, since causes and effects are events and it is precisely the difficulty of establishing event identities which gave rise to the original problem. Someone might reply

here, on behalf of the token physicalist, that it is, after all, possible to break into the circle. For token physicalism does not prohibit the assertion of all type identities, and if the token physicalist were to assert just a few, then it would be possible to establish a few event identities without relying on the Davidsonian criterion, which could then, and without circularity, be used to establish the rest. But this reply has an air of desperation about it. It states what is, no doubt, a possibility; but I doubt that it could secure anything like all the event identities the token physicalist needs, or at least not without requiring the assertion of a good many claims of property identity of the disapproved sort.

Finally, let us examine whether the account of the identity conditions of events devised by Myles Brand can help the token physicalist (see Brand 1976). Expressed without a complication irrelevant to current concerns, the account is as follows, where the quantification is over a domain of spatio-temporal regions and "R" stands for the relation "... occurs within...":

$$e=f \text{ iff } \Box(s)(eRs \leftrightarrow fRs)$$

Perhaps there are cases where the token physicalist could come to a reasonable decision on the basis of this criterion, though one might wonder how sheer observation, as it were, of

an individual event could establish the nature of its modal properties; I am thinking here of claims of biological/physical event identity. But there are also plenty of cases where even this seems most unlikely, and here I am thinking of claims of sociological/physical or mental/physical identity. The problem with these cases is that it is very hard to specify with any precision even the actual spatio-temporal locations of sociological events or mental events. For example, where, precisely, are we to locate increases in the rate of inflation or instances of enhanced class consciousness? Or my noticing a mark on the ceiling? (Cf. Hornsby 1980-1 for a development of a very similar point.) The same question can even be asked about common sense/physical event identities: in an example due to Haugeland (Haugeland 1982, pp.100-101), how can we determine the spatio-temporal boundaries of the movement of a wave on a pool? There is, I think, an explanation of the difficulties illustrated by these examples, and it leads to a further difficulty. The explanation is that our identification of times is closely bound up with our identification of events, and that our identification of spaces is closely bound up with our identification of substances, substances very likely to be involved in the events we are interested in. The further difficulty is that, in exploiting the Brand criterion, the token physicalist may well be able to pick out the

spatio-temporal region in which a special scientific event occurs. But the specification of this spatio-temporal region will be thoroughly steeped in the vocabulary of the special science in question; that is, the spatial component of this region will be specified by reference to some substance, special scientifically described, and the temporal component of the region will be specified by reference to events, also special scientifically described. Moreover, since the description of any region picked out as that in which some favoured <u>physical</u> event occurs will itself be specified in <u>physical</u> vocabulary, it will just fail to be obvious whether the two regions are one and the same. The Brand criterion could help the token physicalist only if there were available a vocabulary for describing spatio-temporal regions which was somehow <u>neutral</u> between the many sciences; but surely no such vocabulary exists. So the epistemological problem for token physicalism which I sketched at the start of this section remains unresolved: the token physicalist, as such, cannot come justifiably to believe the specific claims of token (i.e. event) identity which his view requires. This does not, of course, show token physicalism, or the specific claims of token identity, to be false (on pain of accepting a verificationism which I reject). But it does, I think, reveal a major embarrassment for token physicalism. For it shows that even if

token physicalism is true, we cannot justifiably believe that it is.

6. The Generalised Davidsonian Argument

It might be thought that the extended exploration in the previous section of possible ways in which the token physicalist could come justifiably to believe the required claims of special scientific/physical event identity has overlooked an important and relevant contribution to recent debate. I refer, of course, to Donald Davidson's argument for the token physicalist component of his doctrine of anomalous monism (see Davidson 1980, pp.207-225).

Davidson's conclusion, however, is merely the restricted one that every mental event is identical with some physical event. For his argument to support the general thesis of token physicalism, it would have to be generalised. There is doubtless more than one way in which this could be done, but I shall only present and discuss the version which seems to me to be the most promising.

The generalised Davidsonian argument has four premisses:

(1) Special scientific events cause and are caused by physical events.

(2) Events related as cause and effect fall under strict laws.

(3) There are strict physical laws.

(4) There are no strict special scientific laws on the basis of which special scientific events, qua special scientific events, can be predicted and explained.

The reasoning by which these premisses lead to token physicalism closely parallels the Davidsonian original. Suppose a physical event causes a special scientific event, by (1); by (2), the special scientific event must fall under a strict law; but since there are no strict special scientific laws for it to fall under, by (4), it must fall under a physical law, which exists, by (3), and hence must be a physical event; and what goes for this special scientific event clearly goes for them all.

I shall not challenge any of the steps in this reasoning. So let me turn to the premisses, of which the first three seem to me to be acceptable. Premiss (4), however, I shall challenge, and in the following fashion. I do not say that (4) is false, or that it could not be known or reasonably believed. My claim is that we could not know it without thereby ceasing to have good reason to take special scientific properties to exist: the price we pay for reasonable acceptance of (4), and hence the cogency of the generalised Davidsonian argument, is that

acceptance of claim (c) of token physicalism becomes unreasonable. So the generalised Davidsonian argument is useless to the token physicalist.

Before I elaborate my challenge to (4), I should like to remark upon one alarming consequence it has. Premiss (4) is the analogue, in the generalised Davidsonian argument, of the premiss in Davidson's original argument which affirmed the anomalism of the mental (psychological). Therefore, in accepting (4) as a premiss in an argument for token physicalism, the token physicalist is committed to affirming the anomalism of <u>all</u> the special sciences, not just psychology, i.e. to affirming that there are no strict special scientific laws of any kind. Since for any form of the Davidsonian argument to work, "strict" laws must include all those laws sufficient to back up claims of singular causation, i.e. all causal laws, the token physicalist who uses the Davidsonian argument is committed to denying that there are any special scientific <u>causal</u> laws. But this is a strong commitment, and a token physicalist might well prefer to eschew the generalised Davidsonian argument precisely in order to avoid it, especially if he accepts a suitably generalised version of Ted Honderich's objection to the anomalism of the mental that it deprives the mental of genuine causal efficacy (see Honderich 1982).

Let me return now, however, to the main issue: can (4) be established in a way consistent with continuing to have good reason to affirm claim (c) of token physicalism, the claim that the special scientific properties exist? I shall argue that there is no such way. Let me consider first a putative a priori method. (This method roughly corresponds to Davidson's strategy in defending the anomalousness of the mental, though I shall not argue this interpretative point.) The method consists in coming to adopt a view of the special sciences, conceived of as systems of discourse, according to which we are not doing what we thought we were doing in uttering sentences belonging to these systems. An analogy would be with a prescriptivist or emotivist account of ethical discourse, according to which ethical utterances are not, despite superficial appearances, fact-stating utterances at all; their function lies elsewhere. Just as, by a priori analysis, we come to see that ethical utterance is not fact-stating, so too we could come to see that special scientific discourse is not law-stating, and hence that there are no special scientific laws.

To this possible method I have two objections. First, I doubt whether anyone would actually accept the non-law-stating analysis of special scientific discourse unless they were

already convinced on other grounds that there was something terribly wrong with taking it to be, as it seems to be, law-stating. This observation is prompted by the analogy with metaethics, where it is generally only after becoming convinced that ethical utterance cannot be fact-stating that philosophers try to salvage it by adopting non-cognitivist accounts.

But secondly, and more seriously, the attempt to reanalyse special scientific discourse is almost bound to undercut the rationale for believing in special scientific properties (i.e. claim (c) of token physicalism). For example, if one argued a priori for an instrumentalist interpretation of the special sciences, one would certainly support the denial of strict special scientific laws, but one would also cease to have a reason for taking the properties mentioned in the special sciences to be real. The comparison with metaethics suggests just this. For having accepted prescriptivism or emotivism one will hardly go on believing in moral properties (if one had not already ceased to believe in them). And, quite generally, it is not open to the token physicalist to adopt any non-fact-stating analysis of special scientific discourse, since this will eliminate our ontological commitment to the properties apparently invoked in it. So what the token physicalist needs is just this: an a priori analysis of special

scientific discourse which is fact-stating but not law-stating. I have no idea what this could be. With an eye on the Davidsonian prototype of the argument now being considered it is tempting to offer some interpretative account of special scientific discourse. But even if this were plausible for all the special sciences, it would still involve the difficulty that interpretative science does not seem to be a fact-stating affair, and hence is unsuited to current purposes. Is there another possibility? Perhaps there cannot be. For if we accept the principle that a predicate stands for a genuine property only if it figures in a statement of law, then there cannot be a non-law-stating analysis of special scientific discourse which also preserves the special scientific ontology.

The only other methods for establishing (4) are thoroughly a posteriori. The first involves the empirical discovery of the sheer breakdown of the special sciences, that is, of the systematic predictive failure of their supposed laws. I do not mean the uncovering of a few exceptions to special scientific laws; for in the face of such exceptions the correct response might well be to retain the laws as decent approximations to the truth pending further research. I mean something so drastic that it would lead us to reject all special scientific

laws root and branch as irredeemably false. But while it is manifest that such empirical discoveries would establish (4), it is equally clear that, in exposing the special sciences as false, they would destroy our reasons for believing in special scientific ontology. And this objection applies also to the second a posteriori method, which would involve the elimination of the special sciences, on the ground of their theoretical dispensability, as recommended by eliminative type physicalism. So I conclude that the generalised Davidsonian argument is powerless to justify the set of claims constitutive of token physicalism. The Epistemological Argument against token physicalism remains undefeated.

CHAPTER TWO

Physicalism: From Supervenience To Elimination

Token physicalism, like the eliminative type physicalism which I favour, constitutes a response to the problem of the many sciences. But token physicalism, as I have just been arguing, is unsatisfactory. In recent years, however, several philosophers have advanced a new version of physicalism which is claimed to enjoy the main advantage of token physicalism, which is that it is in some appropriate sense non-reductive, without at the same time suffering its drawbacks. This version we may call supervenience physicalism (see Lewis 1983; Horgan 1981 and 1982; Haugeland 1982), and I shall in this chapter be presenting an argument against it. The argument I present is not the only one I can imagine being raised against it, and indeed I strongly suspect that analogues to the arguments offered in the previous chapter against token physicalism could also be devised to its detriment.

Unlike token physicalism, supervenience physicalism does not affirm the token identity thesis that every event is a physical event. Instead, the domains corresponding to the many sciences

are to be thought of as glued together with the cement of supervenience; and the primacy of physics will consist accordingly in the thought that all the special scientific facts supervene upon the physical facts. It is, of course, the primacy of physics, understood in this way, which makes supervenience physicalism a form of monism, consistent with our metaphysical intuition to the effect that there is only one world.

Supervenience physicalism must involve at least these four claims:

(a) All facts supervene upon the physical facts.

(b) The special scientific properties exist (are real).

(c) It may not be generally possible to identify special scientific properties with physical properties.

(d) It may not be generally possible to assert special scientific/physical event identities.

Each of these claims is essential to supervenience physicalism if it is to represent a distinctive physicalist position immune to the criticisms levelled at alternative versions of physicalism. (a) is evidently crucial, since it is an explication of the distinctively physicalist notion that physics is in some sense fundamental. (b) is necessary to

distinguish supervenience physicalism from eliminative type physicalism, which is consistent with (a), (c) and (d). Eliminative type physicalism maintains, of course, that every property is a physical property, and that special scientific properties which cannot be identified with physical properties must be eliminated. (c) ensures that supervenience physicalism, like token physicalism, is non-reductionist and hence impervious to traditional objections to non-eliminative type physicalism. Finally, the inclusion of (d) makes it clear that supervenience physicalism is not committed to the controversial event identity component of token physicalism or, indeed, to an ontology of events.

1. Supervenience and Supervenience Physicalism

The concept of supervenience is evidently made to bear a heavy load in supervenience physicalism, and we need to know a little more about it before we can decide whether it really is strong enough to take the weight. Unfortunately, supervenience is not a notion which leads a very active life in the world of ordinary language, and so attempts to make it more precise, by appeal to ordinary usage, inevitably have an air of stipulation about them. Nonetheless, the intuitive idea involved can be stated by the use of any of three equivalent slogans: "Supervenient difference guarantees subvenient difference!", or

"Subvenient sameness guarantees supervenient sameness!", or "No supervenient difference possible without a subvenient difference!" This idea is held by supervenience physicalists to express their distinctively physicalist claim that the physical in some sense fixes or determines the non-physical (for doubts about its adequacy in this regard, see Grimes 1988).

Taking as an example the claim that the mental supervenes upon the physical, we can expand these slogans in three equivalent ways:

> (1) Necessarily, if there's a mental difference, there's a physical difference.
>
> (2) Necessarily, if there's no physical difference, there's no mental difference.
>
> (3) It's not possible that there be a mental difference but that there not be a physical difference.

In these formulations the relevant difference may be either that between two objects at one time or that between the same object at two times.

Jaegwon Kim, taking supervenience as a relation between sets of properties, offers this account of what he calls strong supervenience:

> The set of properties, A, strongly supervenes upon the set of properties, B, iff necessarily, for any object x and any property F in A, if x has F, then there exists a property G in B such that x has G, and necessarily if any y has G, it has F.

Kim's definition of weak supervenience is exactly the same as this, except that it omits the second "necessarily" (see Kim 1987, pp.315-6).

Can Kim's supervenience, strong or weak, help in the precise formulation of the crucial claim (a) of supervenience physicalism? It cannot. The problem is that Kim's supervenience defines a relation between sets of properties of a single object, whereas this is often not adequate to what appears to be the truth. Two kinds of cases exhibit this inadequacy. In the first kind, the supervenient property of an object supervenes not merely upon some subvenient property of the object, but also upon properties of other objects. So, for example, one might claim that the epistemic property, had by Catherine, of knowing that James is fat supervenes upon certain physical properties; but the relevant physical properties are not only those instantiated by Catherine, but also, since knowledge entails truth, at least one instantiated by James, namely that of being fat. In the second kind of case the trouble is even more serious, since the having by an object of

a property seems to supervene upon no property had by that object at all. So, for example, one might claim that all the facts about inflation supervene upon physical facts. But if having a certain rate of inflation is a property of anything, it is a property of a national economy, and it is nonsense to ascribe physical properties, by which I mean, roughly, properties mentioned in the laws of physics, such as microphysical properties, to a national economy. (My point here could also have been made by appeal to so-called "wide" psychological states, but it is important to see that the problem is a more general one.)

Kim himself has disputed the allegation that his definitions of supervenience have this limitation (Kim 1987, pp.323-4). His response to examples of the sort I have given is to suggest that they call for "an explicit recognition of relations and relational properties". So in my first example, he would want to say, perhaps, that Catherine's knowing that James is fat indeed supervenes only upon properties had by Catherine, but that these properties will include the relational property of being causally connected in the right sort of way with fat James. And in my second example, which I think gives more trouble, Kim might want to say that the national economy's having a certain rate of inflation supervenes upon the national

economy's being constituted by certain physical objects having certain physical properties, or some such thing.

But I do not think Kim's efforts here can succeed. The problem is with the relations and relational properties which he must invoke. It is not just that they are contrived, though they are and this should arouse suspicion. It is that they are not physical, and it is after all a version of physicalism which we are trying to formulate: we want to be able to claim that the instantiations of all properties supervene upon the instantiations of physical properties. When I say that these relations and relational properties are not physical, I mean simply that they are not relations and relational properties mentioned in the laws of physics. (The crude, though for current purposes adequate, account of the physical that I am working with in this paper makes the physical a narrower category than the merely non-mental.) Consider a relatively straightforward case of supervenience, that of a state of belief, internalistically construed to make it easier. It is tempting, and no doubt Kim would be tempted, to say that Jane's believing that grass is green supervenes upon some physical state of Jane. But what physical state? The same temptation might lead us to reply: the physical state consisting in Jane's having a brain exemplifying _____ physical properties. But

the property thus ascribed to Jane, alleged to be the one upon which the property of believing that grass is green supervenes, is the property of having a brain exemplifying _____ physical properties; and of course this is simply not a physical property, even with the blank filled in. (It is not even a neurophysiological property; for such properties are had by brains or cells or what have you, not by persons.) Moreover, if in the light of this difficulty you wish to claim instead that Jane's believing that grass is green supervenes upon some property had by Jane's brain, rather than Jane herself, then Kim's supervenience is no longer available to you as an explication of your claim, since the supervenient and subvenient properties are now no longer had by the same object (i.e. Jane).

The problem for Kim's appeal to relations and relational properties that this example illustrates is even clearer when what we are after is a formulation of physicalism quite generally, especially if we take seriously, as Kim himself once did (Kim 1978, p.155), the idea that the microphysical facts determine all the facts. For the notion that there is any one kind of object which has both (genuine) microphysical and special scientific properties is absurd (no table has, for instance, spin, and no quark is brown, or viviparous); and the

ersatz relational properties which might in other cases seem to be able to save the day will in this case be all too obviously non-physical.

So we need something other then Kim's supervenience, weak or strong. What we must do is to globalise the notion of supervenience, so that we express the idea that the way the world is in respect of its supervenient properties supervenes upon the way it is in respect of its subvenient properties. We should start by taking the objects whose differences are in question as being whole worlds; and then, to capture the idea of modality essential to the intuitive concept of supervenience, we should quantify over these worlds. We shall then have, as formulations of claim (a) of supervenience physicalism, one of these three equivalent claims: any two worlds which differ in any way will differ physically; any two worlds which do not differ physically will not differ in any way; no two worlds differ in any way without differing physically.

But these formulations are too rough to serve. Their most serious drawback is that the quantification in them is over all possible worlds, which would turn claim (a) of supervenience physicalism into an implausible claim of necessary truth (see Lewis 1983a, p.362 and Petrie 1987, p.120). So the domain of

worlds over which we quantify must be restricted in some fashion. Terence Horgan, in his most recent statement of supervenience physicalism, argues that the domain of worlds should include only those which are, in his defined sense, physically possible (Horgan 1987, p.491). David Lewis suggests that the domain should contain only worlds "where no natural properties alien to our world are instantiated", where a property is alien to a world iff (1) it is not instantiated by any inhabitant of that world and (2) it is not analysable as a conjunction of, or as a structural property constructed out of, natural properties all instantiated by some inhabitant of that world. So there is certainly room for dispute over the details of the best formulation of supervenience physicalism. But they will not matter for what follows.

2. Supervenience Physicalism and Possible Worlds

Claim (a), spelt out as some thesis of global supervenience, is clearly the nub of supervenience physicalism. But I do not propose to challenge its truth. Rather, I want to pose a question about its interpretation which, properly answered, makes it impossible to assert it together with claim (b) of supervenience physicalism, namely the claim that the special scientific properties exist or are real. So my argument will be that supervenience physicalism is at war with itself.

Supervenience physicalism holds that there are physical properties, which are real, that there are special scientific properties, which are also real, and that these two kinds of properties are related to one another in the way described by some claim of global supervenience. The supervenience physicalist therefore requires the notion of supervenience to play a very special theoretical role. Supervenience must be some relation both strong enough to ensure that the physical properties determine the non-physical ones and weak enough to allow to the non-physical properties an appropriate ontological autonomy. But what kind of relation is this supervenience, which holds between real properties, and which enables the physical to hold in check the non-physical? It is unlike logical necessity, since it is supposed to hold between distinct existences, namely distinct kinds of real properties; and yet it is not the same thing as causal or natural necessity. So what on earth is it? This is the question about the interpretation of claims of global supervenience to which I referred above. (My question resembles one pressed in Blackburn 1984, pp.182-187, especially pp.186-187 where he speaks of the need for an "intelligible connection"; but how closely, I cannot say.)

Something must be said about this real relation of

supervenience, supposedly capable of holding between real properties, because if it is not, then we will just be left with a situation in which claims of global supervenience may often, and with plausibility, be asserted, but in which it will be uncertain what accounts for their assertibility, i.e. in virtue of what reality they are assertible; and it may be that an account in terms of our projective tendencies, an account which may not view our ascriptions to the world of special scientific properties as fact-stating, should be accepted as the explanation of this assertibility. (A related possibility is discussed in Klagge 1988.)

The extensive literature on the topic of supervenience seems to me not to address this question (though see, perhaps, Bonevac 1988). It is concerned instead with the question of what claims of supervenience mean, rather than the more metaphysical question of what the real relation of supervenience, capable of holding between real properties, is. My intuition is that we have not yet been told enough about supervenience, especially in the light of the many philosophical tasks it is introduced to perform.

It might be suggested that supervenience just is that real relation which holds between real properties whenever an appropriate claim of supervenience may be asserted, and that

when it obtains it is simply a primitive and irreducible modal fact about which nothing further can be said. It might be further suggested that in this respect supervenience is just like causal or natural necessity, which is also notoriously hard to illuminate. But I do not think that these suggestions are satisfactory. Supervenience compares unfavourably with causal or natural necessity in one important respect, which is that while we can argue plausibly that causal or natural necessity is a well-entrenched part of our common sense conceptual scheme, this is just not true of supervenience. For, as I have already suggested, although the word, "supervenience", occurs, though not commonly, in ordinary language, philosophers have given to it a sense, or rather senses, far more precise than it bears in ordinary language; indeed, the word in its philosophical usages is best seen as a theoretical term, with its meaning defined by its theoretical role. But in the light of this, the notion of supervenience comes to look rather suspicious. It is such an extraordinarily convenient notion for philosophers, and yet they seem simply to have made it up precisely to suit their purposes. I think that this makes the suggestion that supervenience is just some brute metaphysical fact altogether too easy.

So the challenge remains in place: just what is the real

relation of supervenience which is alleged by the supervenience physicalist to hold between real physical properties on the one hand and real special scientific properties on the other, and which gives an appropriately realist account of the assertibility of claims of supervenience?

One way of responding to the challenge just laid down is to adopt Lewisian modal realism (see Lewis 1986b). The language of possible worlds features prominently in the formulation of (a), which states an extravagant modal claim ; so if we take possible worlds to be real, as real as the actual world, then we have a neat account of the assertibility of claims of global supervenience: such claims simply record the brute facts about how things are in certain possible worlds, and possible worlds are comparatively harmless because we have always believed in ways the world could have been, and that is all, according to Lewis, that possible worlds are.

But this response, of course, is only as good as the modal realism on which it depends, and modal realism is open to a standard objection. The objection is that since, by hypothesis, possible worlds other than our own are causally isolated from our world, it is very hard to see how we could possibly learn of their existence. For it seems confirmed by all our experience that we can only get to know about entities

of a given kind if they play a part in causing effects either directly upon us or indirectly via the effects they exert upon other things which in turn exert effects directly upon us; and possible worlds are claimed to be large objects of the same sort as our own world. This objection does not depend upon the acceptance of a full-blown causal theory of knowledge, but only, as I have indicated, upon a sketch of a plausible naturalistic theory of how we know about the existence of kinds of entity. Nor, accordingly, does the objection consist in an unreasonable demand that modal knowledge be supplied with foundations; for all that is being requested is some naturalistic story about how we get to know about possible worlds which is roughly as plausible as the story we can tell about how we get to know of tables and electrons.

Lewis essays a reply to this objection (see Lewis 1986b, pp.108-115). He thinks "it is true that causal acquaintance is required for some sorts of knowledge but not for others" (p.111). So, he adds: "If you think that all knowledge requires causal acquaintance with the subject matter, I think that is just hasty generalisation" (p.112). But although Lewis here says what he clearly has to say, in order to save his position, what he says is not plausible. For it is hardly fair to call the claim that all knowledge requires causal acquaintance with

the subject matter a "hasty generalisation" if no account is offered, and none seems possible, of how we do acquire knowledge of objects with which we cannot be causally acquainted. Lewis declines to offer such an account, but instead pretends to misunderstand what is being asked of him (pp.113-5). He makes some remarks about how we acquire modal knowledge (pp.113-4), but in doing so he sets himself too easy a problem. For what we want to know is not merely how we acquire modal knowledge, but how we acquire modal knowledge <u>if</u> modal claims <u>are</u>, as Lewis holds, descriptions of causally inaccessible possible worlds. This harder question is left unanswered.

Lewis offers a second line of reply which might be thought to show that knowledge of causally inaccessible objects must be <u>possible</u>, even if we cannot as yet say <u>how</u> it is possible. His claim is that we have knowledge of <u>mathematics</u>, despite our lack of causal interaction with the subject matter of mathematics (pp.109-10). And in case you suspect that if mathematics is Lewis's life-raft, then he just pulls it down with him, he says this: "Our knowledge of mathematics is ever so much more secure than our knowledge of the epistemology that seeks to cast doubt on mathematics" (p.109).

But I find this unconvincing, since the doubt I have about

the possibility of our knowledge of causally inaccessible possible worlds carries over to mathematics, if mathematics is conceived as being about a causally inaccessible subject matter; for the same challenge can then be made against mathematical knowledge. Moreover, it is far from obvious that it is our knowledge of mathematics which is so secure. For although, like everyone else, I am as sure as anything that 2+2=4, I am by no means so certain that what I thereby have is an item of knowledge. I am of course, quite certain that in response to the question, "What is 2+2?", the proper thing to say in reply is "4". But there are other ways of construing the assertibility of mathematical claims than by taking them to state potential truths about a mathematical reality, and our certainty about how to answer the question, "What is 2+2?", cannot by itself rule those other ways out. (A way for which I have some sympathy is Hartry Field's instrumentalism; see Field 1980.) So the epistemic objection to modal realism stands undefeated, and claims of global supervenience cannot be taken to be simple descriptions of the brute facts about possible worlds.

3. The Grand-Property Hypothesis

If the relation of supervenience, then, is not to be a metaphysical mystery, some other explanation than that of modal

realism must be given of its nature. Such an explanation has been proposed by Peter Forrest, and he calls it the Grand-Property Hypothesis (see Forrest 1988a).

Grand-properties are properties of properties, rather in the way in which grand-parents are parents of parents. Their role in explaining the nature of the relation of supervenience is set out non-technically by Forrest, who uses the example of "the supervenience of an object's (or state of affairs') being intrinsically good on its naturalistic properties" (Forrest 1988a, p.2). He writes:

> Let us therefore think of <u>intrinsic goodness</u> as a property of the object's <u>nature</u>. Here, by a nature I mean the naturalistic nature, composed of the object's naturalistic properties.....Now consider two objects with the same naturalistic properties and hence the same nature. If one has an intrinically good nature, so must the other. For their natures are identical. This explains the supervenience of the object's being intrinsically good on its naturalistic properties. Yet there is no reduction.....A crucial step in this argument is that when we say two objects have the <u>same</u> naturalistic property we mean it. The properties are identical. Hence they are to be thought of as universals. (p.2)

If supervenient properties are grand-properties, that is, properties had by properties, and if properties are universals, and hence literally present in every exemplification of the property, then two objects with the same naturalistic

properties cannot fail to have the same grand-properties: the Indiscernibility of Identicals suffices to ensure this result (see Forrest 1988a, p.8). Moreover, to the extent that we may assume it to be a contingent truth that a given naturalistic property has a given grand-property, there will be possible worlds wherein an object has the naturalistic property but simply lacks the supervenient property. Thus, though Forrest does not point this out, the Grand-Property Hypothesis has the resources to explain why the presence of a supervenient property is not <u>entailed</u> by the presence of the subvenient property on which it supervenes.

Now the Grand-Property Hypothesis is intended by its author to provide an explanation of cases of supervenience where the properties linked by supervenience are properties had by a single substance. These cases are divided by Forrest into two kinds, which he calls Narrow Supervenience and Broad Supervenience (p.5). Accordingly, he does not explicitly address the question of whether it can also serve to explain claims of global supervenience, such as a supervenience physicalist must make. I shall now argue that it cannot thus serve. The reason why it cannot is that the necessary appeal to the Indiscernibility of Identicals is illegitimate in the global case. As it happens, the objection I shall raise

applies also to Forrest's attempt to use the Grand-Property Hypothesis to explain Broad Supervenience, though I shall not stop to show this. In his discussion of Broad Supervenience, he does not address my objection (p.8).

If we focus attention upon the simplest case, that of Narrow Supervenience, then all is well. Suppose two objects, x and y, have the same subvenient property, P. Then, given that P is a universal, there is literally just one property in which both x and y participate. So if P has a certain grand-property, it will have it in connection with both x and y, i.e. both x and y will have the supervenient property. And so in any world in which x and y have P, either they will both have the supervenient property, the grand-property, or they will neither of them have it. (The worlds in which neither of them has it are worlds in which, I take it, P simply lacks the relevant grand-property.) But there is no world in which both x and y have P but one of x and y has the supervenient property while the other does not.

Let us turn now, however, to cases of global supervenience. A simplified, though adequate, statement of a claim of global supervenience might be this: no two worlds differ in respect of their special scientific properties without differing in respect of their physical properties. By "their special

scientific properties" and "their physical properties" I mean the special scientific and physical properties instantiated in those worlds, rather than had by them. (So the physical properties involved can perfectly well be physical in the sense I earlier sketched.) Now, even though in such cases we cannot rely upon the existence of any single objects to be the bearers of both supervenient and subvenient properties, we may still invoke the idea of grand-properties. For we will be able to advance the hypothesis that, very roughly, the special scientific properties are grand-properties, that is, properties had by the (in this case) physical properties. Indeed, there is this advantage in the global case, that we do not have to claim, as we do in the simplest case, that the supervenient property apparently of an object is strictly speaking not a property of the object at all, but rather a property of some property of the object.

But there is a major problem. It is that a claim of global supervenience involves quantification over distinct possible worlds, which raises a difficulty for the would-be explanation of the supervenience claim. For how is this explanation to proceed? It is tempting to reply that two worlds alike in all their physical properties must, since properties are universals, participate in the very same properties, so that,

by the Indiscernibility of Identicals, if one world involves special scientific properties, then so must the other. But this is an illegitimate application of the principle of the Indiscernibility of Identicals. For the principle simply does not apply across possible worlds, i.e. to individuals in different possible worlds. Let me illustrate this point. Nixon in this world must of course share all his properties with Nixon in this world, i.e. with himself. But Nixon may well be identical with Nixon in another possible world. If so, then Nixon in that other world may in that other world never have been a US President: the two Nixons, in different worlds, need not have all their properties in common. And what goes for particulars, such as Nixon, will go for universals also. The physical universals of our world may well be identical with the physical universals of another possible world; but they may differ with regard to the properties they have, that is, with regard to their grand-properties, in this case special scientific ones. Since a universal may have a grand-property in one world but lack it in another, the Grand-Property Hypothesis does not seem to be able to explain why two physically indiscernible worlds must be indiscernible special scientifically.

This argument could be resisted by claiming that physical

universals have all their properties essentially, so that if two worlds contained the same physical universal, the universal in each world would have to have the same grand-properties. This would, of course, rule out the situation in which a given universal has a certain grand-property in one world but lacks it in another. But this response is doubly objectionable. First, the supposition that either grand-properties in general or special scientific grand-properties in particular are essentially had by the properties which have them seems entirely ad hoc. It would need independent motivation if it were to be used in reply to my argument. Secondly, the supposition has the unwelcome consequence that the presence of the subvenient properties would entail the presence of the supervenient properties. For, given the essentialist claim, the physical properties will have their grand-properties, i.e. the special scientific properties, in every world in which they exist, which will secure the relevant (broadly logical) entailment. In the case of supervenience physicalism, this would mean (roughly) that every world physically indiscernible from our own world would instantiate every special scientific property instantiated in our world. This consequence is unwelcome because it runs counter to our intuition that there is a possible world indiscernible physically from our world, but lacking some of the special scientific properties, perhaps

one or two instances of a mental property, instantiated in our world (see Petrie 1987, p.120). If traditional dualism is so much as coherent, then this must be a possibility. The supervenience physicalist wants the physical properties to fix the special scientific properties in our world, but in some other worlds surely the latter must be permitted to come and go as they please. So the Grand-Property Hypothesis, even when coupled with the suggestion that grand-properties are had essentially, cannot serve to explain the supervenience physicalist's claims of global supervenience.

4. The Best Explanation of Claims of Supervenience

Neither modal realism nor the Grand-Property Hypothesis can provide an explication of the nature of the alleged real relation of supervenience, understood as holding between real properties and as enabling the physical to fix or determine the non-physical. I shall now propose a new explanation of the assertibility of claims of global supervenience, an explanation which does not even try to construe supervenience in the realist way just described. Since it is, I believe, the only such explanation, it is by default the best explanation, and hence should be accepted by anyone who endorses claims of global supervenience. But this explanation of claim (a) of supervenience physicalism is radically subversive of that

doctrine, since it involves a denial of the reality of those special scientific properties which cannot be identified with physical properties; that is, a denial of claim (b) of supervenience physicalism. This is why I said earlier that supervenience physicalism is at war with itself.

My proposed explanation of the assertibility of claims of global supervenience rests upon an <u>error theory</u> of our ascription to the world of those special scientific properties which cannot be identified with physical properties. (For the idea of an error theory, see Mackie 1977, Ch.1, especially pp.35 and 41.) That is to say, while we ordinarily take our ascription of such properties to be true and accurate, we do so erroneously, since in reality there are no such properties. We project these non-existent properties onto the world as part of our attempt to explain it: the special sciences are ways of thinking of the world which aim to explain it, but the explanatory theories constituting them are false.

These non-existent special scientific properties are projected onto -- that is, offered as part of an explanation of -- a real world, constituted entirely of physical properties; it is in response to physical stimuli (the only stimuli there are!) that the special sciences are fashioned. And in postulating special scientific properties as part of an

explanation, we are committed to taking them as real. (So there is no scope here for Blackburn's quasi-realism; see his 1984, pp.171 and 180.) Now, since this explanatory response is not an arbitrary matter, but instead occurs systematically, it would occur in the same way in any world in which the physical stimuli were the same, that is, in any world physically indiscernible from our own. Moreover, this can be generalised: any two physically indiscernible worlds, that is, any two worlds where the physical stimuli are the same, will inspire the same explanatory response, that is, will inspire the erroneous projection of the same non-existent special scientific properties onto the real world of physics. But if this is so, then any two worlds which are physically indiscernible will be special scientifically indiscernible also; and this last claim is equivalent (or nearly enough so) to claim (a) of supervenience physicalism interpreted, as it must be, as a claim of global supervenience. Hence, the error theory of special scientific properties not identifiable with physical properties can explain the assertibility of the relevant claims of global supervenience. These claims are assertible, not in virtue of the holding, between real properties, of a real (though obscure) relation of supervenience, but in virtue of a certain systematicity in our explanatory efforts. We need no longer worry about how exactly

it is that the physical properties fix or determine, while at the same time giving some free rein to, the special scientific properties, since on the error theory these special scientific properties are not real, and hence cannot stand in any real relation to the physical properties.

The error theory enjoys two other advantages, which I shall now mention. The first is that, even if the Grand-Property Hypothesis constituted any explanation at all of global supervenience, and I have argued that it does not, it would still constitute an inferior explanation to that provided by the error theory. This is because it involves so many more properties than does the error theory, namely all the requisite grand-properties, and is therefore much less simple than its rival.

The second advantage enjoyed by the error theory is that it removes a mystery which would otherwise afflict supervenience physicalism. The mystery concerns the causal efficacy of the special scientific, and hence supervenient, properties, and it is best brought out by presenting a dilemma for supervenience physicalism. Either the supervenient properties are causally efficacious, or they are not. If they are, then since the subvenient physical properties are supposed to fix or determine the supervenient states of the world, those supervenient states

must be overdetermined, being determined both by earlier supervenient states and by contemporaneous physical states, and it must be a happy, but cosmic, coincidence that these independent determinations work in harmony and do not instead pull in conflicting directions. If, on the other hand, the supervenient properties are not causally efficacious, then it is something of a mystery how we get to know of their existence, since they will be capable of playing no part, either directly or indirectly, in the causal affecting of our sensory apparatus.

Thus, the dilemma, and the ensuing mystery. The mystery is avoidable by taking the first horn of the dilemma, though I doubt that this will be thought of as a very attractive option. So if the mystery must be confronted, what is to be said about it? The error theory can make light work of it. For if, as the error theory maintains, the supervenient special scientific properties are projected and not real, there can be no problem about how we get to know about them, since there is nothing to get to know about.

5. Conclusion

I conclude, then, that supervenience physicalism is untenable because it involves incompatible commitments. Its crucial

claim (a) stands in need of philosophical elucidation, and the error theory, which alone can supply such elucidation, entails that claim (b) of supervenience physicalism is false. In effect, my arguments show, if they are correct, that the supervenience physicalist is committed to the view I call eliminative type physicalism. This is a bad position for the supervenience physicalist to be in. For either eliminative type physicalism is false or, as I believe, true. If the eliminative type physicalism to which he is committed is false, then he is committed to a falsehood and had better abandon supervenience physicalism. But if, on the other hand, eliminative type physicalism is true, then, since eliminative type physicalism is incompatible with supervenience physicalism (they differ over the reality of the special scientific properties), he is committed to a contradiction.

CHAPTER THREE

Problems For Pluralism

I have argued so far that neither token physicalism nor supervenience physicalism constitutes an acceptable account of the relations between the many systems of discourse which form the many sciences. In this chapter I shall expound and then criticise an account of these relations which is not a form of physicalism at all, and which I propose to call pluralism.

I need to begin, however, by stating what the doctrine of pluralism is. Towards this statement I shall gradually work. My contention will be that pluralism, as I define it, is the position that one is compelled to adopt if one wishes to avoid being a physicalist and takes to heart the arguments against token physicalism and against supervenience physicalism which I have offered in the preceding two chapters. Pluralism, I shall contend, is the only non-sceptical or non-agnostic response to the problem of the many sciences available to the would-be non-physicalist.

If both token physicalism and supervenience physicalism are, as I have argued, inadequate responses to the problem of the

many sciences, what alternative responses are possible?

One class of responses would consist in doctrines modelled on the forms of physicalism I have already considered but assigning the privileged position not to physics but to some other of the many sciences instead. Such a doctrine would be token S-ism, or supervenience S-ism, for some science, S, distinct from physics. Thus we could conceive of a thesis of, for example, token economism, according to which every event was identical with some economic event, or supervenience biologism, according to which all facts, of whatever type, supervened upon the biological facts. But although we could conceive of such doctrines, they are, I take it, very implausible. I shall not consider them further.

A second class of possible alternative responses to the problem of the many sciences would be what we may call hybrid responses. A hybrid response is any response which combines elements of token physicalism, supervenience physicalism or even non-eliminative type physicalism in its account of the relations between the many sciences and their respective domains. To see how this is possible, consider again for a moment the thesis of token physicalism. Token physicalism asserts that every event, for instance every biological event, is identical with some physical event, but it explicitly denies

that every physical event is identical with some biological event. Token physicalism thus asserts the holding of an asymmetrical relation between physics and each of the special sciences, a relation which has the effect of privileging physics. We can call this relation the token identity relation. Consider now the thesis of supervenience physicalism. Supervenience physicalism asserts that all facts, for instance all biological facts, supervene upon the physical facts, but it explicitly denies that all the physical facts supervene upon the biological facts. Supervenience physicalism, therefore, like token physicalism, also asserts the holding of an asymmetrical relation between physics and each of the special sciences, and this relation, too, has the effect of privileging physics. Let us call the relation asserted by supervenience physicalism to hold between physics and each of the special sciences the supervenience relation. Now, what I have called hybrid responses to the problem of the many sciences are possible because one could assert the holding of either the token identity relation or the supervenience relation between any two sciences. Here is an example of a bizarre but possible hybrid response, devised for purposes of illustration only: every social scientific event is identical with some psychological event (though not vice versa); and all psychological facts supervene upon neurobiological facts (but

not vice versa), even though psychological-neurobiological token identities are not assertible; and all neurobiological properties are identical with physical properties.

Hybrid positions, then, constitute the second class of possible alternative responses to the problem of the many sciences. But although hybrid positions are possible, they are not acceptable. Here is why. My earlier arguments against token physicalism and supervenience physicalism did not depend upon its being physics which was the beneficiary of the asymmetrical relations asserted to hold by those doctrines. Nor did those arguments depend either upon the idea that, in token physicalism, it is every non-physical science which is said to stand in the asymmetrical token identity relation to physics or upon the idea that, in supervenience physicalism, it is every kind of fact which is said to supervene upon the physical facts. It follows from these features of my arguments against token physicalism and against supervenience physicalism that those arguments will have equal force against any hybrid view which asserts the holding of the token identity relation or the supervenience relation between any two sciences. If a hybrid view asserts the holding of the token identity relation between, for instance, social scientific events and psychological events, then the view will be liable to the

Argument From Overdetermination and the Epistemological Argument. If a hybrid view asserts the holding of the supervenience relation between psychological facts and neurobiological facts, then the view involves the unacceptable mystery which in Chapter Two I showed to be entailed by supervenience physicalism. So hybrid positions cannot be accepted as responses to the problem of the many sciences, responses alternative to token physicalism and supervenience physicalism.

A third class of possible alternatives to token physicalism and supervenience physicalism contains as its star member my own favoured view, eliminative type physicalism, which asserts the exclusive existence of the physical (see Chapter Four and Part 2, passim). As a brand of physicalism, this view can hardly be adopted by the pluralist, who is attempting, of course, to occupy a non-physicalist position. But although eliminative type physicalism is the star member of this class, it is not the only member. For one could in principle embrace eliminative type S-ism, where 'S' stands for some science other than physics. The only example of such a view which I have ever seen defended is what we could call eliminative type mentalism, which asserts the exclusive existence of the mental. This view is better known as idealism. Although

idealism has had, and continues to have, distinguished defenders, my arguments in this thesis are directed mainly at philosophers who take the sciences very seriously, and who are therefore unlikely to regard idealism as a live option. Accordingly, I shall not argue explicitly against it. My arguments in Chapter Five against external world scepticism are, in effect, my reasons for rejecting idealism, but they are not presented as such.

Let me briefly take stock before proceeding. I have been searching for responses to the problem of the many sciences which are (a) distinct from either token physicalism or supervenience physicalism or eliminative type physicalism, (b) immune to the arguments presented in Chapters One and Two and (c) at all likely to find champions. No response so far considered has satisfied all these requirements. (I shall be referring to these requirements again.) But we can describe a response which does, and that response is pluralism.

What, then, is the thesis of pluralism? It is, in essence, a claim of ontological egalitarianism. No one science is privileged over any other science by its standing in some asymmetrical relation, such as the token identity relation or the supervenience relation, to that other science. Hence pluralism asserts that the many sciences constitute an

irreducible plurality of vocabularies each of which supplies a more or less adequate description of one of an irreducible plurality of domains: to put it dramatically, the different sciences describe different worlds. (I shall use 'worlds' and 'domains' interchangeably.) Pluralism is the positive view which might be offered by the opponent of physicalism who cannot see any reason to think that physics, and physics alone, gives us insight into the nature of reality, and who suspects that many other sciences, and perhaps all of them, also yield such insight.

Can we say more about the nature of pluralism? We can. And indeed we must. For if pluralism is to count as a possible response to the problem of the many sciences which satisfies the requirements (a), (b) and (c) listed above, it must have another feature, which will prove to be crucial. On pain of failing to qualify as a possible response to the problem of the many sciences which satisfies requirements (a), (b) and (c), pluralism must <u>deny that there is any causal interaction between the many worlds which pluralism supposes the many sciences to describe</u>. There may be casual interaction <u>within</u> a world, but not <u>between</u> worlds. Let me now present an argument to show why pluralism must make this -- apparently absurd -- denial.

For the purposes of this argument, let us consider an artifically simplified version of pluralism, a version according to which there are exactly two sciences, and hence exactly two worlds, one world for each science to describe. Let us label these two sciences S1 and S2. The argument has the form of a reductio ad absurdum, and the assumption to be reduced to absurdity is the assumption that there is interworld causation, i.e. the assumption that S1-events cause S2-events (or vice-versa).

Let us suppose, then, that a certain S1-event, c, causes a certain S2-event, e. Then, in accordance with the widely accepted principle that occurrences of singular causation must be backed by a causal law, there must be some causal law backing up the causation of e by c. There are exactly three possibilities for the needed law here. Either (i) there is a causal law linking c, qua S1-event, with e, qua S2-event; or (ii) there is a causal law linking c, qua S1-event, with e, qua S1-event; or (iii) there is a causal law linking c, qua S2-event, with e, qua S2-event. However, none of these three possibilities is open to the pluralist, given that pluralism is to be a response to the problem of the many sciences which satisfies requirements (a), (b) and (c). Let us consider each possibility in turn.

If possibility (i) were adopted, then it would have to be the case that there were laws linking S1-events, qua S1-events, with S2-events, qua S2-events. But such laws, as far as we know, do not in general exist. There seem to be no causal laws linking biological events, qua biological events, with microphysical events, qua microphysical events (e.g. no causal laws linking kidney-events as such with neutrino-events as such); there seem to be no causal laws linking economic events, qua economic events, with microphysical events, qua microphysical events (e.g. no causal laws linking increases in the money supply as such with emissions of electrons as such); and so forth. Indeed, it is this apparent fact -- what Jaegwon Kim called nomological incommensurability (see Kim 1978) -- which gives rise to the problem of the many sciences in the first place (see my Introduction). So pluralism cannot embrace possibility (i), on pain of failing to address a key component of the problem of the many sciences.

Possibility (ii) claims that c causes e in virtue of a causal law linking c, qua S1-event, with e, also qua S1-event. But e is by hypothesis an S2-event. So if e is to be an S1-event as well, it must be that an S2-event is identical with an S1-event. It follows from this that a pluralist who adopts possibility (ii) must assert of every S2-event which is caused

by an S1-event that it (the S2-event) is identical with some S1-event. But notice what the pluralist has thereby done: the pluralist has affirmed the holding of the asymmetrical token identity relation between S1 and S2. Such an affirmation is, however, as I argued a few paragraphs above, liable to the arguments against token physicalism which I presented in Chapter One. Since it is a condition of adequacy on pluralism that it satisfy requirement (b) -- the requirement that pluralism be immune to arguments presented in Chapters One and Two -- pluralism combined with possibility (ii) must be rejected. Possibility (ii) is useless to the pluralist.

The final possibility, possibility (iii), claims that c causes e in virtue of a causal law linking c, qua S2-event, with e, qua S2-event. But c is by hypothesis an S1-event. Exactly the same problem therefore arises for possibility (iii) as arose for possibility (ii). The only difference is that in this case the asymmetrical token identity relation which the pluralist is obliged to affirm privileges S2 over S1, rather than the other way around. The problem, of course, is that pluralism combined with possibility (iii) also falls foul of requirement (b). So possibility (iii) is also useless to the pluralist.

I can now complete my argument for the claim that pluralism,

if it is to be a response to the problem of the many sciences which satisfies requirements (a), (b) and (c), cannot permit interworld causation. If an S1-event causes an S2-event, then one of the three possibilities, (i), (ii) or (iii), must obtain. But the pluralist cannot allow <u>any</u> of these possibilities. So the pluralist must deny that S1-events ever cause S2-events. And this argument, though based on a simplified form of pluralism involving only two worlds, can easily be seen to be generalisable, so as to sustain a general conclusion: any pluralist must prohibit all interworld causation.

This result, that pluralism must prohibit interworld causal interaction, is sufficiently striking, and sufficiently surprising, to merit three supplementary comments.

First, in including in my characterisation of pluralism the requirement that there be no causal interaction between the many worlds of pluralism I may be accused of having burdened pluralism with an unnecessary commitment. But this is not the case. For, as my preceding arguments in this chapter are intended to show, the commitment is necessitated by the theoretical role that pluralism must play. That role, of course, is that of constituting a response to the problem of the many sciences, a response which satisfies requirements (a),

(b) and (c). It may appear as if there could be a version of pluralism which avoided this commitment, but if my preceding arguments in this chapter are correct, that appearance is illusory.

Secondly, the ban on interworld causal interaction, to which pluralism is committed, may be said to be unrealistic. For, after all, is not the world full of interworld causation? If I kick a table, electrons are moved; if I believe the dog is dangerous, my pulse rate will rise; if I remove all the carbon atoms from a room, I destroy all the plants and animals in the room. All these examples are cases of interworld causation. So if the pluralist has to deny interworld causation, pluralism is thereby shown to be a view unworthy of discussion.

I sympathise with this response, but I do not think that it shows pluralism to be unworthy of discussion, for the following reason. Certainly interworld causation <u>appears</u> to be a pervasive feature of the world. But the pluralist can try to argue that this appearance is deceptive. Taking a leaf from the book of the <u>parallelist</u>, who maintains that there is no causal interaction between mental events and physical events, appearances to the contrary notwithstanding, the pluralist can say that because the many worlds run in parallel, there are many observable regularities between events in one world and

events in other worlds; because there are these regularities, we mistakenly suppose that there are causal relations between events in different worlds; but in truth there are no such causal relations. This reply which I have put into the mouth of the pluralist may be held to be unconvincing in the end; but it does suffice, I think, to reinstate pluralism as a view worthy of discussion. Of course, since I am not a pluralist, I need not defend pluralism further. If the pluralist's ban on interworld causation is enough all by itself to discredit pluralism, then the arguments which I shall present against it in the balance of this chapter may be judged superfluous; but the cause of eliminative type physicalism will not have been harmed.

My third comment is this. Pluralism is an anti-physicalist position. Indeed, I have argued that it is the only anti-physicalist position which is immune to arguments presented in Chapters One and Two and which is at all likely to find champions. Traditionally, however, it has seemed natural to set in opposition to physicalism the doctrine of dualism. But pluralism is misunderstood if it is thought of as a mere generalisation of dualism. For dualists, or at least Cartesian dualists, have always wanted to affirm the existence of causal connections between the two worlds of the physical and the

mental; whereas the many worlds of pluralism, as I have argued, are not linked together causally at all. Traditional dualism has therefore always been monistic to the extent that it regarded everything in the world as belonging to a single causal network. (It is interesting to note that Jaegwon Kim, in his recent Presidential Address to the American Philosophical Association, notes the possibility of a non-traditional form of mind-body dualism which, like pluralism, prohibits causal interaction between the the two domains of the mental and the physical. See Kim 1989a, esp. p.47.)

Let me now attempt a summary description of the character of pluralism. The irreducible plurality of pluralism's many sciences and correspondingly many worlds can be expressed in the following four claims. First, there is no causal interaction between objects belonging to different worlds. This is the surprising ban on interworld causal interaction to which I have argued that pluralism is committed. Secondly, no property belonging to one domain is identical with a property belonging to another domain. This requirement is included to save pluralism from having to assert obviously false claims of type (or property) identity such as non-eliminative type physicalism, implausibly, maintains. Thirdly, the many

sciences are not united by their supplying explanations and descriptions, albeit at different levels, of a single world of events. This is the requirement that the token identity relation not be asserted to hold between any two sciences. Finally, it is not the case that one particular science, for example physics, is such that the facts describable with it fix or determine, in the manner of supervenience physicalism, some or all of the other facts. This is the requirement that the supervenience relation not be asserted to hold between any two sciences.

In the light of this summary description, it is clear that, as I remarked in the Introduction, pluralism is not to be confused with what might be called <u>regional specialism</u>, which holds that the one world falls into distinct spatio-temporal parts and that each science deals with its own part. Regional specialism is perhaps the view that most naturally springs to mind in those who consider for the first time what are the relations between the many sciences. But regional specialism is not the same as pluralism, because if the many regions are to be thought of as spatio-temporal parts of one world, then we should expect there to be real connections of various kinds -- perhaps causal or spatial -- linking the regions one to another so as to form a single world, while according to pluralism

there are no such causal connections, and there need be no such spatial ones.

Is anyone a pluralist? It would not matter if the answer were "No-one", since, as I have been arguing, pluralism is a possible position to occupy, and if one seeks a viable anti-physicalist position, the only position to occupy. But as it happens there are pluralists -- or at least people who come extremely close. There is perhaps a hint of it in Ryle's discussion of the compatibility of common sense and physics, though there are other, and conflicting, hints (Ryle 1954, p.79). And much more recently Strawson has suggested, in connection with a variety of apparent incompatibilities, that we must relativise our talk of reality to various incommensurable standpoints (Strawson 1979; and 1985, especially pp.37-8). There is also an explicit, though highly indirect, defence of pluralism provided by John Dupre (see Dupre 1988). And a chemist, Hans Primas, has produced a remarkable account whose main thrust is, I think, pluralist (see Primas 1981). But by far the closest approximation to what I call pluralism is defended in the writings of Nelson Goodman and, recently, Hilary Putnam. It is on these treatments, therefore, that I shall mainly concentrate.

1. An Initial Problem For Pluralism

Let me begin, however, by raising a problem which will afflict any kind of pluralism. Any kind of pluralism, if it is not to be a mere exercise in wishful thinking, must include some account of how the many sciences are to be individuated; and this account must be rich enough to enable us to decide to which science a given set of claims properly belongs. Without such an account, pluralism is powerless to resolve any ontological controversies. For example, suppose that we are interested in whether or not there is a Judaeo-Christian God. If talk about God properly belongs to the same system of discourse as that which includes physics, then we may well see physics and theism as rivals and then be led to deny God's existence. But if talk of God can be permitted to belong to, or even to constitute, a system of discourse of its own, from which physics is excluded, then the threat to the rationality of theistic belief posed by physics is neutralised. And, quite generally, unless the pluralist offers the required account, he will be in the absurd position of being able to vindicate any ontology he likes by simply stipulating that the apparent descriptions of it form a science of their own.

But the pluralist's account of the individuation of

vocabularies must steer a narrow course between two extremes. On the one hand, if vocabularies are too <u>finely</u> individuated, then pluralism could justify the acceptance of discredited theories, together with their ontologies. For instance, one could endorse both Newtonian mechanics and relativistic mechanics by alleging that each theory belongs to a separate system of discourse and hence deals with a distinct domain of objects, the objects described by the former theory having just (e.g.) mass, while those described by the latter have relativistic mass. This path could even lead to the complete elimination of inter-theoretic rivalry. On the other hand, if the vocabularies are too <u>coarsely</u> individuated, then two theories, or even two sciences, may become potential rivals. For instance, one might regard physics as basically continuous with common sense, both being attempts to understand the empirical world, so that they belong to the <u>same</u> system of discourse; but then the question could arise of whether they are <u>rival</u> attempts, and if it was answered affirmatively we would lose either the physical or the commonsensical ontology.

My point here is not that the problem of individuating the vocabularies cannot be solved, but that it must at least be addressed, and plausibly, if pluralism is to avoid the pitfalls I have indicated. I shall not discuss the problem further.

(Gerald Vision raises a closely-related problem for a different view in Vision 1988, pp.226-8.)

2. Two Pluralist Tendencies

In defending pluralism, both Goodman and Putnam represent it as made possible by certain general views in metaphysics and epistemology, views which may fairly be described as being anti-realist and anti-empiricist (by which I mean opposed to knowledge empiricism); and I shall in my fourth and fifth sections be examining and rejecting their defences of these views. But I want now to argue that this association between pluralism and both anti-realism and anti-empiricism is not, as it were, accidental, but is very likely to represent a commitment that pluralism incurs. Yet this commitment is in fact a liability, since most philosophers would certainly like to be both realists and, in the relevant sense, empiricists if they could be. So if pluralism requires the abandonment of realism and empiricism, it must be for good reasons. But then, if the Goodman and Putnam defences are unsuccessful, as I shall argue that they are, and thus do not supply such good reasons, pluralism is just left with its liability and deserves to be rejected. (The idea that physicalism depends on "categorial realism", so that by abandoning it we are free to adopt pluralism, is suggested also in Dupre 1988, pp.51-52; and there

are anti-realist strains in the pluralist-tending account of Primas 1981, e.g. pp.30-1.)

Why does pluralism lead to anti-realism, that is, to the denial of the idea that the non-mental things that exist do so independently of our epistemic and conceptual capacities? For the sake of simplicity, let us concentrate on an imaginary case where we have just two systems of discourse. One system entails that there are F's, whereas the other entails that there are G's. F's and G's are to be thought of as being, in an intuitive sense, different kinds of things. (Perhaps they are tables and neutrinos.) The first point to notice is that the two existence claims made by the different systems of discourse are not formally inconsistent, even when they refer to the same time; for there could easily be a world in which, simultaneously, there were both F's and G's. We could manufacture an inconsistency, of course, by making one system entail that there are F's but no G's, and the other entail that there are G's but no F's; but I do not see why each system cannot simply be <u>silent</u> about the other system's ontology.

Something genuinely alarming about the simplified pluralism under consideration is that, according to pluralism, we must not expect to find causal connections between F's and G's: F's and G's do not belong to a single causal network. Also, it is

unclear whether the pluralist can regard F's and G's as being located in the same space. It is unclear for two reasons. First, the same volume of space may contain both F's and G's, without it being the case that the F's fit in the gaps left vacant by the G's (or vice-versa); for instance, a table-shaped space may contain both a table and neutrinos. This might lead one to say that the F's and the G's occupy different spaces. Secondly, it will not always be the case that it will make obvious sense to ask what is the position of a particular F relative to a particular G; let the F's be beliefs or high rates of inflation.

Now it might be said that there is still nothing here to prevent a pluralist from being a realist. For there might be many spaces, and each space might be as mind-independent as any realist could hope for. Moreover, the idea of their being more than one causally-isolated system has even been advanced by realists to illustrate the possibility of radically verification-transcendent facts.

But the problem for an attempt to combine pluralism and realism does not stem solely from the fact that pluralism is committed to the existence of causally-isolated systems and, perhaps, more than one space. It arises from this, together with the reasonable pluralist assumption that we humans somehow

get to know about the objects forming these causally-isolated systems; for it is not as if the pluralist takes his many worlds (or domains) to lie forever beyond our ken. And this prompts the following awkward question: to which world do we humans belong? If we reply that we belong to only one world, then the causal isolation of the many worlds would prevent the assumed knowledge. For we could not infer conclusions about the other worlds in order to explain the causal effects of those worlds on our world, since there would be no such inter-world causal effects. Nor could our beliefs about the other worlds be reliably formed on the basis of some reliable causal connection involving the other worlds, for the same reason.

Alternatively, then, we might answer the awkward question by suggesting that we humans somehow straddle many worlds, belonging to each either entirely or partially. But both forms of this suggestion are absurd. The first would have us existing entirely in many, causally distinct worlds, which is a notion that is hard even to understand, still less believe. The second form of the suggestion would have us existing partially in many, causally distinct worlds, in some kind of fragmented fashion. But then, especially in view of the causal distinctness of the many worlds, it is hard to see in what our

identities could possibly consist.

What are we to say, then, to the awkward question? The place of knowing humans in a pluralist reality becomes much easier to understand if we think of the many worlds, not as constituting a mind-independent framework into which humans must be fitted, and perhaps squeezed, but as being in some sense the <u>products</u> of the human mind; for if humans have, in some suitably non-causal sense, <u>created</u> the many worlds, then it is unsurprising either that humans should be able to get to know about them without benefit of causal interaction with them, or that inter-world causation should conveniently work to human cognitive advantage by occasionally occurring extra-systemically. But of course to think in this different way is to think in an anti-realist way. It is to abandon the realist idea that non-mental reality is at most <u>causally</u> dependent upon the mental. And that is why it is very likely that pluralism is committed to anti-realism.

Why does pluralism lead to -- I do not say entail -- anti-empiricism? Empiricism, as I am understanding it, is simply the idea that, notoriously difficult though it is to articulate convincingly, there is an important sense in which sensory experience is the touchstone for all our knowledge, including, of course, our knowledge of what there is. Why is a

pluralist under pressure to reject this idea? The reason is that pluralism must ensure that its many worlds are, as it were, epistemically insulated from one another; and when this is done, empiricism must be repudiated.

Pluralism's many worlds must be epistemically insulated from one another in just this sense: it must never be the case that types of entity from different worlds are justifiably believed to exist on the basis of the same sorts of epistemic criteria. The reason why this must be so is that, if there were a single sort of epistemic criterion relevant to the assessment of ontic claims in respect of different worlds, then it might happen that only the entities belonging to one of the worlds should satisfy the criterion, so that we would lose reason to believe in the entities of the other world. Evidently this would tend to push us away from pluralism towards some kind of monism; and it is almost equally evident that how far we would be pushed in this monistic direction would be directly proportional to the number of worlds to which some one sort of epistemic criterion was appropriate. But now, of course, the threat to pluralism posed by empiricism is clear. For if empiricism is correct, then we will have the limit case where just one sort of epistemic criterion is appropriate for all worlds, and hence the possibility of monism will be opened up. (Whether the

possibility is actual will then be the crucial question; perhaps compare Part 2, passim.)

I can think of only one way in which the pluralist might seek to guarantee the required epistemic insulation, and it entails the rejection of empiricism as I have defined it. It would be to defend the claim that for each world there is one and only one appropriate sort of epistemic criterion. Justification would in that case be relative to systems of discourse -- sciences, in my sense -- and no single criterion would be fundamental or essential to justification as such. Since it follows at once that sensory experience would not be fundamental or essential to justification as such, this pluralist gambit would necessitate the abandonment of empiricism. Such a gambit might well take the form of a modified and restricted coherentism: a belief would be justified not if it simply cohered with all other beliefs, but only if it cohered with the other beliefs belonging to the system to which it itself belongs. (It is worth remarking that, if this coherentist line is taken, and perhaps Goodman takes it (see Goodman 1978, pp.17 and 124), the individuation problem raised in my previous section becomes particularly acute.)

3. Cookie Cutting Criticised

There is a certain challenge to my contention in the last section that pluralism leads to anti-realism and anti-empiricism which I shall now set out and then reject. The challenge takes the form of the sketching of a metaphysical position which is consistent with pluralism but which nonetheless remains recognisably realist and empiricist. If this position is consistent, and is also properly describable in these terms, then it will constitute a counterexample to my contention. The position, which has never to my knowledge been fully expounded, rests heavily upon a metaphor, that of the Cookie Cutter, which is invoked quite commonly. (For some partially relevant discussion, see Wiggins 1980, Ch.5, especially pp.133-4; 136-7; 139-40.)

Here is an outline of the position. There is a world as it is in itself, and perhaps we may call it the noumenal world, which is thoroughly mind-independent. But this world, as it is in itself, is formless, undifferentiated, unclassified into kinds, and so forth, <u>until</u> the active mind imposes form, order, structure, classifications, and so forth, onto it by imposing its concepts onto it. Let us decide to call a system of concepts a <u>conceptual scheme</u>; then we can say that there can be -- indeed there are -- many different and incommensurable conceptual schemes: we can and do divide up the world in many

differing ways, in line with our many differing purposes. For example, there is the scheme of physics, and that of each of the special sciences, and doubtless there are others. So although the world as we know it and think about it is mind-dependent in the sense that it is conception-dependent, it is nevertheless mind-independent in the sense that the world, once it has been carved up by our conceptual schemes, is perception-independent, and is, quite generally, the way it is regardless of our epistemic abilities.

But how are we to understand this activity of the active mind, its imposition of concepts, its dividings up of the noumenal world? For there is evidently no literal, physical imposing or dividing -- no one is ascribing to the mind something analogous to glorified telekinetic powers. Well, this activity is to be understood by reference to the metaphor of the Cookie Cutter, as follows.

The neutral stuff which makes up the noumenal world is like shapeless cookie mixture -- we may call it dough -- and our conceptual schemes are like cookie cutters, which come in different types and which we can use to form the dough into shaped cookies. Now the dough is real enough, and once cookies have been cut they have the shape they have regardless of our ability to tell what shape that is. But of course until we get

to work with our cutters, the dough has no shape at all.

Thus the metaphysical position, and the metaphor intended to illuminate it. Let me turn now to my two reasons for rejecting it as a counterexample to my contention that pluralism leads to anti-realism and anti-empiricism.

First, the claim that the position is a realist one seems to me to be implausible, since the limited mind-independence which it assigns to such objects as stars and horses, namely perception-independence, does insufficient justice to widespread realist intuitions. It is a part of the content of those intuitions that stars and horses, if they do exist, exist quite independently, not only of perceptions, but also of any concepts or conceptions which any human or indeed anyone else might have; and in particular that stars and horses might have existed even if there had been no concepts or conceptions at all. But of course the position under consideration denies this.

Secondly, the metaphor of the Cookie Cutter, far from supplying the position with much-needed intelligibility, in fact reveals its dubious coherence. To see this, we must take the metaphor seriously and attempt to spell out its implications. Now, the metaphor certainly enables us to grasp

how the imposition of one conceptual scheme could introduce order, form, structure, and so on, into the noumenal dough. But how does it illuminate the possibility, essential to pluralism, of there being more than one, actual, alternative conceptual scheme? I cannot see how it can do this. It is clear how it illuminates the possibility of more than one, possible, alternative conceptual scheme: we need only imagine having used cookie cutters different from the ones we actually did use -- square ones, perhaps, instead of round ones. Actual, alternative conceptual schemes, however, seem to be impossible; for, as every pastry cook knows, we cannot make square and round cookies out of the same dough! It may well be that, from the point of view of the universe, square cookies are no better or no worse than round ones; but the point remains that having made square ones we must not think that we can then make round ones without destroying the first batch. How could the noumenal dough, as it is in itself or in any other way, be such as to permit itself simultaneously to be formed into both square and round cookies? Evidently there could be no literal dough of this required sort. And if it is replied that the noumenal world is not exactly like dough, then it needs to be shown how and why this reply does not amount to an admission that the noumenal world is not real; for if it is just such an admission, then it is unclear on what grounds it

can still be claimed that the world as we know it and think about it is perception-independent, and therefore that the position has not simply lapsed into idealism.

Similarly, it might be objected that the metaphor has been taken too seriously: whereas the cookie cutter makes a real difference to the dough which it cuts, conceptual schemes do not make a real difference to the noumenal world which they classify. Hence, one conceptual scheme can do what it does to the noumenal world consistently with another scheme's doing what it does. But the difficulty with this objection is that the noumenal world can now fall away as redundant. For if the noumenal world is not being structured, formed, classified, and so forth, by our conceptual schemes -- i.e. suffering real changes -- then what role is it playing? None at all, I suggest; in which case we are returned to idealism.

Nothing in the Cookie Cutter metaphor, therefore, helps us to understand the key question confronting this attempt at formulating a metaphysical position which is fit to underlie pluralism, but which remains realist and empiricist: how could any kind of real world be such as to be carved up simultaneously by different conceptual schemes? Perhaps it will be said that it is inconceivable in any other terms than its own, since it is sui generis. But this seems to me to be a

response of desperation, and it should in any case be possible to supply some model which helps us, as the Cookie Cutter metaphor does not, to understand even the possibility of the metaphysical position on offer. Until that is done, there can be no obligation to treat it as a serious contender.

There is, however, a broader moral to be drawn from the failure just described, and I shall briefly draw it. The realist contends that non-mental reality is at most causally dependent upon the mind (see Devitt 1988, p.165). The anti-realist is apt to object that this is all too simple: there are other modes of dependence than the causal, and non-mental reality is dependent upon the mind in one of these other, more than causal, ways. But the problem for the anti-realist is then to describe one of these other modes suitable for his purposes. The difficulty to which the Cookie Cutter metaphor leads is, I suggest, just a special case of this general anti-realist difficulty; and it is worth noticing how the metaphors of cookie cutting and carving up are both causal metaphors, and hence useless for anti-realist purposes.

4. The Goodmanian Case Examined

Having defended my claim that pluralism leads both to anti-realism and to anti-empiricism, let me now turn to a

critical examination of the case for pluralism presented by the most eloquent of pluralism's advocates, Nelson Goodman.

Goodman's pluralist credentials seem impeccable. Hear him speak:

> ...in what non-trivial sense are there...many worlds? Just this, I think: that many different world-versions are of independent interest and importance, without any requirement or presumption of reducibility to a single base. The pluralist, far from being anti-scientific, accepts the sciences at full value. His typical adversary is the monopolistic materialist or physicalist... (Goodman 1978, p.4)

> I maintain that many world versions -- some conflicting with each other, some so disparate that conflict or compatibility among them is hardly determinable -- are equally right. (Goodman 1980, p.211)

But if it is clear that Goodman is a pluralist, it is much less clear why he is a pluralist. I think I can discern in his writings two main lines of argument, and I shall examine them in turn. There may in fact be a third (cf. Goodman 1978, pp.110-116), but it does not seem to differ importantly from an argument I ascribe to Putnam and criticise in my next section.

The first line of argument can be presented as follows (Goodman's versions are roughly equivalent to what I have called sciences or systems of discourse):

(1) In making versions, we thereby make worlds.

(2) We make many versions.

(3) So we make many worlds.

(4) So there are many worlds.

The inferences in this argument cannot be faulted, and I have no quarrel with premiss (2). Yet premiss (1) seems outrageous. Does Goodman really believe it? Hear him speak again:

> ...when I say that worlds are made, I mean it literally...Surely we make versions, right versions make worlds. (Goodman 1980, p.213)
>
> We can have words without a world but no world without words or other symbols. (Goodman 1978, p.6)

So Goodman certainly endorses (1), and his insertion of "right" does not diminish the outrageousness of the claim.

(1) appears to be just boringly false: we simply do not make, for example, the stars (cf. Scheffler 1980). So it is natural to conjecture that Goodman is using "make" in some non-standard sense. Indeed, he writes that "we do not make stars as we make bricks; not all making is a matter of moulding mud...The worldmaking mainly in question here is making not with hands but with minds, or rather with languages or other symbol systems" (Goodman 1980, p.213; cf. also his 1984, pp.34-36).

But now the problem is not so much that Goodman's claim is false, as that it is senseless. For what is this alternative sense of "making" to which he appeals? It cannot be the causal sense in which one makes bricks, that is, the sense in which making just is causally producing, since on that interpretation Goodman is offering us the boring falsehood. Nor can it be the criterial sense, which we intend when we say that one swallow does not make a summer or that four legs and a top make a table. For it seems equally obviously false that neither we nor our versions make, i.e. constitute, worlds: we and our versions are one thing, worlds are another. (Goodman could respond to this by insisting that there is no real distinction between versions and worlds; but that would just be a re-assertion of the conclusion, not an argument for it.) Since there seems to be no other sense of "make" available, we have to conclude that (1) is just false.

But perhaps my reply has been too quick. In response to Israel Scheffler's objection that there is no sense of "make" in which it is true to say that we make the stars, Goodman challenges his critic to say "which features of the stars we did not make, and...to state how these differ from features clearly dependent on discourse" (Goodman 1980, p.213; cf. also his 1984, p.33). It is perhaps unclear whether this statement

is meant to imply that there are features of stars dependent on discourse. But whether it is or not, the statement is suggestive.

What it suggests is two lines of thought which may be Goodmanian, of which the first is the following argument against realism. If realism is correct, then there is a description of the world independent of all concepts. But it is a conceptual truth that all descriptions involve concepts. So there cannot be a description of the world independent of all concepts. So realism is incorrect.

The argument is fallacious. Its first premiss is ambiguous, and when it is disambiguated it becomes clear that what realism is committed to is unaffected by the conceptual truth about descriptions, whereas what that conceptual truth refutes is not something realism is committed to. The first premiss could, on the one hand, mean this: if realism is correct, then there is a description of the world, which world is independent of all concepts. Perhaps realism, understood purely as an ontological claim, does not actually involve this commitment, but read this way the first premiss is nearly true, because the realist is likely to hold not merely that the mind-independent world exists, but that we can describe it to some extent. But on this reading of the first premiss, the argument does not go

through, since it does not follow from the fact that all descriptions involve concepts that there cannot be a description of a world, which world is independent of all concepts. This leads us straight to the second reading of the ambiguous premiss: if realism is correct, then there is a concept-independent description of the world. Now, of course, if the existence, or even the possibility, of a concept-independent description of the world really did follow from realism, then realism would indeed be refuted by the argument under discussion. But it does not follow from realism: the realist can be perfectly content with a concept-dependent description of a concept-independent world.

It might be held, however, that if the description of the world is concept-dependent, then the world thus described must be concept-dependent also. But this claim could only be maintained by confusing concept-independence with a certain notion of neutrality. Accordingly, the thought would be that if the world satisfies one description rather than another, then the world loses its title to be genuinely independent of that description. But of course if the world satisfies one description rather than another, all that is shown is that the world is not neutral in respect of descriptions of the world. But the realist does not have to suppose that the world is

neutral in this sense, and such neutrality does not reveal the world to depend upon concepts or descriptions in any way.

So much, then, for the first line of thought suggested by Goodman's quoted reply to Scheffler. The second line of thought is suggested also by Goodman's remark that, "No firm line can be drawn between world-features that are discourse-dependent and those that are not" (Goodman 1980, p.212). Let us interpret Goodmanian discourse broadly, to include concepts and conceptions. Then what is perhaps suggested here is, in effect, a slippery-slope argument: there is no principled way of drawing a sharp line between discourse-dependent and discourse-independent features of the world, and so all its features are really discourse-dependent. As stated, the argument is invalid. Like all slippery-slope arguments, it bends both ways. For all that it claims, we could as well conclude that all features of the world are really discourse-independent. If the argument is to serve pluralist purposes, it must include an additional premiss to the effect that at least one feature of the world is certainly discourse-dependent. But I doubt whether this additional premiss could be demonstrated with more plausibility than could a rival realist premiss to the effect that at least one feature of the world is certainly discourse-independent; and of course

such a realist premiss would enable us to conclude that all features of the world are discourse-independent.

But leaving aside all that, must we really accept that we can make no sharp distinction between discourse-dependent and discourse-independent features of the world? Surely we can say this: a feature of the world is discourse-independent iff it is at most causally dependent upon discourse (cf. Devitt 1988, p.165). On this simple and plausible account, which seems on the whole rather easy to apply, the chief, and perhaps the only, members of the class of discourse-dependent features of the world are discourses themselves. And their more-than-causal dependence upon discourse is evidently to be explained by appeal to the fact that they are identical with discourses. We should notice that we have here two relations, causation and identity, which although not perfectly understood, are nevertheless partially understood and entirely familiar. If it is still going to be maintained that stars are discourse-dependent, then it must be claimed either that our discourse causally produces stars, or that our discourse is identical with stars, or that stars are dependent upon discourse in some new and hitherto undisclosed sense. Of these three options, only the third is at all promising. But what on earth is this new type of discourse-dependence?

If there existed a class of features of the world which were clearly discourse-dependent, but equally clearly not dependent upon discourse either causally or in respect of identity, then we should have reason to think that there must be a third type of discourse-dependence, even if we cannot as yet give an account of it. But is there such a class? I think there is not. Let us consider some possible examples.

Our examples must, in effect, be of entities or features of the world which are not themselves identical with (or partially constituted by) concepts or conceptions, but which depend upon concepts or conceptions in some more-than-causal way. (Here I interpret Goodmanian "discourse", I hope plausibly.) It might be said that examples are these: famines, the feature of being red (on some analyses, anyway), money, chairs. Evidently these things are not identical either with concepts or with conceptions. But they are dependent; for it is true to say that, if there had been no humans, there would have been no famines, red items, money, or chairs. So they are examples of the required kind.

But they are not, and the reason is that they are not dependent in some more-than-causal way upon concepts or conceptions. The claim that they are thus dependent is supported by the alleged truth of the counterfactual beginning

"If there had been no humans...". Is this counterfactual true in respect of all the examples? Perhaps it is true of most of them, but it is surely false of red items. For, at least according to my intuitions on the matter, there could have been a red flash just after the Big Bang, even though there were no humans around then, and even if no humans had ever evolved. Where the counterfactual is true, does it indicate more-than-causal dependence? Not, surely, for the case of chairs, or indeed other artefacts; for their existence is causally dependent upon humans, who make them. But we must ask a third question: in the case of the examples where the counterfactual is true, and where it does not signal merely causal dependence upon humans, do we really have dependence upon concepts or conceptions? The correct answer is negative. All the candidate examples are entities or features whose natures make them relative to human interests and/or physiology. That is why they may be dependent in a more-than-causal way upon the existence of humans. They are, in effect, partially constituted by human interests and/or physiology. But this relativity to human interests and/or physiology is not the same thing as mind-dependence or, more specifically, dependence upon concepts or conceptions. And it is, of course, this latter kind of dependence in which we are interested. I see no difficulty in the idea of a famine's

afflicting some people who lacked the concept of famine, and who were the only people around. Similarly, there might be a people, and again we must think of them as the only people around, in whose society some type of thing functioned as, and therefore was, a money, despite their lacking the relatively sophisticated notion of a medium of exchange. So an entity or feature of the world can be partially defined in terms of human interests and/or physiology and still not be concept- or conception-dependent. It could even be the case that an entity or feature was defined partially in terms of concepts or conceptions, so long as the concepts and conceptions were not those of the entity or feature being defined. So even the truth of the relevant counterfactual beginning "If there had been no humans...", in connection with an entity or feature, does not show that the entity or feature is dependent upon concepts or conceptions. I conclude, therefore, that there is no third type of discourse-dependence, distinct from that dependence explained either by causation or by identity. Not only do we lack an account of its nature, but there are no clear examples of entities or features which exhibit it.

I said that I could discern in Goodman's writings two main lines of argument in support of pluralism. Having examined the first, let me turn now to the second. It is best introduced by

quoting from those writings:

> ...right versions are different from wrong versions...Rightness, however, is neither constituted nor tested by correspondance with a world independent of all versions. (Goodman 1980, p.211)
>
> We cannot test a version by comparing it with a world undescribed, undepicted, unperceived... (Goodman 1978, p.4)
>
> The overwhelming case against perception without conception, the pure given, absolute immediacy, the innocent eye, substance as substratum, has been so fully and frequently set forth...Talk of unstructured content, or an unconceptualised given, or a substratum without properties is self-defeating; for the talk imposes structure, conceptualises, ascribes properties. (Goodman 1978, p.6)

The general thrust of these remarks is presumably that, without some kind of suitable Archimedean point, the epistemologist will be in no position to reject some world-versions as defective. Thus he will be led either to scepticism or to pluralism, of which the latter is the preferable option.

Let me begin by clearing the ground of one obstacle to progress. In the first and second of the quoted passages, there is a hint of the idea that what the anti-pluralist realist must believe is that we can take a version of the world, take that portion of the world of which it is meant to be a version, and then compare the two, rather in the manner in which we can compare a photograph with the person of whom it is

a photograph in order to see if it is a good likeness. Despite the fact that it is analogous with something which is possible, such a comparison of versions with the world is, of course, not possible. But I see no reason why the anti-pluralist realist should deny this.

What the realist opponent of pluralism does believe, however, is that we can <u>test</u> our versions of the world against the world, though not of course by means of any attempt at direct comparison. Aware of this, and aware, too, that perception will play a crucial role in this testing of versions, Goodman argues that such testing is impossible, since perception is conception-laden, there is no unconceptualised given, and so forth.

Now the idea that perception is theory-laden is the first dogma of anti-empiricism. It is usually held to have two implications, both of which undermine the notion that we can make choices, on the basis of perception, between different world-versions. The first supposed implication is that we are rendered unable to test our theories against the world; the second is that, whether or not we can test any given theory against the world, we can never <u>compare</u> the merits of different theories by reference to perception: they are incommensurable. I shall begin with the first alleged implication.

Intuitively, the notion that by means of perception we can test our beliefs against the world involves the thought that there is a point, constituted by perception, at which the world simply forces itself upon us, regardless of our wishes or of our beliefs. But the fact that perception is theory-laden casts no doubt upon this thought. For the world can force itself upon us even if our contact with the world is always conceptually mediated or would change in character if our concepts were different. In an imperfect analogy, a punch from a gloved man may still knock me out, even though my contact with his knuckles is mediated by his gloves. Empiricist testing of realist theories of the world does not require experience whose character is uncontaminated by theory. What it does require is the possibility of recalcitrant experience, and theoretical contamination is, so far as I can see, quite compatible with that.

I suspect that any inclination to deny my last remark stems from putting a very strong, and indeed over-strong, interpretation upon the claim that perception is theory-laden. This claim is, of course, a metaphorical one, but it can perhaps be expressed more literally like this: the very phenomenal character of our sensory experience, and/or the perceptual beliefs we form, are affected by our non-perceptual

beliefs, theories, expectations and so forth, so that two subjects could enjoy experience of a different character, and/or form different perceptual beliefs, even though the stimulus in each case was exactly the same, if their beliefs and so on differed. On this interpretation of the claim that perception is theory-laden, the claim may well be true, though its extent has doubtless been exaggerated (for useful correctives, see, e.g., Kordig 1971, Ch.1; Fodor 1984 and 1988; Vision 1988, pp.271-285). But it is also possible to give to the claim the stronger interpretation that I mentioned. According to the stronger interpretation, the literalised claim is this: the very phenomenal character of our sensory experience, and/or the perceptual beliefs we form, are <u>wholly determined</u> by our non-perceptual beliefs, theories, expectations and so forth. Now it is evident that, if this stronger claim is correct, then perception cannot involve the world's forcing itself upon us. For if the character of our sensory experience, and/or the perceptual beliefs we form, really are <u>wholly</u> determined by our beliefs and so on, then no part of them remains to be determined, even partially, by the world's forcings. But the stronger claim is surely not acceptable. It does not seem to follow from its weaker sibling, and I cannot imagine what evidence could establish it. In addition, it makes it utterly mysterious how it could

ever happen, as it seems so often to do, that we suffer recalcitrant experience, that is, experience contrary to our expectations, beliefs, theories and even knowledge.

Let me turn now to the second alleged implication of the idea that perception is theory-laden, namely the claim that we can never compare different theories by reference to perception. Why is this claim of incommensurability supposed to follow from the idea that perception is theory-laden? The first point to notice is that if the theory with which perception is said to be laden can be any theory, then the claim of incommensurability does not follow. For the theory with which perception is laden can simply be irrelevant to the theories which we are trying to compare. For example, an observation might be made to the effect that a certain quantity of fluid turned red. Let us allow that this observation is theory-laden in the sense that the corresponding observational judgement has implications that go beyond the contents of the immediate experience. Nevertheless, it is hard to see why this observation could not serve to confirm some chemical theory and disconfirm another. For it does not seem to be laden with any significant chemical theory. We just do not need either theory-free, conceptually uncontaminated, perceptions or perceptual judgements in order to compare theories by reference

to perception (cf. Fodor 1988, p.189).

In response to this, it might be claimed that every perception is theory-laden in the sense that it is laden with the theory currently under test. But this claim is plainly false. A key observation for deciding some issue between Darwinian evolutionism and creationism might concern the colour of the shell of some rare type of beetle. But I can see no sense at all in which it would be true to say that one's observation that the shells of the beetles were blue was laden with the theory either of Darwinian evolutionism or of creationism, even if, as I allow, the observation is in some sense theory-laden. Moreover, it is hard to see what the mechanism by which perception becomes contaminated by theory is meant to be, especially when we remember that a scientist trying to compare two theories by reference to perception does not have to believe either of them.

What may be said in reply to this last point is that a theory under test contaminates perception because we have to use the vocabulary of the theory in order to report our perceptions. Well, that is certainly not always so, as my examples of the beetle shells and the mixture's turning red show. But perhaps it is sometimes so. Still, it is quite unclear, even in those cases, why incommensurability should be thought to follow. For

suppose we have a theory, T1, which entails a set of observational consequences described in the vocabulary of T1, and a second theory, T2. What rules out the possibility that T2 should beat T1 <u>at its own game</u>? That is, what rules it out that T2 should entail a set of observational consequences described in the vocabulary of T1, and that it should prevail over T1 when the relevant observations were made? It might be said that what rules this out is that if one describes the observational consequences in the vocabulary of T1, then one is thereby committed to believing T1, regardless of any empirical superiority revealed by T2. But this reply is mistaken. For, as Michael Levin has shown, there are linguistic devices available which enable us to describe observations in the vocabulary of one theory while explicitly cancelling our commitment to believing the theory. (In effect, this means denying the extreme meaning-holism which the theorist of incommensurability seems at this point to assume.) One such device is what he calls the "so-calling" operator. Using it, I can refer to an episode in my behaviour as "my so-called rudeness" without committing myself to the view that what I said was in fact rude. Similarly, the advocate of T2 can describe its observational consequences using the vocabulary of T1, but avoid any commitment to the truth of T1. If this is done, then the way is open for T2 to score a convincing victory

over T1 in point of empirical adequacy. (Here I draw on Levin 1979b, especially pp.416-419.) The theory-laden character of perception does not make it impossible to compare the merits of rival theories by reference to perception.

So much, then, for Goodman's second line of argument in favour of pluralism. Since it is as unsuccessful as his first, I conclude that he makes no sound case for pluralism, or for the anti-realism and anti-empiricism to which pluralism leads.

5. Putnamian Internal Realism

Hilary Putnam has been in his recent writings a defender of pluralism:

> There <u>are</u> tables and chairs and ice cubes. There are also electrons and space-time regions and prime numbers and people who are a menace to world peace and moments of beauty and transcendence and many other things. (Putnam 1987, p.16; cf. also Putnam 1981, pp.145 and 211)

There is a hint of back-sliding (e.g. Putnam 1987, p.20), but nothing is made of it. This pluralism is made possible, Putnam holds, by the adoption of a general metaphysical view called by him <u>internal realism</u>: "...one of the attractive features of non-realism is that it allows the possibility of alternative right versions of the world" (Putnam 1981, p.79).

So what is internal realism? It is evidently meant to be some kind of realism, and so we should begin by asking <u>why</u> it is felt by Putnam to qualify as a kind of realism. The question is pressing because he often contrasts it with what he calls <u>metaphysical realism</u>, a doctrine certainly very close to what one ordinarily thinks of as realism. This metaphysical realism is characterised in these terms:

> ...the world consists of some fixed totality of mind-independent objects. There is exactly one true and complete description of "the way the world is". Truth involves some sort of correspondance relation between words or thought-signs and external things and sets of things. (Putnam 1981, p.49)

It is natural to suppose that any view that contrasted too sharply with metaphysical realism thus defined could not itself be a brand of realism. More recently, Putnam has spoken of the need to abandon the dichotomy between real and projected properties, where a real property is one something has apart from any contribution made by language or the mind (Putnam 1987, p.8). So we can, and indeed must, endorse David Lewis' judgement here: "Internal realism...is realism feigned" (Lewis 1984, p.231). And yet there is a point to Putnam's insistence that internal realism is a genuine kind of realism, as we shall see in a moment (Putnam 1987, p.17).

Let us now consider the force of the epithet "internal"; what does it signify? The idea seems to be a denial that there is any absolute, non-relative notion of existence:

> ...what objects does the world consist of? is a question that it only makes sense to ask within a theory or description. (Putnam 1981, p.49)

> ...according to me, how many objects there are in the world...is relative to the choice of a conceptual scheme...How we go about answering the question, "how many objects are there?" -- the method of "counting", or the notion of what constitutes an "object" -- depends on our choice (call this a "convention"); but the answer does not thereby become a matter of convention. (Putnam 1987, p.33)

And we see from this last remark why Putnam feels that his position is genuinely realistic.

So Putnam is a pluralist, and his pluralism is made possible by his internal realism. Can we say in more detail how the internal realism leads to the pluralism? In a little more detail, I think we can. The idea is presumably that we have many conceptual schemes, and that in each of them questions of existence can be raised and, in most cases, answered affirmatively. So we will end up with a domain of entities for each conceptual scheme. Of course, we need to understand a conceptual scheme as a system of discourse such as physics or common sense or even, though with a little more strain, moral

assessment; but I think this is in line with Putnam's intentions.

The crucial question for Putnamian pluralism can now be posed: what arguments can be offered in defence of the internal realism which, it is plausible to claim, leads to pluralism? Putnam's writings contain, I believe, two main lines of argument. They are, first of all, the model theoretic argument, and then the argument from so-called conceptual relativity. I shall discuss each in turn.

The model theoretic argument (see Putnam 1978; 1981, ch.2; 1983, ch.1), even if successful, does not support internal realism directly. For it is presented in the first instance as an argument against metaphysical realism, so that it will support internal realism only if internal realism is the only alternative to metaphysical realism. It is perhaps doubtful whether this necessary condition is satisfied, but we should probably let that pass. More seriously, we might wonder whether, by abandoning the detailed formulation of metaphysical realism offered by Putnam (and quoted above) while retaining its spirit, we could evade the impact of the model theoretic argument. But there is no such easy way out. For the argument has as its target any view according to which the truth of a theory depends on (a) what the terms in the theory refer to and

(b) the way the mind-independent world is. Anyone who believes in a mind-independent world and who further believes that we can have theories about it which are true or false is threatened by Putnam's argument. One could certainly hold the first of these beliefs without also holding the second (see Devitt 1984), but the resulting position is, I think, unattractive.

So what is the model theoretic argument? Its interpretation is, unfortunately, not obvious, and I shall adopt David Lewis', possibly controversial, reading of it (Lewis 1984). The argument is an attempt to show that from the assumption that we have a theory about a mind-independent reality there follows an absurdity. The absurdity is that every theory, so long as it is consistent, is true of the world, so long as the world is large enough. And the metaphysical realist cannot avoid this absurd conclusion, Putnam contends, because reference to a mind-independent world is radically indeterminate. So metaphysical realism must be rejected, by an application of <u>modus tollens</u>.

In more detail, the argument runs like this. Let T be some theory purportedly about a mind-independent world. We must think of T as expressed in first-order predicate logic. If T is consistent, then T has some model on the basis of which we

can assign an interpretation to T. But then, if the world only contains enough objects, T will have a model in the world; indeed, it will have any number of models, and on one of them T is bound to come out as true. Let us call the interpretation based on this model -- the assignment of sets of objects in the world to the predicate letters of T -- the intended interpretation, i.e. the interpretation which specifies the references of the terms in T. Then what is to stop us saying that T is true? Nothing, it seems, at all: there are no further constraints to which the realist can appeal in order to exclude this interpretation from counting as the intended one. For suppose there was some further constraint; then this constraint, being describable, would itself be a theory which we could set alongside T, and which could be subjected to the same treatment as T. But then we would have got nowhere, since there is bound to be an interpretation, in the world, of T together with the proposed constraint according to which T and the constraint are true. And what goes for T goes too for any other consistent theory expressible within first-order predicate logic. So any such theory is true, and indeed true of the world.

Let me now respond to the argument. (In doing so, I take myself in the main to be following David Lewis (Lewis 1984),

Michael Resnik (Resnik 1987) and Alvin Plantinga (Plantinga 1982).) The alleged absurdity is certainly an absurdity, and there can be no question of trying to live with it. I shall argue, however, that the model theoretic argument demonstrates less than Putnam supposes, and that what it demonstrates the metaphysical realist can (just about) live with. In more detail, this response involves three claims: first, that what the model theoretic argument certainly shows is that <u>if</u> there is no constraint on what can count as the intended interpretation of a theory, then the absurd conclusion that (roughly) every theory is true whatever the world is like indeed follows; secondly, that Putnam does not show that any constraint which might be proposed <u>must</u> fail to determine reference; and, thirdly, that there is reason to think that there has to be some constraint, even though we cannot as yet say what it is. If these three claims are accepted, then the metaphysical realist will have achieved a stand-off. I acknowledge that this is not very satisfactory; but it will suffice to repel Putnam's challenge, and anything better would require a full theory of reference which I do not possess.

The first of the three claims will, I presume, pass unchallenged by a defender of the model theoretic argument. The same, however, is not true of the second, and so I shall

now defend it. Putnam's reason for supposing that any proposed constraint upon what may count as the intended interpretation of a theory must fail is that it will be possible to trivialise the constraint -- to show that it is too easily satisfiable -- by appeal to the same considerations which showed that T was bound to be true. But Putnam here misunderstands the nature of the proposal. The proposal is not that some <u>sentence</u> expressing the constraint is true, though this need not be denied; rather, the proposal is that there <u>is</u> a certain constraint which delimits the class of intended interpretations: it is the operation of this constraint, its operation being a non-linguistic fact that can obtain whether or not we can or do describe it as obtaining, which determines reference. Of course, it is necessary to assert the existence of this constraint using language; but this is harmless, because if there is such a constraint, then the very operation of it can ensure that the assertion that it exists has determinate reference. The realist can allow that nothing that we <u>say</u> -- nothing <u>internal</u> to language -- is sufficient by itself to fix one interpretation as the intended one; but the possibility is left open that something <u>external</u> can join with language to do so. Several specific proposals have, of course, been offered, for example, the causal theory of reference (see Devitt 1984; Brueckner 1984b) or a theory of eligible referents

(see Lewis 1984), though I do not necessarily endorse either of these. Putnam has claimed, in response to this kind of objection, that he cannot understand external constraints of this sort (Putnam 1983, p.xii); but the candidates offered so far may not be the only ones possible, and in any case he may not be trying hard enough.

So much, then, for my second claim. If I am right, the dialectical position is that, for all that Putnam has shown, there might be an external constraint sufficient to delimit the class of intended interpretations of T. My third claim was that we have reason to think that there has to be such a constraint, even if we cannot say what it is. This reason is simply the intuitive unacceptability of Putnam's conclusion that metaphysical realism is false -- that there is no thoroughly mind-independent world. Instead of reading his argument as an argument to this conclusion, we can read it as a <u>reductio</u> of one of the premisses, namely the premiss that there is no external constraint fixing reference (Lewis 1984, p.225). If we do so, we can set up a contest of intuitions: we have the realist intuition that realism must be right <u>versus</u> Putnam's intuition (for that is what it is) that external constraints make no sense. Now my official view on the role of intuitions in philosophy is that they are to be avoided. But in cases

where, as here, we have no arguments, perhaps we are obliged to resort to them. If so, I know for my own part which way I shall jump. An intuition to the effect that metaphysical realism is correct cannot, of course, establish metaphysical realism; but it can defeat Putnam's intuition that external constraints make no sense. The model theoretic argument certainly does not force us to forsake metaphysical realism in favour of the internal realist's cold embrace.

I turn now to what I called the argument from conceptual relativity. The term, "conceptual relativity", is Putnam's (Putnam 1987); and he seems to regard it as denoting a certain phenomenon which metaphysical realism, to its discredit, cannot recognise (Putnam 1987, p.19), but which is an integral part of internal realism. Indeed, Putnam's order of exposition strongly suggests that conceptual relativity provides the core of the defence for his idea that "how many objects there are in the world...is relative to the choice of a conceptual scheme" (Putnam 1987, p.32-3).

So what is conceptual relativity meant to be? Does it really exist? And if it does, how exactly does it lend support to internal realism and thence pluralism? These questions are not easy to answer, because Putnam's argumentation is not explicitly set out. In attempting to answer them, therefore, I

shall consider two different interpretations of the notion of conceptual relativity, though I am not entirely confident that either interpretation faithfully captures Putnam's intentions. At any rate, on neither interpretation does conceptual relativity lead us, via internal realism, to pluralism.

On the first interpretation, the notion of conceptual relativity is the notion of a phenomenon exemplified whenever there exists a situation which can properly be described as follows: there are two descriptions, D1 and D2, of what is in some sense the same thing, such that (a) there is excellent reason to accept each of D1 and D2, (b) there is no possible reason for preferring one of D1 and D2 to the other, and (c) D1 and D2 make different and incompatible ontological claims. The significance of conceptual relativity, on this interpretation, is that whenever it occurs we may, and indeed must, accept <u>both</u> D1 <u>and</u> D2, together, of course, with their respective ontological commitments. Relative to each description, there is an ontology, and we can accept both ontologies. This pluralist conclusion follows because there is no other response possible to such a situation. We cannot accept just one of D1 and D2, because by hypothesis there is nothing to choose between them; nor can we reject them both, because, again by hypothesis, each description taken independently is very

plausible. It might be said that this has not exhausted the alternatives, because there is the neglected possibility of accepting a disjunctive description, [D1 or D2]. But this possibility seems to be excluded by clause (b); for if there is no possible way of choosing between D1 and D2, so that there is no prospect of further considerations emerging to resolve the dispute between D1 and D2, then it is hard to see the point of affirming a disjunctive description. Whether a realist should allow this last move will depend on how "possible way of choosing" is to be understood. I assume that it can be understood in a sufficiently liberal way. (For relevant remarks, see Devitt 1984, p.187.)

This interpretation is suggested by Putnam's examples intended to illustrate conceptual relativity (Putnam 1987, pp.18-19 and 32-33), and by various remarks (Putnam 1987, pp.20 and 29; and 1981, p.73-4; cf. also Goodman 1978, pp.110-116). But if this interpretation of the notion of conceptual relativity is genuinely Putnamian, then it is inadequate as a foundation for pluralism. An initial doubt would concern whether the alleged phenomenon of conceptual relativity is ever in fact instantiated, that is, whether there ever is or has been a situation satisfying the specifications given in my last paragraph. Putnam offers two alleged examples, one

mereological and one geometrical (Putnam 1987, pp.18-19). But since such examples by their very nature do not involve the notion of explanatory or empirical adequacy, their relevance to a full-scale pluralism of the sort I am interested in is unclear. (Also, they might be held to compel only very local retreats from realism; cf. Wright 1985, pp.510-511.) Putnam claims, however, that the phenomenon of conceptual relativity "has become pervasive in contemporary science" (Putnam 1987, p.29), but he supplies no examples in that passage. An example does appear, however, in an earlier work (Putnam 1981, p.73); but it is open to the objection that it fails to satisfy the required conditions (see McMichael 1988, p.62ff.).

But I shall leave aside this initial doubt. For even if we allow that the phenomenon of conceptual relativity is genuine, we still do not reach a sufficiently impressive conclusion. The existence of conceptual relativity, as defined, shows that ontology is relative to choice of description (and hence, by extension, conceptual scheme), and that we can and must accept apparently incompatible ontologies, only if we have a situation satisfying the peculiar conditions specified in the definition of conceptual relativity. But there is no reason to suppose that those rather exacting conditions are in fact satisfied by the conceptual schemes with which we are concerned, namely

physics and the special sciences. Certainly these schemes make incompatible ontological claims (if we assume that their ontologies are located in a single spacetime), but it can hardly be said that there neither are nor could be any grounds for choosing between them. To make these further suppositions it would be necessary to discredit sensory experience in the way attempted, though unsuccessfully, by Goodman. It would not be enough to show that we humans, through contingent epistemic limitations, could have no grounds for choice between the schemes. What Putnam needs to claim is that physics and the special sciences are empirically equivalent theories. But, unsurprisingly, he does not demonstrate this.

So I do not think that, on this first interpretation, conceptual relativity serves to advance the internal realist and pluralist cause. What of the second interpretation? According to it, conceptual relativity just is the fact that the word, "exists", or the logical constant, "Ex...", has many meanings, and that corresponding to each meaning there is a distinct kind of existence. If this is indeed a fact, then the door to pluralism is opened wide. For if there are many kinds of existence, then there really can be many realities, and the monistic temptation to suppose that everything that is real must somehow be able to fit together so as to constitute one

reality could be effortlessly resisted. And if there is a distinct sense of "exist" for each conceptual scheme, then one scheme could entail that there exist F's but no G's, and another could entail that there exist G's but no F's, and there need be no inconsistency between them, since "exists" need not mean the same thing in each claim. To deny this would be to equivocate on "exists".

This second interpretation is suggested, for example, by Putnam's insistence that the phenomenon of conceptual relativity "turns on the fact that the logical primitives themselves, and in particular the notions of object and existence, have a multitude of different uses rather than one absolute 'meaning'" (Putnam 1987, p.19; italics removed). Other remarks echo this theme (Putnam 1987, pp.20, 35-6).

Putnam's contention that there are many different senses of "exists" is hard to believe, especially since it is such an extraordinary boon to the pluralist. Moreover, unless we are offered some principled way of telling how many such senses there are, nothing prevents the defender of some very weakly confirmed theory from continuing to believe in the theory's ontology on the grounds that its ontology exists in some hitherto unrecognised sense. So the defence of it must be powerful if it is to carry conviction. What, then, is Putnam's

defence? It is embodied, I think, in a number of brief and inexplicit remarks, which I shall now examine.

It is Putnam's opinion that "the idea that there is an Archimedean point, or a use of 'exist' inherent in the world itself, from which the question 'How many objects <u>really</u> exist?' makes sense, is an illusion" (Putnam 1987, p.20). Now there is a truth lying behind this claim, though it goes no way to support the desired conclusion. The truth is this. If I ask the question, "How many things are there in this room?", then I have not asked a question in good order: as it stands, the question cannot be answered -- or rather there could be many perfectly good answers. What I should have asked is, "How many chairs are there in this room?", or, "How many tables are there in this room?". That is, I should have specified what <u>kind</u> of thing I was enquiring about; and this specification would have required the use of a concept, that of <u>chair</u> or <u>table</u> or whatever.

So much for the truth, which reflects the metaphysical fact that it makes no sense to suppose that a thing can exist without being a thing of a particular kind. Does it support the idea that there are many senses of "exist"? I do not see how. For it seems manifestly consistent with this truth that one should answer the questions just mentioned with the

(admittedly unidiomatic) replies "There exists one chair in this room" and "There exists one table in this room" without any suggestion that "exists" is not univocal. Does the truth perhaps show instead that existence is in some way dependent upon concepts or conceptual schemes? Again, I do not see how; though it does show, what no realist would deny, that saying or asking what exists depends upon concepts and the schemes in which they occur.

In his next paragraph, Putnam speaks of "the mistake of supposing that 'Which are the real objects?' is a question that makes sense independently of our choice of concepts" (Putnam 1987, p.20). Now if the mistake mentioned here is that of denying the truth that we can only ask how many F's there are, not how many things (or objects) there are, then we have to allow that it is truly a mistake. But why does this tend to establish that there is more than one sense of "exist", so that it is harmless if different schemes involve apparently incompatible claims about what exists? Of course, what is involved in a table's existing differs from what is involved in a chair's or an electron's existing; but what we are looking for is variations in the sense of exists across schemes, not across kinds.

Putnam extends the same idea to the notion of fact, and in a

spirit similar to that animating the two remarks just quoted, claims that "What we cannot say -- because it makes no sense -- is what the facts are independent of all conceptual choices" (Putnam 1987, p.33). But this is ambiguous. Is it our saying what the facts are which cannot be independent of all conceptual choices? Or is it the facts themselves which cannot be thus independent? The answer to the first question is surely "Yes"; that to the second, in the light of my earlier discussion, is "No"; and there is no obvious inconsistency in returning these different answers. Saying is evidently a linguistic, and hence conceptual, affair; whereas facts, except for linguistic and conceptual ones, are not. There is nothing here to dissuade us from our habitual realism.

So far, then, and on both my interpretations of the notion of conceptual relativity, I have found in it no satisfactory defence of either internal realism or pluralism. My suspicion is that what lies at the back of the notion is the following sort of case. We can say that there are two chairs and one table in a room and that there are three items of furniture in it. That does not make six items in total, of course: there are no items of furniture in addition to the chairs and the table. So we can make these two claims, which are in some sense different, and they can both be true of the same

situation. Now it might be thought that cases of this sort provide a model for pluralism, showing how recognisably different ontological claims can both be true of the same portion of spacetime without there arising awkward questions about why the objects said to exist by the two claims do not bump into one another. But it is most implausible to try to generalise from this sort of case to a full-scale pluralism of the kind which interests us. For it is analytically true that a table (or a chair) is an item of furniture, just as it is analytically true that a red object is a coloured object. Tables and chairs are, as it were, determinates of the determinable, item of furniture. And all cases of this sort will, I think, involve a relation analogous to this. But if this is so, and I know of no exception, then cases of this sort can provide a model for pluralism only if the kinds of things mentioned in the many sciences can be made out to stand in an appropriate determinate/determinable relation to one another. Now this is most unlikely to be possible, especially if we require the relation to be analytic. Someone prepared to give a disjunctive definition of every non-physical property in terms of very many (perhaps infinitely many) disjuncts each of which mentioned a physical property could perhaps do it; but the resulting view would, it seems to me, be token physicalism, which is not pluralism and against which I have presented other

arguments.

6. Conclusion

I have argued that neither Goodman nor Putnam has presented arguments adequate to support pluralism, and in particular that neither of these philosophers has made plausible the anti-realism and anti-empiricism to which pluralism leads. So pluralism must be rejected. It cannot serve as an account of the relations holding between the many sciences and between their many domains.

CHAPTER FOUR

Eliminative Type Physicalism

As I indicated in the Introduction, my own favoured response to the problem of the many sciences is the brand of physicalism which I call eliminative type physicalism. (Though I am the first to name it thus, the view, or something very much like it, has been defended in the past; see, for example, Sellars 1963, pp.106-126, Maxwell 1968, and Feyerabend 1981, pp.44-96.) The aim of the preceding three chapters has been to provide indirect support for eliminative type physicalism by arguing that the only serious rivals to it, namely token physicalism, supervenience physicalism and pluralism, are, in their different ways, defective. That phase of my argument is now complete. The aim of this chapter is to begin the task of providing direct support for eliminative type physicalism by constructing a positive argument for it. More specifically, I want in this chapter to do three things: first, make clearer exactly what the thesis of eliminative type physicalism is, qualifying it in important ways; then, give some account of the fate of the eliminated sciences; finally, sketch an outline of the positive argument for eliminative type physicalism, the

elaboration of which will follow in Part 2.

1. "Physical"

Eliminative type physicalism is the claim that every property -- every property that genuinely exists -- is a physical property, so that if some putative special scientific property cannot be type identified with a physical property, then we cannot suppose the putative special scientific property to exist. And when I say that this, that or the other does or (more usually!) does not exist, I mean to use the word, "exist", in the everyday sense in which everyone thinks that tables and chairs and planets exist, and people dispute whether or not God, or Yetis, or electrons, exist. This everyday sense of "exists" is, in my view, the only sense of the word.

But before the claim that every genuine property is a physical property can be either defended or attacked, I need to explain with some precision which properties are to count as physical. If I do not do this, then I am open to the charge of having said, quite literally, nothing. But although I am under this obligation, which I shall shortly discharge, of explaining clearly what I mean by saying that a property is physical, it is important to see that I am under no obligation whatsoever to supply anything like a conceptual analysis of the word,

"physical". This is because I am trying to delimit a class of properties, rather than investigate the uses of a predicate. Accordingly, it will not do to respond to what I have to say in this section by providing alleged counter-examples. Counter-examples are appropriate arrows to hurl at an analysis; but I am not offering an analysis.

Which properties, then, are the physical properties? One suggestion is this: a property is physical iff it is mentioned in the laws of completed physical theory. But this suggestion is worthless. For even if such a theory ever exists, and nothing, so far as I can see, guarantees that it will, we have at the present time only a vague and rather negative idea of which properties will rate a mention in the laws of completed physical theory. So if the suggestion is used in a formulation of physicalism, it will have a tendency to evacuate physicalism of interesting content. Here, then, is a second suggestion: a property is physical iff it is mentioned in the laws of current physical theory. This suggestion is better, but it binds too tightly the hands of a physicalist who uses it. Specifically, it entails, absurdly, that we could not call a newly discovered property physical.

Now some philosophers write as if these two unsatisfactory suggestions exhaust the field (see perhaps Crane and Mellor

1988, p.4). But of course they do not. For there is an intermediate position, and it exploits the intuition that current physics could change quite a bit without ceasing to be physics, or even current physics. This intuition could no doubt support a variety of different suggestions, but the one I shall use is essentially this: a property is physical iff it is mentioned in the laws of a science most of whose properties are mentioned in current physics. This explanation of which are the physical properties is still vague, and we could squabble over the precise interpretation of "most"; but I doubt that it would make an important difference. According to it, however, membership of the class of physical properties, while selective -- it is constrained to a considerable extent by the content of current physical theory --, is not absurdly exclusive.

This explanation, however, is still not quite right. It will need one more amendment, but since this is best made only after I have sketched the outline of the positive argument for eliminative type physicalism, I shall delay its description until then.

2. The Fate Of The Eliminated Sciences

The response of eliminative type physicalism to the problem of the many sciences is, of course, harshly monistic. It

assigns to the domain of just one of the sciences, physics, the strongest possible form of metaphysical primacy: exclusive reality. But what is to be the fate of the other, special sciences, which are to be eliminated? To the extent that the properties included in their domains cannot be identified with physical properties -- and I take that extent to be great -- the existence of those properties must be denied. Does this mean, then, that the special sciences are to be junked, and left to lie in the dustbin of history alongside the phlogiston theory of combustion and Ptolemaic astronomy? Not necessarily. The reason is that we can give an instrumentalist interpretation to the eliminated sciences. Thus they can be regarded as useful, perhaps indispensably so, in everyday life, but as useful fictions. I myself see no prospect whatsoever of our coming to learn to respond to the world spontaneously in terms of current physics, as some philosophers sympathetic to my position have suggested; indeed, I would even venture the speculation that we are hard-wired, for evolutionary reasons, to think of the world in folk-physical terms. (There have apparently been experiments on young babies which support this speculation.)

This instrumentalist interpretation could assume either of two forms (see Peacocke 1983, p.203). First, it could purport

to be an account of how practitioners of some given special science themselves regard their practice. The claim would then be that the special science in question was not even an attempt to reveal the nature of reality, but on the contrary an effort in some other direction, perhaps purely predictive. If such a claim was accepted, then we would be in a position to defend our everyday special scientific practices from the charge that they systematically get the nature of that reality wrong by their making of erroneous claims about the existence of properties. (See Blackburn 1984, pp.171 and 180 for the closely-related quasi-realist project.)

The second form that an instrumentalist interpretation of the eliminated sciences could take would be, in essence, a salvage operation. It would be conceded that a given science was in radical error, but maintained also that there was a certain utility in continuing to speak as if it were not. And we would not then, of course, be obliged to send the science the same way as phlogiston theory.

The fact that we can offer instrumentalist interpretations of the eliminated sciences is of some importance. Here is why: it enables the eliminative type physicalist to give an ontologically harmless sense to the widely-voiced feeling that the special sciences are autonomous, or irreducible, or

ineliminable. For he can say that the special sciences are autonomous and irreducible and ineliminable in the sense that they are practically indispensable for beings such as ourselves, with our peculiar cognitive limitations. But in saying this, and thereby adopting the instrumentalist interpretations, he is not committed to the special scientific ontology which he officially disallows. In accepting eliminative type physicalism, we do not even have to forgo the explanatory advantages of the special sciences, so long as we realise that all we are entitled to are what we may call pseudo-explanations, i.e. explanations the statements of which are explicitly given instrumentalist interpretations.

Of course, the defender of the special sciences wants rather more than I have given him: he wants the ontology too! He will claim that the explanatory advantages of the special sciences are that they explain -- really explain -- certain regularities left unexplained by physics, and that is why we must keep the special scientific ontology (see perhaps Fodor 1980, p.131, and certainly Owens 1989, pp.78-79). But what are these unexplained regularities? It turns out, of course, that they are regularities holding among special scientific events. But then we have in fact got nowhere. For the defender of the special sciences needs the claim that realistically interpreted

special sciences are explanatorily advantageous, so that we have to retain the special scientific ontology. But he supports that claim only by assuming the existence of at least some special scientific ontology, namely those events among which hold the regularities alleged to be unexplained by physics. Thus we have been given no non-question-begging reason why instrumentally interpreted, and hence ontologically non-committal, special sciences are not enough (unless his position is allowed to collapse into The Obvious Objection (see the Introduction pp.10-11)).

So the fate of the eliminated sciences is for them to be regarded properly as constituting no more than useful fictions. But useful for what? One use I have in mind is, of course, that of helping us get by in daily life, and the importance of this should not be downplayed. But there may be something else for which they can be useful. I am alluding here to the possibility that the eliminative type physicalist is not committed to denying Jerry Fodor's claim, upon behalf of token physicalism, that "The point of reduction is...to explicate the physical mechanisms whereby events conform to the laws of the special sciences" (Fodor 1980, p.127). Fodor's claim need not be denied so long as -- what seems perfectly possible -- one construes the laws of the special sciences

instrumentally. If this is right, the eliminative type physicalist can allow a role for the special sciences in guiding the research of physicists. But I have no further account of how this guiding is to be theorised.

3. The Positive Argument Outlined

One very important challenge to physicalism, especially eliminative type physicalism, consists in asking for the reason why we should believe that physics, and physics alone, is the source of our knowledge of what there is: what is so special about physics? (See especially Crane and Mellor 1988.) The positive argument for eliminative type physicalism, which I shall in this section sketch, supplies an answer to this question, and hence a response to the challenge.

The point of departure for the argument is the traditional problem of scepticism about the external world. Such a point of departure should occasion no surprise. For the question of physicalism just is, to a very great extent, the question of which external objects -- of what sort of external objects -- we are justified in supposing to exist, and the sceptic is someone who alleges that we are justified in supposing no external objects to exist. If we can show why he is wrong, and hence that we <u>are</u> justified in taking some external objects to

exist, then we are almost bound to cast some light on the question of which external objects we are justified in taking to exist. I argue that this expectation is not disappointed.

So I begin by taking scepticism about the external world seriously. My response to it builds upon an idea I take from the late John Mackie, though he did not originate it, and it now has several supporters (see Mackie 1976, pp.62-67; and, for instance, Goldman 1988, pp.203-215 and Lycan 1988). The idea is that external objects are postulated as part of a realistic hypothesis intended to constitute the best explanation of the course of our perceptual experiences; and since we are justified in taking the best explanation of some set of phenomena to be literally true, we are justified in taking the postulated external objects to exist. But this idea, I argue, has implications which subvert Mackie's original intention of vindicating commonsense realism. For I try to make plausible the claim that it is physics, and physics alone, which, with a relatively small amount of adjustment and supplementation, constitutes the realistic hypothesis best explanatory of the course of our perceptual experience. But if this is right, then it is physical properties, and only those, which we are justified in taking to be real. Mackie was a believer in the reality of physical properties, but he wanted to validate that

belief through the intermediary of justified belief in the existence of the macro-objects of common sense. My suggestion is that this is an unnecessary detour. We can cut out the middle-man. Someone in the egocentric predicament of traditional scepticism need only postulate the world of physics. He need go no further: the hypothesis of a special scientific ontology is a hypothesis he can do without. If my argument is correct, it gives sense to the claim that physics and the special sciences are rival theories in respect of what is basically the same set of data, so that if we accept physical theory, then we simply do not need the other theories. The ontology of physics crowds out that of the special sciences, not because it is <u>inconsistent</u> with the special scientific ontology, but because it makes the latter explanatorily redundant, and hence epistemically gratuitous (cf. Crane and Mellor 1988, p.5).

This exposition of the general form of the positive argument is crude. A more refined version is this:

> (P1) We are justified in taking a property to exist if, but only if, it is mentioned in the best explanation of the course of our perceptual experience.
>
> (P2) PHYSICS is the best explanation of the course of our perceptual experience.

So: (C1) We are justified in taking a property to exist if, but only if, it is mentioned in the laws of PHYSICS.

(P3) The properties mentioned in the laws of special science, S, are not mentioned in the laws of PHYSICS.

So: (C2) We are not justified in taking to exist the properties mentioned in special science, S.

The conjunction of (C1) and (C2) are to be understood as representing the thesis of eliminative type physicalism. Notice that it involves an epistemic dimension so far unmentioned.

Several of the claims made in this refined version of the positive argument for eliminative type physicalism need elucidatory commentary.

(P1) is evidently the key philosophical premiss in the argument. What it states is a specific version of what is sometimes called the explanatory test for reality, the thought that for a property to count as genuine it must do necessary explanatory work. As I noted in the Introduction, although the explanatory test for reality is often invoked in philosophical discussion, it is unclear what reason there is for adopting it, and accordingly some philosophers reject it (see Wiggins 1976, pp.361-364). (P1) will be defended at length in Chapters Five and Six. This defence will provide a philosophical motivation

for at least my version of the explanatory test for reality.

(P2) is the key empirical premiss of the argument. It is, in effect, a statement of the intuitive idea that physics is by itself explanatorily adequate. (For doubts about one version of the idea, see Stroud 1986-7, p.271ff.) Since this idea could no doubt be cashed out in other ways, (P2) is just one interpretation of the idea. It is defended in Chapter Seven.

I have said that (P2) is an empirical premiss, and I wish now to emphasise its empirical and contingent character. Physics, or rather PHYSICS, owes its preeminence in the eliminative type physicalist scheme of things to contingency, to the contingency of its explanatory adequacy. It is in no way conceptually or logically or metaphysically guaranteed that physics, or rather PHYSICS, should constitute the best explanation of the course of our perceptual experience. Moreover, the contingency of the premiss, (P2), carries over to the conclusion, (C1), which is derived from it, thus making eliminative type physicalism a contingent thesis. If (P2) is false, or if physics undergoes revolutionary changes, then eliminative type physicalism will be refuted. In fact, I prize this empirical character of my view. But it is in any case inevitable. For what else could any thesis of physicalism be than a large-scale claim about the nature of empirical reality, a claim which might turn out to be

false?

Premiss (P2), as well as the conclusion (C1), speak of PHYSICS. But what is PHYSICS? PHYSICS is closely related, of course, to physics, though it is not identical with it. It is not identical with it, because I do not believe that physics, exactly as it is, constitutes the best explanation of the course of our perceptual experience, though it does, I shall argue, supply the materials for something which does, namely PHYSICS.

PHYSICS is both less and more than current physics. It is less than current physics, because PHYSICS is likely, I think, to include only the microphysical portions of physics. That is to say, non-microphysical physics will probably suffer the fate of the eliminated sciences, being a superfluous part of the best explanation of the course of our perceptual experience. And PHYSICS is more than current physics, because, as I indicated in Section One of this chapter, physics could change to some limited extent, perhaps embracing newly discovered properties, without ceasing to be physics. (And PHYSICS may need a few properties not mentioned in physics if it is to constitute the best explanation of the course of our perceptual experience; see Chapter Seven.) But although PHYSICS is not quite the same as physics, PHYSICS is still pretty closely

related to physics, and the differences between PHYSICS and physics can give scant comfort, so far as I can see, to the enemies of physicalism.

I can now state the final modification, deferred at the end of Section One, to my account of which properties are to count as the physical properties. The physical properties are, of course, the PHYSICAL properties, those properties mentioned in the laws of PHYSICS. These are the properties which eliminative type physicalism claims to be the only ones that there are.

(C1) states the essence of eliminative type physicalism. Since it claims that we are justified in believing in a PHYSICAL ontology, and since I have conjectured that PHYSICS is roughly equivalent to microphysics, it follows that eliminative type physicalism involves the claim that we are justified in believing in the ontology of microphysics. Believing in the ontology of microphysics means believing that it exists, and indeed that it exists in a robustly realist, thoroughly mind-independent, way. I have elucidated and defended this robust realism, together with notion of thorough mind-independence, in Chapter Three (Sections Three, Four and Five).

So belief in the ontology of microphysics is implied by (C1).

But belief in the ontology of microphysics is plausibly regarded as a version of the doctrine of scientific realism, the doctrine that we should believe in some, at least, of the unobservable entities postulated by science. But since premisses (P1) and (P2) entail (C1), they must be seen as constituting a defence of, and argument for, this version of scientific realism, this belief in the ontology of microphysics.

But this defence, this argument, involves a highly controversial assumption, and it is one which I shall mention only to ignore. The assumption is involved in premiss (P2). (P2) claims that PHYSICS, which I have conjectured to be approximately equivalent to microphysics, constitutes the best explanation of the course of our perceptual experience. The assumption involved in this is that it is possible to give a realistic interpretation to quantum mechanics, the core of microphysics. By a "realistic interpretation", I mean an interpretation which at least represents quantum mechanics as purporting to be a description of the world (whether a true one or not), and which means that quantum mechanics can constitute at least a candidate explanation (whether correct or not) for the course of our experience. Realistically interpreted, quantum mechanics will be viewed as describing real items and

the ways in which they behave, items whose postulation would explain the experimental data which confirm quantum mechanics. But if this realistic interpretation is impossible, then (P2), as I have elucidated it, cannot be correct, and the positive argument for (C1), the essence of eliminative type physicalism, will accordingly fail. This would be very bad news for eliminative type physicalism, because I have no idea how eliminative type physicalism could be argued for unless by appeal to my positive argument or to some argument so similar to it as to encounter the same difficulty.

Why is a realistic interpretation of quantum mechanics so controversial? The reason, as I understand it, is as follows. The theory of quantum mechanics consists in a highly sophisticated mathematical formalism. Although empirical predictions can be derived from this formalism, which is why the theory is so well confirmed experimentally, it is entirely unobvious, intuitively, what reality, if any, this formalism describes. This is due partly to the weirdness (at least to the mathematically untrained) of the formalism, and partly to the strangeness of certain experimental results (such as two-slit experiments). (For a survey of the problems which is, nevertheless, optimistic about the prospects for a realistic interpretation, see Forrest 1988b.) The reader may be

wondering, since the possibility of a realistic interpretation of quantum mechanics is so crucial to premiss (P2) of the positive argument for eliminative type physicalism, why I do not defend the possibility in this thesis. The reason is that such a defence lies beyond my resources of both space and, especially, competence. Indeed, I doubt that anyone who, like me, does not truly understand the formalism of quantum mechanics could contribute usefully to discussion of its proper interpretation.

But although I cannot defend the possibility of a realistic interpretation of quantum mechanics, I will briefly indicate what I take to be the consequences for eliminative type physicalism if my assumption is false and no such interpretation is possible. First of all, as I have said, if no realistic interpretation is possible, then (P2), elucidated in the way I intend, cannot be correct and so the positive argument for eliminative type physicalism would collapse. But, in spite of this, there are a couple of fall-back positions to which I could retreat. First, since the question of the interpretation of quantum mechanics has no tendency to affect premiss (P1) of the argument, (P1) could in principle be combined with some appropriately weakened analogue of (P2) in order to support a correspondingly weakened conclusion. What I

have in mind is this. It could perhaps be argued that some physico-chemical theory which took atoms and their tendencies to combine as basic, and which therefore did not need a (quantum mechanical) theory of the atom, constituted the best explanation of the course of our perceptual experience. This claim, conjoined with premiss (P1), would permit the inference of some doctrine of physico-chemicalism, according to which we would be justified in believing in a property if, but only if, it was mentioned in the requisite physico-chemical theory. Such a conclusion would, I suspect, offer scant consolation to the traditional enemies of physicalism, and would, indeed, amount to almost as grave an offence against common sense as eliminative type physicalism. I remarked a moment ago that the failure of a realistic interpretation of quantum mechanics would be very bad news for eliminative type physicalism, and that is true, but a doctrine of physico-chemicalism could survive this failure and would certainly constitute a robust form of physicalism. It could survive the failure of a realistic interpretation of quantum mechanics because arguments for the existence of atoms and molecules, arguments of the sort which convinced Einstein, would retain their force whatever view one took of quantum mechanics. (On Einstein's argument, see Miller 1987, pp. 470-482.) But I shall not comment further on the plausibility of this possible fall-back position of

physico-chemicalism.

There is, however, a second possible fall-back position. This fall-back position consists simply in defending premiss (P1) and then suspending judgement about what, if any, empirical premiss should be joined with (P1) in order to sustain a conclusion about what we are justified in believing to exist. Such a fall-back position would still be interesting, because it would have elucidated the epistemological principles relevant to assessing large-scale claims about the nature of reality such as the varieties of physicalism and anti-physicalism make, and it could in the future even be combined with some hitherto unimagined empirical premiss in order to support an exciting ontological conclusion. So even if I am wrong to assume that a realistic interpretation of quantum mechanics is possible, the spirit (though certainly not the letter) of eliminative type physicalism might still live on.

Finally, let me make some remarks about (P3). (P3) is the premiss which, when asserted for a specific value of S, seals the fate of that special science. But it is important to see that it will not always be perfectly obvious whether or not (P3) can be truly asserted for a given value of S. This is because there are two possibilities which could save the

special science (or part of it). The first possibility is that some or all of the properties should turn out to be type identical with PHYSICAL properties, in which case they would after all be mentioned in PHYSICS, though by a PHYSICAL name. But although I shall not defend (P3) -- or any of its instantiations -- further, I take it as obvious that there is little scope for the assertion of special scientific/PHYSICAL type identities. The second possibility is that some or all of the properties should be among those which may harmlessly be added, for whatever reason, to (basically) microphysics to produce PHYSICS. For example, it has been suggested that in order to secure a realistic interpretation of quantum mechanics it is necessary to suppose that the microphysical interacts with the macroscopic, qua macroscopic. (For the suggestion, see Putnam 1965, especially pp.91ff.; for criticism, see Forrest 1988b, p.62) This second possibility is, in effect, the possibility of exploiting the lack of full determinacy which I deliberately introduced into the characterisation of PHYSICS so as not to tie the physicalist's hands too tightly. The existence of this "play" in the system will turn out to be significant when I come, in Part 3, to discuss the threat or otherwise posed to physicalism by so-called qualia. It would also be significant for the important, but enormous, question, which I therefore do not discuss, of the proper response of the

eliminative type physicalist to the propositional attitudes.