Physicalism Meets the Knowability Paradox

David M. Lewis II

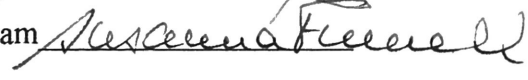
University Undergraduate Research Fellow, 1995-1996

Texas A&M University

Department of Philosophy

APPROVED

Undergraduate Advisor 

Exec Dir., Honors Program 

Thesis Abstract

Physicalism meets the Knowability Paradox. David M. Lewis II (Jonathan L. Kvanvig), Philosophy, Texas A&M University

This research is an attempt to shed light on a serious difficulty for contemporary physicalism. The methodological approach includes both historical critique and logical analysis. The first section shows that one type of physicalism is superior to other versions. In the second section, I explain the knowability paradox and argue that it is entailed by this superior physicalism. Finally, I argue that there are no adequate solutions to the paradox, and that the one possibility to rescue the physicalist from the paradox fails, making the prospects for physicalism poor.

"There's physics, and there's stamp collecting."
-- Rutheford

In the following pages, I attempt to shed light on a serious difficulty for contemporary physicalism. The first section shows that one type of physicalism is superior to other versions. In the second section, I explain the knowability paradox and argue that it is entailed by this superior physicalism. Finally, I argue that there are no adequate solutions to the paradox, and that the one possibility to rescue the physicalist from the paradox fails, making the prospects for physicalism poor.

I Physicalisms

What is physicalism and why should anyone care whether it is true? People rarely question whether there is such a view[1], however, there is a great deal of confusion surrounding the first half of this question. So as a place to start we can say that physicalism is the view that everything is physical. The view has a grand tradition in the history of philosophy, descending directly from materialism, the view that all is matter. Physicalism, however, is a more contemporary and fashionable version. The name 'materialist' shall refer to earlier, non-contemporary adherents of a view which is best thought of as a precursor to the view to be discussed below as substantive physicalism.

There are of course, many reasons why we should care whether such a view is the right one. First, the truth of physicalism entails the denial of many claims of orthodox religion. Each of the three major monotheistic religions is committed to the existence of a

[1] for exceptions, see Crane and Mellor, 1990

non-physical God. Supernatural causation is particularly offensive to physicalists, who make much of the causal simplicity of a closed physical system. In addition, orthodox versions usually support claims of immortality by postulating souls which are independent of physical reality (on this count most eastern religions are also inconsistent with physicalism). Of course, let us not forget spirits, angels, demons etc. which can have no place in a physicalist's ontology.

Not everyone is interested in religious questions; however, the implications of a physicalist view for philosophy are astounding. Idealists have long claimed that if everything is physical then skepticism quickly follows. It has also been widely contended that if everything is physical, then everything is governed by physical laws (or whatever determines the behavior of strictly physical entities). Hence everything is determined, excluding the possibility of a free will. This holds enormous implications for ethics, especially those areas which seem to require freedom of choice.

The truth of physicalism would have several implications for contemporary epistemology, philosophy of mind, psychology, theoretical mathematics and of course physics itself. The ontologies which members of these disciplines will be allowed to utilize in describing and explaining would of course be limited by the success of the physicalist program; the norms for acceptable modes of inquiry in any of these disciplines, perhaps in all disciplines of any substance, would be determined by the physicalist axioms as well. Physics would bear the burden of being the paradigm of rational inquiry, the sole legitimate among a host of bastard disciplines. Currently, many psychologists assume either explicitly or implicitly a physicalist view as foundational to their work. That is, much

of the philosophy of mind is carried on under the very banner of physicalism and psychologists, in an effort to further a friendly philosophical cause and establish their own position among 'physical' scientists, through either reductionism or eliminitavism are even now attempting to turn psychology into a highly sophisticated neurophysiology.

And so we see that the question of the truth of physicalism is no trivial concern. The physicalist offers a view of the fundamental nature of reality which, if true, entails the falsehood of many deeply entrenched beliefs, both popular and academic. Yet there are serious questions to be asked with regard to the nature of the view itself. It is not difficult to imagine what the physicalist intends to say about the world, we know at least what kinds of things are supposed to be physical, and which kinds are not; the problem, as will be demonstrated in the following pages, is that nobody has found a coherent way to state the view.

Physicalism, simply phrased, is the view that everything that exists is physical. The view dates to the Presocratics but began its assent to its current level of dominance during the Enlightenment. Its ancient adherents, most notably Democritus and later Lucretius (Their view was called atomism, the view that there is nothing but atoms and the void, or space), were taken with an agenda which was substantially anti-religious. Both desired to free their fellow man from the fear of religion: get people to believe your story about how lightning *really* happens and they'll no longer think Zeus is punishing their infidelity every time they see a thunderstorm. This early atomism claimed that nothing existed but atoms, and that everything could be explained in terms of the interaction between those atoms. It became necessary (in order to make the view seem plausible) to provide an explanation of

ordinary events in terms of these interactions. Indeed, clever theories were devised. To account for perception for instance, it was said that thin, ethereal atoms traveled through the air, penetrated the eyes or ears and eventually reached the soul atom.

With the Enlightenment, however, and the exponential growth in scientific knowledge, this chore of explanation began to look much less daunting, and the stories involved became much more believable. As increasing numbers of phenomena were explained by the new science, there grew a healthy disdain for the mysterious and illusory explanations that remained. Hobbes published a serious work expounding the thesis that all phemomena could be so explained, including even the most intractable problems of human consciousness. And so it was that the physiclalist's image was transformed from medieval heretic to modern sage. Surprisingly, though, in spite of the manner in which the materialist position has been historically buttressed by scientific advance, the relationship between science and physicalist theses has not always been congenial.

As we have noted above, the success of the science served to render the explanatory power of the supernatural and incorporeal less and less essential to understanding the world. Anti-scholastic sentiment made conditions ideal for the advent of an ontology that disposed of all occult powers. In so doing, the newfound lust for science was satiated by adjusting the ontological paradigm to exclude anything beyond the grasp of the new science: materialist metaphysics and the new science, it was thought, would make great bedfellows. Science was to be entrusted with producing material explanations to replace the mysterious and materialism would preside over the

dismantling, predicting that eventually these explanations would encompass all there is (i.e. "Everything is matter").

It was understood that a meaningful definition of matter (other than "the stuff of which everything is made") was necessary to safeguard the view from the charge of triviality. If nothing was known of matter but that everything was composed of it, then the proposition that everything is matter would be true but it would tell us absolutely nothing about what does and does not exist. According to this formulation, a ghost could exist, would therefore be material, since everything is made of matter, and would fall neatly within the materialist ontology!

The goal of any ontology is to tell us what there is. The Materialist/Physicalist tradition is different, however, in that it is focused largely on a negative ontological task: to say precisely what there is not. The tradition has always disguised this task by making the apparently positive assertion that everything is physical. The movement has been marked by the need to deny that there exists anything like ghosts or souls, angels, gods, or anything else that doesn't lend itself to objective inquiry. Because of the intimate relationship between these ontologies and the physical sciences, however, and indeed because the physical sciences have seemed so obviously unassailable as methods of acquiring knowledge for the last few centuries, allowance has always been made for the existence of any entity of which science may have need when making its hallowed explanations. What was needed, therefore, to get the materialist project off the ground was a definition of material which included all scientifically respectable entities, yet excluded ghosts.

The history of philosophy records many auditions to play such a coveted role, some by players with very impressive credentials (and managers). Hobbes included such an aside in his materialism. According to Hobbes, what is material is whatever has volume. This is reminiscent of Descartes, who defined the material portion of his dualism as extended and non-thinking. Hobbes proceeded to make his case as noted before, but what eventually discredited his view was not the implausibility of a scientific account of the special nature of humanity (this is a goal to which many physicalists and scientists still aspire), but rather, a new revolution in the discipline of science itself. Newtonian mechanics postulated several entities (which were very helpful in explaining little things like planetary motion) which weren't ghosts, but still didn't meet the Hobbesian/Cartesian criterion for materiality. The most significant of these were absolute space and time, and of course, gravity. None of these have volume or are extended, but they each remain indispensable in Newtonian theory.

Other definitions have also been set forward. Peter van Inwagen offers a definition of material in his theory of material beings which does justice to our intuitive conception of material: "A thing is a material object if it occupies space and endures through time and can move about in space (literally move about, unlike a shadow or a wave or a reflection) and has a mass and is made of a certain stuff or stuffs. Or, at any rate, to the extent that one was reluctant to say of something that it had various of these features, to that extent one would be reluctant to describe it as a material object."[2] This definition was intended to serve as a criterion for material objects, a concept related to the notion of 'substance' in

[2] van Inwagen, p. 17

metaphysical discourse. Clearly, many of the entities to which current scientific theory helps itself (fields, quanta, etc.) are not in the extension of this sort of predicate.

Of course, Van Inwagen claims a material atomism, "every material thing is composed of things that have no proper parts, "elementary particles" or "mereological atoms" or "metaphysical simples," as a premise for the above theory of material objects. If this were meant as a criterion to distinguish material from immaterial, there would naturally be a further explication of the nature of these metaphysical simples. Without such a definition, Leibniz' monads would be just as good a candidate as our quarks, or whichever more fundamental particles will be discovered by later physics.

Others, with implicit reference to "Elizabeth's Question," a discussion in metaphysics concerning the apparent impossibility of causal interaction between the material and immaterial, suggest that a thing is physical if it is causally effective. This formulation can only do the work desired by the physicalist if it is buttressed by the additional assumption that physical things are the only things that exist in causal relationships. The addition of this assumption makes the causal criterion viciously circular as a means of designating physical entities. The only available remedy is to stipulate an additional criterion by which to separate the physical from the non-physical, bringing us right back where we started. Also, it happens that the issue of whether physics is in need of causation is debatable. If not, then on this definition the things that physics quantifies over are themselves not physical!

Indeed today physicalism is still entrenched in the same difficulty of specifying just what is meant by 'physical' and hence what sorts of things the physicalist ontology will

really countenance. At this point that there no longer remains any such animal as physicalism "simply put." The contemporary physicalist is painfully aware of his forebear's travails and the price at which his view has been inherited; however, there are two ways in which a physicalist might respond to this difficulty. The first is, like Hobbes, to offer an intelligible account of what is meant by physical in terms of current scientific understanding. Call this substantive physicalism (SP for short). The other option, to be discussed later, involves a completeness claim for science (or some future science) and defines 'physical' entities as all and only the kinds of entities recognized by the specified scientific theory. This view we shall call methodological physicalism (MP).

Consider first the substantive physicalist. In the first three centuries of the history of modern materialism, most of the view's advocates, like Hobbes, offered explicit definitions of 'material' in terms of qualities which were supposed to designate all and only those entities which counted as material and which would exclude all the undesirables. There are few contemporary philosophers who continue to espouse such a courageous, if not terribly prudent view. It is courageous, because it shows a determination to set forth a non-trivial view, an appreciated gesture, to be sure. This course is imprudent in that it must march confidently on in spite of the fact that every historical attempt to so formulate the materialist/physicalist view has been undermined by the very success of science which motivates the view in the first place. These historical events constitute a powerful inductive argument against any new attempt to grab the same golden ring.

The nature of this argument is simple. In the case of substantive physicalism, the defeater arises as follows: We reason inductively from the history of the scientific enterprise, more specifically, the recurrent pattern of scientific revolutions, to the claim that the nature of future scientific explanation of phenomena will be radically different from that of current scientific theory. We notice that with past revolutions, this change has been accompanied by radical changes in the ontology needed to support a new theory. So, if any physicalist view includes a definition of physical which is couched in terms consistent with current science, the above considerations provide a good reason to believe that future scientific theory will postulate the existence of entities which do not fit that definition, thus undermining the view in question. Even if we find the view useful for some present purpose, in light of the inductive argument for scientific revolution, we cannot be sure the view is true. Hence there is a defeater for any view that crams ontology into a mold fashioned by current scientific theory: history tells us that science will almost certainly evolve to postulate new entities which will not fit the prescribed ontology.

So what can the physicalist to do in the face of such a specter? To understand the typical response, recall what has been observed about the substantive physicalist so far. 1. Substantive physicalist views are motivated and supported by arguments from the success of science. 2. History shows that such views have invariably been formulated in scientifically respectable terms, consistent with the language and ontology of contemporary science. 3. Any substantive view is quite likely going to be demonstrated false by science, its champion in the first place.

This is a very embarrassing thing for physicalists, yet they are not without options. I can think of three: First, they could decide that physicalism is the wrong view after all. Curiously, this path is rarely chosen. Second, they could also claim that physicalism is not wrong, but has just been stated incorrectly. They could subsequently offer a new definition of what physicalism really means which is consistent with the new scientific progress. There are few physicalists who take this route. But there is a third, and presently much more popular option. The physicalist could admit that their most fundamental commitment just *is* that their ontology should be in accord with science, and claim that what there is exactly what a true and complete physics would say there is. This is the combination of the Quinean dictum "To be is to be the value of a bound variable" and the coronation of physical theory as the complete description of the world. It is this view that we have chosen to label methodological physicalism.

I have mentioned that MP is presently much more popular than SP. This is in fact so much the case that it is difficult to find contemporary physicalists who do not state their view in terms of a completeness claim for physics. We will examine the wisdom of pinning one's ontology on a mode of inquiry later. First we must examine how much of a difference there really is between the substantive physicalist and her methodological counterpart. Both SP and MP claim to give an account of what there is. We can also be sure that each view, when held at any given time, will be consistent with the truth of contemporary scientific theory. The difference is that a given version of MP will also be consistent with future scientific theory, even if it is radically different from current science, while a concurrent version of SP will have to be altered or replaced altogether to

compensate for the changes in the scientific landscape.

We cannot blame a substantive physicalist for changing his view in light of scientific revolution; we have noticed, after all, how embarrassing it must be to have your view undermined by its very foundation. Likewise, then, we certainly cannot blame a methodological physicalist for protecting his own view from such a fate. Is it really fair, however, for the substantive physicalist to adjust his view in such a fundamental way at each revolutionary turn, and yet continue to call it by the same name as always? There have been in fact dozens of versions of SP which differ in important ways but which are all alike in the most important way: Every time an SP is revised because of scientific revolution it is always altered in a manner parallel to the scientific changes which provided the impetus for change.

This suggests that there is a deeper psychological commitment to science than to any one ontological view. If SP takes all of its cues from scientific progress and is always adjusted to fit whatever scientific theory is current, then it, like MP, will offer a true, complete and intelligible description of what is only when science has in fact been made true and complete. The upshot of all this is that if SP is able to make these repeated adjustments to scientific change while ostensibly maintaining the same identity, then SP is at bottom MP, disguised by some very fancy (though predictable) footwork.

And so it seems that the first category of physicalist views collapses into the second. If this is the case, then all physicalist views entail the completeness of science (or future science) and claim that nothing exists beyond those entities posited by the scientific ontology. Regardless of the strength of this claim, however, it is a matter of empirical fact

that most contemporary physicalists are of the methodological variety. The preference is so dominant that Crane and Mellor lumped the view that I have called SP in with Materialism, describing the lot as attempts to limit physics a priori by placing restrictions on 'physical' similar to those examined above[3] . Crane and Mellor appear to make the claim that none of these represent a viable option because physics has undermined the view, and that physicalists have "lost their metaphysical nerve. No longer trying to limit the matter of physics a priori, they now take a more subservient attitude: the empirical world, they claim, contains just what a true complete physics would say it contains." Their opening remarks summarize much of what has been said above:

> Many philosophers are impressed by the progress achieved by physical sciences. This has had an especially deep effect on their ontological views: it has made many of them physicalists. Physicalists believe that everything is physical: more precisely, that all entities, properties, relations, and facts are those which are studied by physics or other physical sciences. They may not all agree with the spirit of Rutheford's quoted remark that 'there is physics; and there is stamp-collecting', but they all grant physical science a unique ontological authority: the authority to tell us what there is.[4]

Crane and Mellor go on to note several difficulties with such a view, most notably that it is very difficult to understand which sciences are physical sciences and why they rate such a distinguished appellation, and also that physicalist theories tend to say either too little, in which case they appear vacuous and trivially true, or too much, causing them to appear obviously false. I do not wish to focus on these arguments, however, but rather on the implications of the physicalist's claiming the completeness of physics and their simultaneous entrusting of physics with 'unique ontological authority'.

---

[3] Crane and Mellor, 1990
[4] Crane and Mellor, 1990

The best way to understand this claim is to realize that if physics is completed, then it has answered every question it set out to answer about the way the world works and what things it contains. The physicalist claim, then, is that nothing else exists but the things referred to in that description. No one thinks that contemporary physics is complete, so physicalists of this variety must of course be thinking of a future physics which will have reached this zenith. So according to the physicalist, physics paints a picture of the world which explains everything that is or was... and no other picture of the world is correct.

There are a variety of consequences of this view, including the subjugation of all other disciplines to the ontological fiat of physics. Psychologists cannot tell us what there is unless it can somehow demonstrate that psychology as a discipline is itself 'reducible' to physics.[5] Of immediate interest to this project, however, are the epistemological consequences of the physicalist claim. At first blush, if nothing exists but what falls within the purview of a particular science, then the theory associated with that science implies every true statement about the world. Any statement which is not so implied by the theory in question simply would not be true in any significant sense of the word.

It would be helpful to more clearly understand the manner in which theories are related to ontological commitments. A theory, viewed axiomatically, is a set of propositions. These propositions paint the picture alluded to above by describing events or phenomena. Like all propositions, those of a theory are either true or false. The propositions of a given theory can be true only if each of the entities referred to by those propositions actually exists. So ontological cooperation is a necessary condition for the

---

[5] for a telling discussion of the problems involved in understanding what is meant by such a reduction, see Crane and Mellor, 1990

truth of any theory. This is what is meant by ontological commitment, and this is the spirit behind the Quinean dictum "To be is to be the value of a bound variable" (in formal theories, bound variables are quantified terms that represent things in the world). There are ways to formulate MP without positing an ultimate theory, of course, but my arguments in later sections of this paper are not affected by the details of specific MPs. Rather, my points hinge on the fact that every MP includes the completeness of science and the exhaustive nature of scientific ontology.

Now MP claims that when science is completed (i.e. when it has passed judgment on every possible proposition, declaring it true or false) nothing will exist beyond what is referred to by physical theory. It is a direct consequence of this claim that there will be no truths besides those which are either part of the physical theory itself or are implied by that theory. To demonstrate this fact, let us consider a set of truths about non-physical entities or events. These are propositions which are true but are not implied by scientific theory. If physicalism is true, however, then there is a scientific theory which is true and which implies every true statement about everything that exists. So our set of propositions must either contain truths about physical entities or not contain truths after all. Further, if any proposition is implied by a scientific theory which is itself known, then that proposition can be known. In short, physicalism entails that if anything is true, then it can be known via completed scientific inquiry. This position will soon be seen to be quite detrimental, so we should understand that physicalists are not usually averse to this claim. They most often understand, acknowledge and welcome their commitment to the position that science will make all truth accessible for human knowledge.

II The Knowability Paradox

This last proposition, namely that all truths can be known, is the premise of the knowability paradox. Later I will attempt to demonstrate that physicalists are committed to this paradox arising from truth and knowledge claims, but it will first be helpful to examine the structure of the paradox itself. The knowability paradox arises from two claims, each of which seems reasonable of itself, yet together imply a contradiction. The first of these claims is (1) that if any proposition is true then it can be known (that is, all truths are knowable). This sentence can be translated into symbolic logic as follows: (1)$\forall x\ (x \rightarrow \Diamond Kx)$ where '$\Diamond$' means 'possibly' and K is interpreted 'it is known by someone at some time that' (note that all quantification in the following formalism is substitutional)[6]. The second is (2) that there exist unknown truths, or that some truths are not known. This sentence is formalized as (2) $\exists x\ (x\&\sim Kx)$ where '$\exists$' is a quantifier meaning 'There exists an (x) such that' and $\exists$ is ranging over the domain of propositions. If (2) is true, then there must be an instance of it, which we will call 'p' which is both true and not known, i.e. (3) $p\&\sim Kp$. Since (3) is true, our first premise (1) must apply to it, producing (4) $p\&\sim Kp \rightarrow \Diamond K(p\&\sim Kp)$. This conditional is eliminated by Modus Ponens, so it can be known that p is true and yet unknown, (5) $\Diamond K(p\&\sim Kp)$. Knowledge is distributable. That is, you cannot know a conjunction unless you know its conjuncts, so (5) implies the claim that it can be known that p and it can be known that p is unknown,

---

[6] A substitutional interpretation of the quantifiers allows the domain of quantifiers to range over propositions rather than objects, see Haack, 1983

or, (6) ◊(Kp&K~Kp). Since it is impossible to know a falsity (knowledge implies truth), it follows that it is possible that p is Known and p is unknown at the same time, (7) ◊(Kp&~Kp), which is clearly impossible.

KIT (Knowledge Implies Truth) follows directly from the definition of knowledge (knowledge is minimally defined as justified true belief). The truth of KIT has been recognized since Plato's Theaetetus as a basic and obvious requirement for knowledge. If anyone wished to deny this principle, there would be several undesirable consequences. For example, it would likely be possible to know contradictory propositions.

The second rule of inference, K-&E is a rule about the distributivity of knowledge claims. It is quite intuitive that if a person knows the conjunction, roses are red and violets are blue, that person must also know the propositions stated by the individual conjuncts, i.e. he/she knows that roses are red and knows that violets are blue. There are some peculiar theories of knowledge in which this rule is not allowed, but there are independent theoretical costs associated with such restrictions. This problem lies beyond the scope of this discussion, however, and we will be content to note that the paradox can be reformulated to render the K-&E rule inessential, so we need not be concerned.[7]

---

[7] Timothy Williamson offers such a variation in which we are to assume that a proposition p is 'completely unknown' iff no conjunction is known of which p is a conjunct. No one can know the following, however: p and it is completely unknown that p. For if someone did know that conjunction then p would not be completely unknown. Therefore, (1)p((Kp implies that there are no completely unknown truths. This argument relies on KIT but not knowledge distribution. More formally, p is completely unknown iff ~∃qK(p&q). so there is a world W where K(p& ~∃qK(p&q)) Let r= ~∃qK(p&q) so in W, K(p&r). If so, then in W, ∃qK(p&q) by existential generalization on K(p&r). By KIT, K(p& ~∃qK(p&q)) implies p& ~∃qK(p&q) which in turn gives us ~∃qK(p&q), thereby yielding a reductio of the assumption that p is completely unknown. "Verificationism and Non-distributive Knowledge," Australasian Journal of Philosophy 71, (1993), pp.78-86

The argument is formalized as follows:

(1) $\forall x(x \rightarrow \Diamond Kx)$	Premise
(2) $\exists x(x \& \sim Kx)$	Premise
(3) $p \& \sim Kp$	Existential Elimination on (2)
(4) $p \& \sim Kp \rightarrow \Diamond K(p \& \sim Kp)$	Universal Instantiation on (1)
(5) $\Diamond K(p \& \sim Kp)$	by Modus Ponens on (3) and (4)
(6) $\Diamond(Kp \& K \sim Kp)$	by K-&E on (5)
(7) $\Diamond(Kp \& \sim Kp)$	by KIT on (6)

A normal reaction to a first reading of the above derivation is to believe that it must be the result of some sort of clever sophistry, that the paradox is generated by nothing but a bit of pseudo-logical slight of hand. It is important to examine at this point, then, exactly what it is on which the above derivation depends. The derivation is perfectly valid, and the contradiction is derived from the two original premises using standard rules of logical inference, with the addition of the two epistemic rules, KIT and K-&E.

The only remaining avenue for quibbling then, is to object that one or both of the premises ought not be accepted. There might be any number of reasons for rejecting the first premise, but as we have seen, this avenue is not available to physicalists and other people whose ontological commitments have made sacred the knowable truth. At this point, it will be helpful to understand just how it is that one could be committed to a paradox. It must also be admitted that the term paradox is perhaps a bit misleading. The problems arising from knowability are different from other paradoxes in a crucial way. The Liar's paradox, for instance, (arising upon consideration of sentences such as 'I never tell the truth.') seems to spring from the semantic structure of language itself and applies quite generally, that is, it applies to people of all philosophical positions equally.

The knowability paradox is quite different, however. As mentioned above, the paradox involves a necessarily false claim which is derived from two plausible premises. The interesting fact about these premises is that they are each factual claims, i.e. claims about reality. If reality is such that either (1) we really can't know all truths or (2) we, or something else, does or will actually know all truths, then the proof demonstrated above that I called a paradox shouldn't bother anyone at all. The derivation of an impossible claim from the two premises in question would serve only to demonstrate that one of the premises must be rejected. The knowability paradox only merits its name, then, because there are many people whose philosophical commitments will not allow them to reject one or another of the premises. For such people, however, the paradox presents quite a difficulty: they must resolve the paradox, or concede that their view is untenable.

One might ask at this point, "Why shouldn't the clever physicalist simply reject the second premise, that there exist some unknown truths?" Such a rejection would amount to an appeal to omniscience of some variety, whether the existence of an omniscient being, or a community of cognizers or perhaps a set of such communities that jointly knows all true propositions. (It would suffice for all truths to be known at some point in time, i.e. they need not all be known at the same time.) This sort of omniscience claim would not be a problem for a theist, of course, but the paradox could easily be reformulated with a new interpretation for the K-operator such as 'it is known by an ordinary person at some time that' so as to make the existence of God irrelevant to

concerns related to the paradox. Anyway, a theistic physicalism would be a rare specimen indeed! Seriously, though, the paradox so reformulated would be at least as offensive to the physicalist as the original version because they are committed to knowability in precisely such a way that truths would be knowable for ordinary people (that is, via the ultimate scientific theory).

So all that remains is the possibility of humanity's knowing all truths (or reaching the state of having known all truths). However, to reject the second premise on these grounds would require an unusual optimism. One can imagine many questions, such as whether there is an even or odd number of quarks in the universe, which obviously have true answers, but which are so tedious or pointless that they will never be answered. Even if someone did answer one of these questions, there would be more, and life is too short to answer them all. The ultimate physical theory must provide the tools to make discovery of each of these trivial truths possible, hence the commitment to knowability; to think that such a theory will ever actually be used to learn about these truths, or even about obscure historical truths, seems a little silly.

Once we understand these considerations, then, we understand that the paradox spells trouble for any view that entails the universal application of a factive operator to all true propositions. Factive operators are those which are applied only to true propositions. These operators include, but are not limited to, the following: 'It is known that', 'It is conclusively verified that', 'It is proven that', etc. The paradox demonstrates that if a theory implies that possibly every truth satisfies such a predicate, that theory also entails that every truth actually satisfies the given predicate. Physicalism entails that any truth is

knowable, because, according to physicalism, nothing exists that cannot be explained in physical terms. Therefore, physicalism either entails a contradiction (the conclusion of the proof for the knowability paradox) or physicalism entails that every truth is actually known.

I do not intend to argue that the Knowability Paradox constitutes a positive argument for the existence of any of the entities which are denied by a Physicalist ontology. The moral to be learned is that the physicalist's temptation to pin their ontology on a theory or a mode of inquiry like science is an invitation to a dead end: If the knowability paradox cannot be solved, any physicalist view which specifies its ontology according to science is false. Note that this problem applies equally generally to other ontological and semantic 'isms' as well: The viability of several views (including verificationism, many versions of anti-realism, according to which truth is defined as what is, will be, or is possibly known, and naturalism, "the realist ontology that recognizes only those objects required by the explanations of the natural sciences"[8]) depends upon resolution of the knowability paradox. The fortunes of any view in metaphysics, epistemology, or semantics which implies the knowability of all truths hinge on some forthcoming solution of the knowability paradox.

Given costs as high as these, one would expect to witness a tremendous effort to resolve this little difficulty. Surprisingly, however, there is not a huge literature on the subject. Perhaps too few people realize the breadth of the paradox's implications for

---

[8] as defined by Ed Zalta "Naturalized Platonism vs. Platonized Naturalism"

philosophical views. There have, of course, been some attempts to deal with the problem, however I will argue that none of them is sufficient to deliver the physicalist from the stake of omniscience. Some attempts to resolve the paradox are partial and some are general. General attempts are those which, if successful, would demonstrate that the two premises do not actually entail the conclusion or that knowability of each truth does not really imply knowledge of all truths. An example of this type is the suggestion that we abandon classical logic in favor of some other logic in which the paradox fails. Partial attempts are those which purport to find a way out of the paradox for a particular view, usually by either altering the view or reinterpreting the paradox itself.

Dorothy Edgington offers such a partial attempt at resolution.[9] She argues on behalf of verificationism (according to Edgington, the view that all truths are knowable) that though the first premise is not compatible with the existence of unknown truths, it is compatible with the existence of actually unknown truths. To formalize her argument, Edgington introduces an Actuality operator to be interpreted "in some actual situation." The first and second premises, then, read $1.\forall x(Ax\rightarrow\Diamond AKx)$ and $2.\ \exists x(x\ \&\ \sim AKx)$. Edgington then shows that nothing disturbing follows from these premises and thus claims to have exorcised the omniscience monkey from the verificationist's back. It is not likely that she has been successful however, and even if she were, her solution would not be available to the physicalist.

My remark that her success seems unlikely is in reference to Williamson's discussion of Edgington's solution (Williamson, 1987). Williamson notes that her

---

[9] Edgington, pp. 557

application of the Actuality operator within the context of the knowability claims depends on the possibility of non-actual knowledge of the actual. He examines as many possibilities for such knowledge as are readily obvious and concludes that none are very promising. Williamson is certainly correct in claiming that Edgington's solution must be regarded as unsuccessful pending an as yet unavailable account of such non-actual knowledge. The physicalist should not hold his/her breath, however, for the nature of the physicalist commitment to knowability (or verificationism, if you wish) makes this solution unavailable. The physicalist is committed to the knowability of each truth because each truth is implied by physical theory. Derivability from a theory insures the possibility of knowledge per se, not just a special type of knowledge with or without an actuality operator.

Some general attempts, such as the denial of the knowledge distribution principle, were discussed earlier in the analysis of the logical structure of the proof itself. As observed before, the proof is perfectly valid in classical logic. Timothy Williamson, however, has suggested that one general solution would be to abandon classical logic in favor of intuitionist logic.[10] Accept intuitionist logic, the line goes, then show the argument to be intuitionistically invalid, and there are no more worries. It is not quite that simple. First, the problem is only partially blocked by acceptance of intuitionist logic. It is true that knowability (1) does not actually entail the feared $\forall x\ (x \rightarrow Kx)$ in intuitionist logic, but this is not the end for the knowability paradox.

---

[10] Williamson, 1992 he actually hints that it might be the only way

The paradox, as formalized above, is the derivation of a necessarily false claim from two plausible premises. The argument could be restructured to produce the conclusion $\forall x\ (x \rightarrow Kx)$, but this is a matter of strategy, not fundamental to the paradox itself. The conclusion is derived by reductio. The distinctive treatment of negation within intuitionist logic disallows proof by reductio, however, so all that can be derived intuitionistically is $p \rightarrow \sim\sim Kp$, a claim which is classically equivalent to $p \rightarrow Kp$ but intuitionistically different. The proof at the beginning of this section does not rely on the reductio method, however, nor on any other form of proof that is not accepted by intuitionists. So acceptance of intuitionist logic may reduce the collateral damages caused by the paradox, but does not block the paradox itself.

These considerations aside, however, even if there is a way out of the paradox via intuitionism, it would not be an easy turn for the physicalist to manage. First a word about dealing with inconvenient paradoxes by switching logics: It is possible to weasel out of any apparent problem involving a simply by adopting a different, more convenient logic wherein the problem at hand simply vanishes. The Liar's paradox can be dealt with in this fashion. One needs only to adopt a paraconsistent logic (one in which contradictions are acceptable) in order to render this normally vicious paradox perfectly docile. If you wished, you could adopt a logic that disposes of Modus Ponens too. This would certainly alleviate any worries about the knowability paradox. The point, though, is that you can't just go about switching logics for no reason other than convenience in the current situation. To adopt a non-standard logic in order to deal with a logical problem in the absence of independent reason or motivation for adopting said logic is invariably to

beg the question. If there were such independent motivation available, however, then the resolution of the problem at hand would be only as strong as the motivation to adopt the new logic.

Here the physicalist appears to miss the intuitionist boat at the pier. Intuitionism has been advanced as a possible method of clarifying mathematical knowledge. The majority of mathematicians do not accept intuitionism as the right logic for their concerns, of course, but mathematics is nonetheless the intuitionist's familiar playground. The question remains, then whether there could be any independent philosophical motivation for the physicalist to accept intuitionism. The answer is almost ironic: One of the standard arguments in favor of accepting classical logic is that it remains the logic of scientific theory. The physicalist, as we have seen, is extraordinarily partial to scientific theory and should therefore be very impressed with this argument for classical logic and hence against intuitionism. Further, and perhaps more importantly, the natural home of intuitionism is in the philosophy of mathematics and there has been very little work done to extend intuitionism into the empirical domain. Since it is the empirical domain which is central to physicalism, it seems that the appeal to intuitionistic logic is unmotivated without a defense of how it might help yield a better account of empirical knowledge. This cannot be satisfying for the physicalist, however, because switching logics for utility's sake without independent motivation amounts to sophisticated question-begging.

III On Behalf of the Physicalist

If the attempts at general solution will not lend themselves to the physicalist cause, then the only hope for the view lies in demonstrating that it does not, after all, entail the verificationist premise of the Knowability paradox. We examined earlier the physicalist claim that the ontology required by physical theory is exhaustive of the total ontological inventory, and the required corollary that such a physical theory be a complete description of the world. From this notion of completeness we inferred the verificationist principle, that all truths can be known. But maybe that inference was a bit too hasty. Maybe it is possible that physicalism should be true, i.e. that there should exist a physical theory that provides a complete account of the world, and yet there still exists some truth that can possibly be known. If we were to demonstrate that this scenario is possible, physicalism would be rescued and physicalists could gleefully join in the denunciation of knowability. "Damn the paradoxes, full speed ahead!"

Let us examine this possibility more closely, then, to see if we might be of some assistance. In order to deny knowability, the physicalist must affirm the existence of at least one truth that is not possibly known. Physicalism requires however, the completeness of physical theory. The physicalist completeness claim can be understood as follows: Physical theory will be complete if and only if it accounts for all truths.

The physicalist, therefore, needs an unknowable truth that is accounted for by physical theory. If there exists a theory which accounts for all truths, yet there remains a truth which we cannot know, the only possibility is that we (by virtue of our limited faculties, perhaps) are incapable of using this theory to discover the truth in question. It is

possible to conceive of such a truth. There could be a pair of mutually exclusive propositions{p&~p}, one of which is true, that cannot be known and yet is accounted for by physical theory. What can be said of these propositions? Perhaps they are simply undetectable. Perhaps humans are prevented by the laws of physics from ever discovering whether p or ~p is actually the case. Nonetheless, physical theory could account for either one, much as quantum theory accounts for undetermined events even though it cannot predict them.

In fact, if the truths in question are undetermined, then the theory might still account for the truth in question by assigning probabilities to the pair in question, say .7 and .3 respectively. The theory cannot tell you which is true, of course, for the truth of the matter is undetectable. The truth is, nonetheless, accounted for. So the physicalist can escape knowability if there is an undetermined truth which is accounted for but yet undetectable.

To posit the existence of such a truth is a denial of determinism. If determinism is true, then knowledge of all that happened in the past is sufficient for predicting everything that happens today (using the physicalist's theory). So if there are any truths that cannot be predicted, e.g., our imaginary {p&~p}, then there must not be complete knowledge of the past, assuming determinism. If this is the case, then there are explanatory wholes going all the way back in time, undermining the completeness of scientific inquiry, for there will be "why" questions about which the best science is simply silent. So to explain these unanswered questions, and hence to save the physicalist from knowability, it is necessary to posit indeterminism. If this is the only way for the physicalist to deny the

knowability claim, then (barring a general solution to the paradox, of course) physicalists are philosophically committed to indeterminism!

It would not be shocking if we were to meet a physicalist on the street who happened to mention that she was an indeterminist, for the views are not thought to be incompatible. Indeed, if physicalists were to conduct their scientific inquiry, and at the end of that inquiry discover that some truth is indeterminate after all, this would not be disturbing for their view in the least, as an empirical discovery, but only as an empirical discovery. Given the physicalist commitment to the completeness and pre-eminence of physical theory, however, a philosophical, a priori commitment to any thesis whatsoever about the nature of reality is unacceptable. It matters not whether the thesis in question be friendly to physicalism or otherwise. So it turns out that our attempt to rescue the physicalist might be successful, but only at considerable cost to the rescuee.

Once again, however, it appears that celebration for even this modest success may be premature. Just as the paradox was reformulated to immunize it from attempts at solution by denial of knowledge distribution, the paradox will prove itself a slick one in this assault as well. The quantifiers in the original proof ranged over propositions. It could easily be stipulated that the proof use only restricted quantifiers, restricted so that they range only over determinately true propositions.[11] The only difference in the proof would be in the interpretation of the domain of the quantifiers. The premises for the new proof are (1) $\forall x(x \rightarrow \Diamond Kx)$, which would read "For all determinate propositions, if x is true, then it can be known that x" and (2) $\exists x(x\&\sim Kx)$, which would read "There exists a

[11] Thanks for the suggestion of utilizing restricted quantifiers to deal with this issue go to Michael Hand.

determinate proposition which is true and which is not known by anyone at any time." There is just as much reason for the physicalist to be committed to the new premises as the first, for the only propositions that are excluded are those which are indeterminately true, the most problematic of all in the case of the knowability claim (1), and there would still be little reason to expect the sort of omniscience required to deny (2). The remainder of the proof works out like the original, and the new and improved version will be a thorn in the flesh even of the most humble of physicalists who is willing to accept the a priori commitment to indeterminism discussed above. 6

And so we must leave physicalism on the ropes after all. Before I conclude, however, let us reflect on the moral of this story. First, it is not very prudent to fashion your ontology after a theory or mode of inquiry, empirical or otherwise. Second, if anyone must resort to such an ontological stance (such as physicalism), then he or she needs to pay close attention to the discussion surrounding an obscure little paradox about truths and possible knowledge. If the knowability paradox cannot be solved, then any such ontological view is false. Finally, I must acknowledge that even if this argument succeeds in demonstrating that the physicalists' method of stating their claim is misguided, there is no reason to believe that physicalism will simply disappear. What van Frassen calls 'the spirit of materialism has been around since the beginning of philosophy and will not soon go away. There is simply to much to deny and too much at stake.

Now, after all is said, I suggest we adjourn to enjoy the pleasures of some liquid refreshment, a phenomenon for which there doubtless *is* a satisfying physicalist explanation.

Bibliography

Crane, Tim and Mellor, Hugh, "There is No Question of Physicalism," *Mind* 99 (1990) p. 185

Edgington, Dorothy, "The Paradox of Knowability" *Mind* 94 (1985), pp. 557-68

Fitch, Frederic, "A logical Analysis of Some Value Concepts" *The Journal of Philosophy* 71 (1963), pp. 135-42

Haack, Susan, *Philosophy of Logics*, Cambridge University Press, 1978

Healey, Richard, "Physicalist Imperialism," *Proceedings of the Aristotelian Society*, (1978-1979)

Hellman, Geoffrey and Thompson, Frank Wilson, "Physicalist Materialism," *Nous* 11 (1977)

Quine, W.V., "Facts of the Matter," *Southwest Journal of Philosophy* 9, (1978)

van Fraassen, Bas, "Science, Materialism, and False Consciousness," (forthcoming)

van Inwagen, Peter, *Material Beings*, Cornell University Press, 1990

Williamson, Timothy, "On the Paradox of Knowability," *Mind* 96, (1987p.261.

Williamson, Timothy, "On Intuitionistic Modal Epistemic Logic" *Journal of Philosophical Logic* 21, (1992)

Williamson, Timothy, "Verificationism and Non-Distributive Knowledge," *Australasian Journal of Philosophy* 71 (1993), pp. 78-86

Zalta, Ed and Linsky, Bernard "Naturalized Platonism vs. Platonized Naturalism," *The Journal of Philosophy*

My thanks to Jonathan Kvanvig for helpful criticism, comments and instruction.