FROM PHYSICS TO PHYSICALISM

a version of this paper is published in *Physicalism and its Discontents*
ed, Gillett and Loewer Cambridge University Press 2001

Introduction

Hilary Putnam explains that:

> The appeal of materialism lies precisely in this, in its claim to be *natural* metaphysics within the bounds of science. That a doctrine which promises to gratify our ambition (to know the noumenal) and our caution (not to be unscientific) should have great appeal is hardly something to be wondered at. (Putnam (1983), p.210)

Materialism says that all facts, in particular all mental facts, obtain *in virtue of* the spatio-temporal distribution, and properties, of matter. It was, as Putnam says, "metaphysics within the bounds of science", but only so long as science was thought to say that the world is made out of matter.[1] In this century physicists have learned that there is more in the world than matter and, in any case, matter isn't quite what it seemed to be. For this reason many philosophers who think that metaphysics should be informed by science advocate *physicalism* in place of materialism. Physicalism claims that all facts obtain *in virtue of* the distribution of the fundamental entities and properties –whatever they turn out to be- of completed *fundamental physics*. Later I will discuss a more precise formulation. But not all contemporary philosophers embrace physicalism. Some- and though a minority not a small or un-influential one- think that physicalism is rather the metaphysics for an unjustified *scientism*; i.e. it is scientistic metaphysics. Those among them that think that physicalism can be clearly formulated think that it characterizes a

cold, colorless, unfeeling and uninteresting world and not, they think, the world we live in. In their eyes it is or would be if they thought we had reason to believe it- a doctrine to be feared. In this paper I will argue that these fears have little foundation. (Perhaps they result from traumatic exposure to other doctrines called "physicalism")

Formulating Physicalism

Physicalism is sometimes formulated (e.g. Crane (1991)) as the thesis that all God had to do to create our world is to create its physical facts and laws; the rest followed from these. Fortunately for physicalism's proponents, there are non-theological formulations. The following is due to Frank Jackson (1998):

> (P) Physicalism is true IFF every world that is a minimal physical duplicate of the actual world is a duplicate *simpliciter*.[2]

In (P) a physical duplicate duplicates the laws of physics as well as the physical facts. A minimal physical duplicate duplicates just this and nothing more than is absolutely (i.e. metaphysically) necessary. The idea behind (P) is that once one has fixed the physical facts of our world one has thereby, metaphysically speaking, fixed all the facts. Notice that (P) is itself not metaphysically necessary. There are possible worlds that contain non-physical entities or properties. Minimal physical duplicates of such worlds are not duplicates *simpliciter*.[3]

As Jackson notes, the truth of (P) is necessary to capture the idea that all facts hold in virtue of physical facts. If (P) is false, then there is a world that completely matches the actual world physically, but leaves something out. If there is something over and above the physical that would thus be left out obviously there are some facts that do

not hold in virtue of physical facts. On the other hand, it is not clear that (P) is sufficient for physicalism. The worry is that (P) fails to capture the idea that the fundamental properties and facts are physical and everything else obtains *in virtue of* them. One problem is that (P) does not exclude there being some kind of fact other than physical facts such that minimally duplicating this kind duplicates the world. It helps to add to (P) the claim that there is no other kind of fact and I will henceforth understand (P) with this addition. But this still may not be enough. The worry is that (P) may not exclude the possibility that mental and physical properties are distinct but necessarily connected in a way in which neither is more basic than the other. In this case, it doesn't seem correct to say that one kind of property obtains in virtue of the other's obtaining.[4] Now I am not sure that this is a real possibility, since it may be that if two *fundamental* properties are really distinct then they cannot be connected by metaphysical necessity.[5] But if considerations about the nature of necessity do not rule this possibility out then we must admit that (P) is not quite sufficient for physicalism. However, it seems to me that if we had good reason to believe (P), then, unless we also had some reason to believe that despite (P) mental facts (or some other kind of facts) do not hold in virtue of physical facts, we have good reason to accept physicalism. In any case, even if (P) is not quite physicalism it is close enough to be an interesting claim. So for the rest of this paper I will focus on the credibility of (P).

Some philosophers dismiss physicalism as a serious view because they are skeptical that the notion of a *physical* statement can be adequately characterized. The alleged problem (Crane and Mellor (1990)) is that every way of explicating "physical" will make (P) either obviously false or trivially true (or not consonant with the idea

behind physicalism). If "physical" is characterized in terms of the language of current physics then it is likely that (P) is false since it is likely that the vocabulary of current physics is incomplete.[6] Certainly a complete description of the world in terms of the vocabulary of classical physics is incomplete and there is reason to suspect that there are additions (if not revisions as radical as the replacement of classical physics by quantum physics) that will be made to physics in the future. On the other hand, if "physical" in (P) means facts expressible in the language of the complete physical theory of the world (if there is one), then that threatens to make (P) trivial unless some conditions are placed on what makes a theory "physical." If it were to turn out that to account for certain clearly physical events physicists needed to posit fundamental intentional, or phenomenal, properties then the resulting theory would not be physical.

The most straightforward response to this objection (Papineau (1993a)) is to require that fundamental physical predicates (the atomic predicates of the language in which the complete physical description of the world is expressed) are not mental (i.e. intentional and phenomenal). Given this, (P) implies that mental truths supervene on non-mental truths. This is non-trivial and not obviously false. Other and stronger versions of physicalism can be formulated by adding further conditions on fundamental physical predicates, e.g. no biological predicates, no macroscopic predicates, etc. Another route is to add to (P) the claim that our world has a PC theory and then specify physical vocabulary includes only the vocabulary of the PC theory (plus logic and mathematics). So (P) can be taken as formulating a number of physicalist doctrines with the strongest identifying the physical vocabulary as constructed from the predicates of the true PC theory. In what follows I will generally have this strongest version in mind. [7]

(P) has an important consequence that Jackson calls "the entry by entailment thesis" (ENT). A first, but not quite correct, formulation is:

(ENT) If B is true and $ is the full physical description (including the laws) of the world, then N(If $, then B).

ENT isn't quite correct, since even if (P) is true there will be some true statements, e.g. *There are no spirits*, that are not entailed by $. The problem is that there are worlds at which $ is true but which contain non-physical entities (e.g., spirits) or at non-physical alien properties are instantiated. Jackson recognizes this problem and seeks to handle it by requiring that $ include not just the full physical description of the world but also a statement that says, as he puts it, "that's all." But it is not clear how to formulate "that's all" without going beyond purely physical and mathematical vocabulary.[8] A better approach is to restrict the statements B that appear in ENT to those that are in an intuitive sense "nonglobal" (Witmer (1997)). Global statements are those whose truth-values depend not just on the fundamental facts but also on a collection of fundamental facts being *all* the fundamental facts that obtain at a world. We can make this precise by defining a proposition as global just in case it is true at some world w and for any world w at which it is true, if world w* contains w, then it is also true at w*. "There are no spirits" and "There are exactly 10 conscious beings" are examples of global propositions, while "There are at least 10 conscious beings in region R" and "John is in pain" are examples of non-global statements.[9] Many of the basic predicates of the special sciences are non-global in that statements affirming their instantiations are non-global. However, a true special science generalization may well be global. If so then although it supervenes (and if it expresses a law the law supervenes) on the physical facts it will not be entailed

(and its being a law will not be entailed) by the full physical description.[10]

Jackson thinks that if physicalism is true, then true instances of ENT are analytic and *a priori*. His argument for this surprising claim involves a (controversial) generalization of Kripke's account of how certain statements express necessary truths and yet are only knowable *a posteriori e.g., Water = H20*. In fact, at one time, Jackson thought that statements concerning phenomenal consciousness do not follow *a priori* from the full physical description and that this showed that physicalism is false.[11] But Jackson's argument involves semantical assumptions that are no part of physicalism and, in any case, it can be demonstrated that the claim has exceptions.[12]

It is interesting to compare Jackson-Lewis Physicalism with another influential view in the philosophy of mind that is considered physicalist by some and anti-physicalist by others. I have in mind Davidson' anomalous monism

AM consists of three claims

every event is a physical event

there are no strict psycho-physical or psychological laws

psychological predicates supervene on physical predicates.

None of these claims are free of obscurity,. Davidson's conception of events is one on which a single event can satisfy many descriptions of many different kinds e.g. the pulling of the trigger = the killing of the Archduke = the starting of WWI etc.- and in particular both mental and physical descriptions. Events are particulars with location in time and space (modulo the same kinds of vagueness possessed by objects). All Davidson

says about the individuation of events is that if c and c* have the same causes and effects then they are the same event- but this is not very helpful since causes and effects are themselves events. He also seems to think either that all of an events properties are essential to it or- more charitably- that – talk of whether a particular event could have lacked a certain feature- failed to satisfy a certain description- is not really sensible. This- so called "think"- view contrasts with views of Kim, Lewis, Yablo, on which events are or correspond to instantiations of properties at times (or over durations).

Exactly what Davidson thinks is a law is also obscure and even more obscure is his argument for the non-existence of psycho-physical laws. But it is clear that he thinks those arguments establish that "there are no tight connections between the mental and the physical." In particular, he would certainly reject that the physical state of person and her environment and the laws of physics necessitates her mental state.

Davidson's supervenience thesis is also unclear. He has waffled between the claim that in the actual world (or in each world) no two individuals (objects or events) can differ mentally without differing physically (Weak supervenience) and the much stronger claim that in any possible worlds individuals from those worlds can differ mentally without differing physically (strong supervenience). It has been argued that the latter claim is sufficient for there being psycho-physical laws-whether or not that is right depends on exactly what laws are – but it is certainly sufficient for the physical properties of an individual A and the laws of physics to necessitate A's mental properties. So this reading would seem to conflict with the "anomalous" part of AM. Weak Supervenience is very weak and very trivial if among physical predicates are included spatio-temporal locations. If they are not included then it is stronger but it is very difficult to motivate

why it should hold (Blackburn).

In any case, if AM is understood as denying the existence of tight connections of the sort entailed by Jackson physicalism then it is weaker than it and not physicalist at all since we say that Jackson's formulation is at least necessary for physicalism. Further the event identity thesis is neither necessary not sufficient for physicalism. Event identity isn't sufficient since it may be that there are mental properties whose instantiations are not physically necessitated. It is not necessary since it may be that the modal properties of mental and physical events differ even though the existence and the modal properties of mental events are physically necessitated. Finally if the supervenience thesis is understood as strong supervenience then (if understood as all instantiated properties strongly supervene on physical properties) the thesis is stronger than Jackson physicalism- unreasonably strong- since it means that there can't be – as a matter of metaphysically necessity-a world whose sole inhabitants are souls.!

Fear of Physicalism

I now want to consider some alleged consequences of physicalism:

Physics takes precedence over all the other sciences and other ways of obtaining knowledge.

All sciences (and all other truths) are reducible to physics.

The only genuine properties, events, and individuals are those of fundamental physics.

The only genuine laws are laws of physics. There are no special science laws.

The only genuine causation is causation by physical events; in particular, there is no

mental or other higher-level causation.

Eliminativism: Intentionality, consciousness, rationality, freedom, and norms do not exist.

Sometimes these alleged consequences are taken to be reasons to disbelieve physicalism and sometimes as unpleasant truths that physicalists must learn to live with. I agree that each of the above claims is indeed unpalatable but none is a consequence of (P). Decisively establishing that this is so is beyond the scope of this paper, but a brief discussion may lend it some credibility.

1. It is important to keep in mind that (P) is not an epistemological or methodological doctrine. (P) doesn't imply that physics is epistemologically or methodologically more basic than other sciences. If the claims of current physics should conflict with the claims of one of the special sciences it may very well be that the latter is better confirmed than the former. Nor does the truth of (P) favor allocating federal funds to the proposed super collider over biological research, or to support the arts. (P) doesn't imply that the reductionist methodology of analyzing complex systems into simpler parts that is so successful in physics, is the appropriate method for the enormously more complex systems treated in biology. Nor does (P) imply that scientific methodology is the only, or best, way to acquire knowledge about every topic. It is compatible with (P) that hermeneutics, or *verstehen,* are better methods than experimental psychology when it comes to knowing the minds of our fellow human beings. (P) does require that if that is so, then this fact, like all others, is implied by the totality of physical facts.

2. A reduction of a special science (say biology), or a particular special science theory, to fundamental physics involves systematically locating physical truths that entail

the truths (including the laws) of the special science. (P), of course, does not imply that any special science can be reduced to current physics. More interestingly, it doesn't imply that any of the special sciences can be reduced to completed physics. Since laws are expressed by global statements, (P) doesn't imply that physical truths imply special science laws. And although (P) does imply that non-global truths of the special sciences are metaphysically entailed by the truths of fundamental physics it does not imply that these entailments are systematic or that we can ever locate and know them. Similarly, (P) does not imply that truths that do not belong to any science can be reduced to physics, even though they are implied by statements of physics. It may be, as some have suggested, that limitations in the kinds of concepts that we are capable of entertaining, or the nature of the concepts of the special sciences, or the complexity of the entailments, precludes us from ever knowing the implications that (P) requires.[13]

3. (P) does imply that the only *fundamental* properties, events, and individuals are those of fundamental physics. But (P) doesn't exclude non-fundamental entities belonging to these various categories. Exactly what non-fundamental properties etc. (P) allows for depends on the nature of properties etc. Let me begin with what is sometimes called an "abundant" conception of properties. According to the abundant conception every predicate (or concept) that can be used to make a statement (or is a constituent of a thought) with truth conditions expresses a property, although there may also be properties that are not expressed by the predicates of any language.[14] Further, two predicates express the same property IFF it is metaphysically necessary that they are coextensive. Although the abundant conception is profligate with respect to the existence of properties it is compatible with properties being non-linguistic and (largely) mind-independent, and

with distinct predicates expressing the same property.[15] It is clear that on the abundant conception (P) is compatible with the existence not only of the properties of fundamental physics, but also with many other kinds of properties. On the abundant conception, "is a storm", "is green", "is 5 miles from the Eiffel tower", "is grue", "feels painful," "is a thinking about Vienna," all express properties. It may be that given (P) some of these properties are not instantiable, but that would have to be shown for the specific property.

Contrasting with the abundant conception are so-called "sparse" conceptions of properties according to which only certain predicates express genuine properties. Lewis calls these "perfectly natural properties" (Lewis (1983), Armstrong (1978)).[16] Perfectly natural properties satisfy certain conditions e.g. they involve real similarities, figure in laws and scientific taxonomies, and are causally relevant. Such added conditions presumably disqualify "grueness" (and perhaps some of the other examples mentioned above) from being a perfectly natural property. So whether or not the instantiation of perfectly natural properties that are not expressible in the language of physics (in particular psychological properties) is compatible with (P) depends on whether the existence of special sciences (in particular psychology) and higher level causation is compatible with P; issues to which I will discuss below.

Suppose for now that (P) doesn't exclude the existence and instantiation of perfectly natural properties other than those of fundamental physics. How might they and their instantiations relate to physical properties? If (P) is true, then these instantiations (for the non-global properties) are necessitated by physical facts and laws. But in some cases more can be said. It is widely and plausibly held that some non-fundamental properties are *realized* by fundamental physical properties. There are various views

concerning exactly what is involved in one property instantiation realizing another, but they have in common that if an instance of P realizes an instance of F, then the P instantiation metaphysically necessitates the F instantiation; i.e. any possible world that contains the first also contains the second.[17] For example, the property of being a storm may be realized by various dynamical configurations of physical properties, but is not identical to any specific one or even to their disjunction. For it seems plausible to suppose that storms occur in possible worlds that contain none of our fundamental physical properties (these storms will be realized by alien fundamental properties).

Suppose that events are, or are correlated with, the instantiations (by individuals and times) of certain event constituting properties (Kim and Brandt (1967)). I see no reason why non-fundamental properties cannot be event constituting. For example, the event of a storm striking the coast is composed of various physical events, though it is somewhat vague around the boundaries, but is not identical to any of them. This is clear since the storm's modal properties differ from the modal properties of the events that compose it. That is, it may be that the storm might have struck a bit to the north, but not the case that any of its constituent events might have occurred a bit to the north. Of course, these counterfactual truths must, if (P) is true, supervene on the totality of truths of fundamental physics.[18]

Similar remarks apply to particulars, e.g. particular rocks, trees, and people. If (P) is true then these "higher level" particulars are constituted by the instantiations of physical properties, elementary particles possessing certain states, but are not identical to them. Of course exactly what fundamental physical entities constitute a given particular, e.g. my cat, is a vague matter. The important point is that (P) is compatible with the

existence of my cat even though the cat is not identical to any fundamental physical particles (or mereological sum of them). The reason that my cat is not identical to any sum of fundamental physical entities is that there are counterfactuals true of my cat that are not true of the mereological sum (or any precisifications of the vague sum) that constitute it. Of course, as in the case of events, any such counterfactuals must supervene on the totality of physical facts.

4. Of course, (P) is not compatible with the existence of *fundamental* laws that are not laws of physics. Whether or not (P) is compatible with there being non-fundamental laws depends on exactly what laws are and that is controversial.[19] Jaegwon Kim (1993b) has argued that physicalism excludes there being special science laws. More precisely, he argues that if F and G are non-physical predicates each of which is multiply realized, then the generalization "F's are followed by (or cause) G's" cannot express a law. His argument is predicated on the idea that law statements must involve predicates that are projectible and if physicalism is true then only predicates of physics are projectible, A predicate G is projectable with respect to F (the generalization "F's are followed by G's" is projectable) if F's which are G's confirm that F's are followed by G's.[20] Since he holds that a generalization expresses a law only if its predicates are projectible he concludes (P) is incompatible with the existence of laws expressed by generalizations composed of non-physical predicates. Kim's main argument involves an example. He asks us to consider the property of being Jade which is realized by or equivalent to the disjunction Jadeite or Nephrite. Kim observes that if our evidence consists just of samples of Jadeite and we are interested in a generalization like "Jade melts at such and such a temperature" then it would be foolish to generalize just from this

sample. Of course this is correct, since we know that different molecular structures typically give rise to different melting temperatures. But this example doesn't generalize to other predicates. In the first place, whether or not a generalization is law-like depends on the predicates in both antecedent and consequent. There are some generalizations with "Jade" in the antecedent that are confirmable by their instances; e.g. "Jade is called "Jade" by English speakers." In fact, one can see that (P) is compatible with one's assigning a probability function over certain generalizations expressed in non-physical vocabulary, say the vocabulary of psychology or biology, which allows for their confirmation. That is, predicates of the higher level sciences may be projectable even though they express multi-realizable properties. Indeed, a prohibition against such probability functions would be pig-headed, since we have reason to believe that generalizations that we can express in non-physics vocabulary are sometimes true, or approximately true. We can see that something must be wrong with Kim's argument from the fact that there are well-confirmed generalizations of the special sciences that involve properties that are multi-realized by fundamental physical properties.[21] Perhaps the best and most obvious examples are provided by statistical mechanics; e.g. the probability of the entropy of an isolated body decreasing is negligible.

As far as I can see, (P) does not preclude there being true generalizations couched in non-physical vocabulary (whose predicates do not refer to properties of fundamental physics) that are confirmable by their instances, support counterfactuals, and are entailed by highly informative and simple theories, i.e. have all the usual marks of laws. If this makes them laws (although non-fundamental laws), then (P) does not exclude special science laws.

5. The issue of whether (P) allows for causation by non-physical events and non-physical properties, in particular by mental events and properties, is a vexed matter. Part of the difficulty is that there are no fully acceptable accounts of causation by either events or properties, even for fundamental events and properties. But it seems to me that any adequate account of causation must allow for causation by non-fundamental entities. For example, the storm's striking the coast caused flooding. The storm is a cause and the property of being a storm is causally connected to the flooding. Of course such causation is not fundamental, but will supervene on physical facts and laws. As far as I can see any account that allows non-fundamental physical events and properties to be causes will do the same for at least some kinds of mental events and properties. I won't argue for that here, but I do want to sketch an account of causation, which, though not fully adequate, has this consequence. The account is this: Ms causes Ps* if Ms does not metaphysically entail Ps* and at a time t immediately prior to Ms the following pair of counterfactuals are true:

Ms >t Ps*;

-Ms >t -Ps*.

As far as I can see, (P) is compatible with the pair of counterfactuals that, on this account, ground claims like the storm's striking the coast caused the flooding. And, additionally, (P) is also compatible with the pair of counterfactuals that ground; for example, Fred's wanting a beer causing Fred's body to be in the kitchen. On this account, there may be many events or property instantiations, even ones that occur simultaneously, at different levels that are causally connected to a given event. For example, a neurological event that instantiates Fred's urge may also be causally connected to his body's location.

It has been objected (Kim (1998)) that counterfactual accounts of causation are inadequate, especially in the context of mental causation, since the truth of counterfactuals (like the above pair) is compatible with epiphenomenalism. That is, it is compatible with there being mental properties that are distinct from physical properties to which they are connected by so-called "bridge" laws. If there are no causal relations between these mental properties (or events) and physical properties than the mental properties are epiphenomenal. This view is non-physicalist, since a minimal physical duplicate of a world at which it obtains does not duplicate the mental property instantiations or the bridge laws. Kim's claim is that in the epiphenomenalist world mental to physical counterfactuals may still obtain. But, as we will see in the next section, there is reason to doubt this. In any case, even if it is true it doesn't show that the counterfactuals are not sufficient for causation in a physicalist world. Indeed, if they or some similar account were not correct, then there could be two PC worlds exactly alike in their physical laws and instantiations of physical and mental, properties but only in one are there *genuine (sic) -causal* connections between mental properties and physical properties. These causal connections would, of course, be themselves non-physical, since they don't supervene on the instantiations of physical properties and laws. (If they did, then physicalism could allow *genuine* causal connections). But what are these causal connections doing? They are truly "epiphenomenal"; unneeded and unknowable. The conclusion is that the counterfactual surrogate for causation (or some similar account) is really all we need to account for mental causation, as we know it.

Now, of course, Kim would reject the above reasoning, since what he really wants is for mental properties to make a causal difference over and above causation by physical

properties. But it is obvious from the start that this is something that a physicalist cannot allow. For the physicalist it is enough for our bodily movements to be counterfactually sensitive to our mental states. That the counterfactuals themselves are true in virtue of more fundamental physical facts takes nothing away from that. And one can take comfort in the fact that mental properties are in the same boat as other non-fundamental properties.

The claim that mental causation is compatible with (P) is highly abstract and of, course, doesn't imply that mental causation is implemented by physical processes. Advocates of (P) would like to be in a position to show that certain physical mechanisms do implement mental causation. For example, to explain the causal processes involved when a driver notices a deer running across the road and then applies the brakes. Current psychological research is some distance from coming out with an account, but optimistic cognitive scientists think that the so-called "computational-representational theory of mind" is on the right track. If intentional causation can be shown to be implemented by computational processes then we are on the way towards showing that it is physically implemented since we know that computational processes can be physically implemented; for example, by computers. [22]

6. If P really did exclude the instantiation of mental properties (consciousness, intentionality, rationality) then it would have to be rejected since mental phenomena certainly exist. It is difficult to see how elements of physical reality can add up to intentionality and consciousness. There have been various attempts to show that they do. These "naturalization" projects have not met with much success.[23] But as I mentioned previously, (P) doesn't require that we can ever see how physical statements

metaphysically entail mental statements, just that the entailments hold. On the other hand, attempts to show that (P) excludes mental facts are, if anything, even less convincing. Here I only want to briefly mention two such attempts. One involves so-called "conceivability arguments" which claim that it is *conceivable* that the physical facts be what they are and there be no consciousness, or intentionality, at all.[24] These arguments assume (and sometimes argue for) the claim that, at least in the relevant situations, conceivability implies possibility. However, it can be shown that at most conceivability is defeasible evidence for possibility (Balog (2000)). The second line of argument identifies some feature of mental properties, e.g. alleged normativity, and then claims that it is not the case that physical facts that fail to exhibit this feature can metaphysically entail facts that do. The claim is that two kinds of facts are just too different from one another.[25] While it is not difficult to find philosophers expressing sympathy for this line of thought it is well nigh impossible to find an argument for the conclusion that is sufficiently well articulated to evaluate. In any case, as we will see, given CP if any of these arguments were sound we would be saddled with the conclusion that mental events (or properties) are epiphenomenal, or overdeterminants of physical effects. These consequences seem to me to be so implausible as to cast doubt on the soundness of any argument that claims to show that (P) excludes mentality. [26]

If the previous discussion is on track then reasonable anti-reductionists, humanists, opponents of scientism, and so on, have nothing to fear from (P). But of course the question reamins whether they have any reason to believe (P). I leave that for another occasion.

From CP to P

What would the world be like if (P) were false? First, there would be certain nonglobal contingent facts not metaphysically necessitated by the full physical state description. But how would these non-physical facts, and more specifically the properties and entities that constitute them, be related to physical properties and entities? One possibility is that both physical and non-physical properties are instantiated, but there are no laws or metaphysical connections linking them.[27] This view makes it quite mysterious why it is that mental properties are found associated with only certain physical properties; why, for example, rocks or gasses don't have thoughts. Since there are no laws or metaphysical links that connect these properties such correlations are merely coincidental. Most opponents of physicalism reject this picture and hold instead that mental (and perhaps certain other properties) are emergent in a certain robust sense which involves the existence of *emergent* laws linking them to physical properties.[28] These laws are thought of as vertical since they link the physical state at t with the mental state at t. Emergentist views sometimes also posit horizontal laws linking instantiations of physical and mental properties with each other at different times. Such laws might ground causal relations between mental and physical events (where events are property instantiations at times). If physics is incomplete, then such mental-physical laws might also be required to give full accounts of physical events (or their chances). The important point about these Emergentist laws is that they are not among the fundamental laws of physics or entailed by them and the physical facts. If our world contains them, then God when he made the

world had to make them in addition to the physical facts and laws.

Now I don't think there is much reason to believe that there are Emergentist laws, either vertical or horizontal. But not having a reason to disbelieve is not a reason for belief, so I now want to examine an argument for (P). I know of two ways of arguing for (P). One is by finding reductions of particular higher level facts (properties, events etc.) to lower level, and ultimately physical, ones. While there are some notable reductions (or partial reductions), e.g. of thermodynamics to statistical mechanics, claims of reduction are usually accompanied by a great deal of hand waving. And while each successful reduction provides some reason in favor of (P) each failed reduction provides some reason against it. So this piecemeal approach is far from conclusive. A very different line of argument seeks to establish (P) all at once based on very general considerations about laws and causation. The line of argument I have in mind has been formulated in a number of different ways (for example, McGinn (1982), Peacocke (1979), Papineau (1990), (1993a) and (1995), and Loewer (1995)) The version I will discuss proceeds in two steps. The first step argues that any property (or event) whose instantiation at t (or in time interval d) is causally relevant to a physical event at least nomologically supervenes on the physical state of the world at t. The second step argues that the laws in characterizing nomological supervenience are no more than the fundamental laws of physics. If we further assume that every property instantiation has a physical effect, and that all non-global facts are determined by some collection of property instantiations, then we get (P).

I will assume that the fundamental laws of physics are deterministic. It is not difficult, though a bit messier, to run the argument if the fundamental laws are indeterministic. The main premise of the argument is the deterministic completeness of

physics (PCD):

> (PCD) For any distinct times t and t', the physical state S(t) and the fundamental physical laws entail the physical state S(t')

I will say that a property instantiation M(t) is physically detectable by P*(t') IFF at times immediately prior to t the counterfactuals M(t)>P*(t') and -M(t)>-P*(t') are true. If M(t) is physically detectable by P*(t') and M(t) occurs, then M(t) causes or is actually relevant to P*(t') in accord with the account given earlier. Now, suppose that PCD is true so that at times s prior to t the physical state P(s) and the laws entail that P*(t') (or -P*t')). Then, whether or not M(t) occurs P*(t') (or -P*(t')) will occur. So if M(t) is not connected by law to P(t) it will be physically undetectable. So every property instantiation that is physically detectable at the very least nomologically supervenes on the complete physical state. This is not yet (P). There are two problems: (1) there may be some property instantiations that are not physically detectable. And (2) even for physically detectable properties the most we have shown is that they *nomologically* supervene on the physical state.[29] But among the laws may be laws that are not entailed by the full physical description of the world including physical laws.

As for the first problem, it seems to me plausible that all property instances are either physically detectable or supervene on properties that are physically detectable. But even supposing this is not so, the following principle still strikes me as quite reasonable:

> (U) If some instances of M nomologically supervene on the physical state (or some instances of M supervene on property instantiations that are each physically detectable) then, unless there is some positive reason to think that other instances don't supervene on the physical state, we should suppose that all instances of M

do so supervene.

I know of no reason to think of any instances of a nonglobal property that fails to have instances that are physically detectable (or fail to supervene on property instances that are physically detectable). It follows that it is reasonable to suppose that all property instances nomologically supervene on the full physical state.

Nomological supervenience is compatible both with (P) and with the Emergentist account sketched earlier on which M properties are linked by special "vertical" laws with physical properties (and perhaps by also horizontal laws with physical and mental properties). To establish (P) then I need to show that the Emergentist account is false. I know of no way of conclusively demonstrating this. However, there are some considerations that make the Emergentist picture quite unattractive when combined with (PCD). To begin with, notice that it would be bizarre to suppose that a property like being a rock, or a cloud, etc. is linked by special vertical law (over and above the laws of physics) to the physical state. If that were so there would be a possible world physically identical to the actual world, but where in the actual world there is a rock at a certain location in the other world there is a cloud. It is only for mental properties that the proposal of special vertical laws has any credibility. But even here the result is peculiar. It would entail not merely that zombies (physical, but not mental duplicates of people) are possible, but that either mental properties are completely epiphenomenal or if linked by causal law to physical (or other mental) properties that the instantiations of these properties are pervasively causally overdetermined. The world may be like, that but I think that simplicity considerations suggest that we don't believe it is pending persuasive arguments against (P).

Furthermore, it is arguable that if PCD is true then on the Emergentist account, even with horizontal mental-physical laws, physical events do not counterfactually depend on mental events. So, for example, it will not be the case that the location of my body near the refrigerator will depend on my having desired a beer. Here is why. Suppose that I form the desire to have a beer at t, call this "M(t)," and that there is a physical state P(t) that is linked by Emergentist law to M(t); i.e. "P->M." Suppose now that P(t) leads by the complete deterministic physical laws to P*(t'). Then, I claim that –M(t)>-P*(t') is false. The argument assumes Lewis' (1983) account of counterfactuals and specifically Lewis' account of world similarity. (It doesn't presuppose Lewis account of possible worlds as concrete entities). To evaluate the counterfactual we ask what world (or kind of world) is most like the actual world (where the laws of physics are deterministic and there are Emergentist vertical laws connecting physical with mental properties). For Lewis world similarity is evaluated in terms of two factors: (1) the size of the region in which laws of the actual world are violated; and (2) the size of the departure of perfect match in matters of particular fact. There are two kinds of worlds to consider (where "L" is the conjunction of the complete deterministic laws of physics and "V" is the Emergentist law linking P with M):

W1: -M(t) & P(t) & L & -V

W2: -M(t) & -P(t) & -L & V

If w1 is more similar than w2 to the actual world, then the counterfactual will come out false since in it P*(t') will be true. But if w2 is more similar than w1, then the counterfactual might be true depending on what physical state occurs at t. At first, it may appear that there is a tie in similarity since both worlds involve the violation of a law; w1

an Emergentist law and w2 a dynamical law of physics. But on Lewis' account of evaluating similarity that is not so. The reason is that w2 will differ enormously in matters of particular fact from the actual world, whereas w1 will differ only to the extent required to make -M(t) true. This being so on the Emergentist accounts, and assuming that the fundamental laws of physics are deterministic, it follows that physical events don't counterfactually depend on emergent events. Whether or not I want the beer would make no difference to the location of my body. But, of course, it does. And (P) is no impediment to that being so. Under (P), worlds of type w1 are metaphysically impossible since the physical state P and the physical laws are metaphysically sufficient for M(t). This being so the counterfactual assumption -M(t) requires some violation of the physical laws. If the most similar world in which -M(t) is one at which -Q(t), such that Q(t) leads by law to P*(t), then indeed the counterfactual will come out true. And that scenario is completely compatible with (P) depending on the exact details of the physical realizations of M. So I conclude that (perhaps somewhat surprisingly) if we want the physical to depend on the mental, and if we think that the fundamental laws of physics are complete and deterministic, then we should also accept (P).

The Credibility of PC

I will conclude with a very brief discussion of the credibility of PC. I think it is fair to say that no one thinks that current physics is complete and there is not a consensus among physicists or philosophers on whether PC is true. Some physicists, for example Steven Weinberg (1992), think that we are fairly close to a unified physics that would validate PC. But there is a rather contrary view according to which our world is a much sloppier place than the world of Weinberg's dreams. Nancy Cartwright (1999) is a vigorous advocate of this viewpoint. She has been arguing that the fundamental laws of physics work, to the extent they do, only under very special contrived conditions found in laboratories. If I understand her, the view is that the fundamental laws of physics are woefully incomplete and what laws there are are best understood as containing *ceteris paribus* qualifications. More fundamental than laws are capacities; e.g. the capacity of a charged particle to produce an electromagnetic field. There are many capacities associated with entities and properties at different levels that together determine the tendencies for various courses of events to occur. She suggests that lawful regularities are more or less artifacts of laboratory situations in which interactions are shielded from the various capacities that normally affect them. And she also thinks that emergent *ceteris paribus* laws may appear at higher levels of description that fundamental physics. This is an interesting view and it would definitely be instructive to think through what a world like that would be like in contrast with a world at which physics is complete. Cartwright's main argument for her view is in the form of a challenge to proponents of PC. Her example is dropping a dollar bill from St. Marks tower. Proponents of PC think that the fundamental laws of physics, and to a good approximation Newton's and Maxwell's laws, govern the trajectory of the dollar bill as it floats to the ground, through the use of

those laws to predict the trajectory is not possible due to the complexity of the interactions between the dollar bill and air molecules. But Cartwright thinks this is mere dogma. She seems to suggest that although "F=ma" may hold in specialized laboratory conditions it may be false in this case (presumably that some of the changes in motion of the dollar are not due to forces).

Cartwright's arguments against PC strike me as very weak. They are merely *skeptical* arguments. She doesn't produce a single case in which "F=ma" (and other examples of fundamental laws) fails, but rather claims that extrapolating it beyond controlled laboratory conditions is unwarranted. One would expect that if a putative fundamental law fails outside the laboratory it would be possible to find evidence of it. A Nobel Prize would be in the offing for the discoverer. Even though the fundamental laws (or the equations believed to approximate them) cannot be used to give detailed explanations in complex situations they can be used to give approximate predictions, or be used in connection with statistical models to make statistical predictions, and these are born out. Second, unless we have some specific reason to think that moving outside of the laboratory, or increasing complexity, leads to the failure of these laws then ordinary principles of scientific inference counsel that we should, pending counter evidence, suppose that they hold generally. So I think that while the completeness of physics is a contingent claim, there is now reason to believe that it may well be true and scant reason to think it false.

Conclusion

The journey from physics to physicalism is not an entirely smooth one. As we have seen, the starting place, PC, is plausible, but not obviously true. Given PC, the argument for (P) is fairly straightforward although it requires that one resist epiphenomenalism and nomological overdetermination. Once we have arrived at (P) we still have not quite reached physicalism, since we saw that there may be ways in which (P) could be true, but still not all facts obtain *in virtue of* physical facts. (P) is thus *almost* as credible as PC. I think it fair to say that it is more credible than its denial. And since PC is a scientific claim, physicalism well-deserves Putnam's title of "scientific metaphysics".

Notes.

[1] Putnam himself is no advocate of materialism or physicalism but thinks that metaphysical claims like these involve presuppositions that he rejects. As we will be clear while I like the quote I disagree with his view.

[2] David Lewis (1983) formulates physicalism as the doctrine that "among worlds where no natural properties alien to our world are instantiated, no two differ without differing physically." A property is alien to a world w if it is not instantiated in w and not constructed out of properties instantiated in w. On the assumption that fundamental properties of physics are natural properties Lewis' formulation entails Jackson's, but the converse entailment doesn't hold, at least not without further metaphysical assumptions about natural properties and laws. While both formulations employ possible worlds neither is committed to Lewis' account of possible worlds.

[3] (P) is contingent but, in a certain sense, non-accidental. If (P) is true and if Q is any proposition (true or false) that is compatible with the laws of physics then the counterfactual "if Q were true then (P) would be true" is also true. See Witmer (this volume) for a discussion of this point.

[4] How might this happen? One way would be if properties were individuated by their nomological connections to other properties (Shoemaker 1998) so that e.g. F cannot be instantiated without G's being instantiated. Another way (Russell 1927) would be if fundamental properties possess both categorical and nomological/causal aspects that are metaphysically inseparable and if physical concepts refer to the nomological/causal aspect but not the categorical aspects. Some philosophers inclined to this view think of the categorical aspect as a proto-mental and others as merely non-physical and unknowable. In either case, if the properties referred to in physics possess this kind of categorical aspect it seems wrong to say that all facts hold in virtue of physical facts even

though (P) is satisfied.

[5] According to Lewis's "recombination principle" (1983), if P and Q are distinct fundamental properties then there are possible worlds in which one is instantiated but the other is not.

[6] The trouble isn't that current physics may be false (i.e. physicists are mistaken about the laws) but that it may be incomplete. If its vocabulary were sufficiently complete to specify all the fundamental physical facts and laws then (P) would not be obviously inadequate.

[7] This strong version has the apparent disadvantage of defining physicalism so that it entails PC but if PC but since I think that PC is plausibly true and that it provides the premise for the best argument for physicalism this doesn't strike me as a real disadvantage.

[8] Adding that "$" is the complete description of the world involves semantic notions.

[9] Some global statements e.g. "There are no ghosts" have negations that are non-global while others, e.g. the average age of a ghost is a million years have negations that are also global.

[10] If the reduction of a special science to physics involves the laws of physics and statements of physics implying the laws of the special science then (P) doesn't imply that the special science laws are reducible to physics.

[11] Jackson (1990). Chalmers (1996) develops this argument in great detail.

[12] For general objections to Jackson's claim see Yablo (2000). For an argument that shows that the claim must have at least some exceptions see Balog (2000). Balog shows that statements about phenomenal consciousness must be exceptions to the claim.

[13] McGinn (1996) claims that limitations in our concepts may prevent us from ever solving certain philosophical problems, e.g. How phenomenal consciousness arises out of physical phenomena; Loar (1997) and Balog (2000) suggest that the nature of the phenomenal concepts themselves prevent our finding any identification or realization of phenomenal consciousness by physical properties satisfying.

[14] For Lewis (1983) an abundant monadic property is any set of possible entities.

[15] It is important to distinguish between concepts and properties. Concepts are the meanings of predicates while properties are references of predicates. Distinct predicates can express distinct concepts which unknown to a thinker who understands both express the same property, e.g. is water and is H2O.

[16] Universals can be construed as special abundant properties, or as another kind of fundamental entity, that is correlated with an abundant property. Lewis (1983) points out that the work that is done by universals can also be done by tropes.

[17] Shoemaker (this volume) suggests that a property P realizes F when the set of F's causal powers is a subset of the set of Ps causal powers.

[18] The same holds for other accounts of events e.g. Lewis (19xx), Yablo (19xx), and Davidson's rather different account. An interesting point about Davidson' account is that his event monism, all events are physical events, is neither necessary nor sufficient for (P).

[19] It is plausible that special science laws or law statements include a *ceteris paribus* qualification. There is no fully satisfactory account of exactly what this qualification comes to, but that shouldn't prevent us from recognizing that there are plenty of examples of special science statements that are thought to express laws.

[20] Following Goodman (1979) a hypothesis "Fs are followed by Gs" is projectable (G is projectable with respect to F) IFF positive instances of the hypothesis confirm it (raise its credence and the credence of unobserved positive instances).

[21] More extensive responses to Kim's argument are in Block (1997), Fodor (1997) and Antony and Levine (1997).

[22] See Rey (1997) and (this volume)

[23] For a survey of some recent attempts to naturalize intentionality see Loewer (1997).

[24] Conceivability arguments directed at showing that phenomenal consciousness fails to supervene on physical facts go back to Descartes and have recently been revived by Kripke (1980), Jackson (1998), Chalmers (1996) and McGinn (19xx). They have, in my view been decisively refuted by Balog (1998) and (2000), Loar (1997), and Hill and McLaughlin (1999).

[25] Davidson's (1969) argument against the existence of strict psycho-physical laws and more generally there being any "close connection" between physical and intentional concepts seems to be of this sort but it is very difficult to know since it is very difficult to say exactly what the argument is.

[26] The arguments that are intended to show that (P) excludes mental facts might actually better be understood as showing, if anything at all, that we cannot see how physical statements metaphysically entail mental statements.

[27] This view is sometimes attributed to Davidson (1969) who on the one hand holds that there are no psycho-physical laws, but on the other hand accepts that the mental weakly supervenes but not strongly (as required by (P)) on the physical.

[28] The Emergentist laws may be "deterministic" in that the physical state completely

determines the co-temporal mental state or "indeterministic" in that the physical state only partially determines the mental state or determines the chances of various mental states.

[29] Witmer (1997) commenting on earlier versions of this argument made this point about Papineau (1993a) and Loewer (1995).