

Lewis on Materialism and Experience*

Daniel Stoljar

Australian National University

Introduction

The traditional question in philosophy of mind of the compossibility or co-tenability of materialism, on the one hand, and the existence of experience on the other is a theme throughout David Lewis's work. His first published paper, (1966) 'An Argument for the Identity Theory,' was a defence of materialism. And one of his most famous posthumous papers (2009) 'Ramseyan Humility' contains a discussion of the identification thesis, the rejection of which is crucial to his account of experience. In between, papers such as (1972) 'Theoretical and Psychophysical Identifications', (1978) 'Mad Pain and Martian Pain', (1983) 'New Work for a Theory of Universals', (1988) 'What Experiences Teaches', '(1994) 'Reduction of Mind', (1995) 'Should a Materialist Believe in Qualia?' and others contain a systematic and extremely influential approach to this topic.

This approach might usefully be viewed as the conjunction of four elements. *Element #1* is materialism¹ for which Lewis gave a distinctive and well-known characterization. *Element #2* is an account of what experience is: experience is the occupant of a certain functional role implicitly set out by folk psychology. *Element #3* is an account of the source of the tension between experience and materialism: the truth of materialism is apparently inconsistent with the existence of experience if we adopt a conception according to which experience is construed as satisfying all of the requirements of folk psychology. *Element #4* is a strategy for resolving the tension: Lewis argues that while the existence of experience may preclude materialism on the conception just mentioned, it is possible to articulate a different conception, according to which experience is construed as satisfying most (but not quite all) of the requirements of folk psychology. In this sense "it is not altogether wrong to call him an eliminativist, but see how much he eliminates and how much he retains" (1995, p. 329)

In the first part of this paper I will review these four elements of Lewis's account of materialism and experience. In the second I will consider two recent objections to that account. The first argues that knowledge-how is a certain kind of knowledge-that and in consequence Lewis's well known 'ability hypothesis' fails. The second argues that if Lewis's contextualist approach to epistemology is correct, his rejection of the identification thesis is impossible. I will suggest that Lewis has the resources to answer both objections, but I will end by stating where in my view the real problems for Lewis lie.

Element #1: Materialism

Lewis did not just give *a* distinctive and well-known characterization of materialism, he gave *two*: one in terms of fundamental properties, and one in terms of supervenience. The fundamental properties definition starts, he says, from something he believes a priori (1994, 291), namely, that any possible world at all instantiates a relatively small class of fundamental properties, where 'fundamental' or 'perfectly natural' properties are (among other things) "not at all disjunctive, or determinable, or negative. They render their instances perfectly similar in some respect. They are intrinsic; and all other intrinsic properties supervene on them" (2009, 204). On the assumption that what Lewis believes here is not merely a priori but necessary, one might think that materialism is the thesis that every fundamental property in any possible world at all is physical. But this construes materialism as necessary if true, when it is usually taken to be contingent.

* Acknowledgements: I am grateful to Paolo Santorio and to Wolfgang Schwarz for very helpful comments on a previous draft.

¹ Some might prefer 'physicalism' to 'materialism', but Lewis did not and I will follow him here. For a discussion of the issues lying behind the phraseology, see Stoljar 2010, ch.1.

A better approach is to provide what is sometimes called a ‘world-relative’ definition of materialism, i.e., one that tells us what it is for materialism to be true at some possible world arbitrarily chosen:

M1. Materialism is true at a possible world w iff every fundamental property instantiated at w is physical

Given this definition, the materialist *about a possible world W* believes that every fundamental property at W is physical. But of course we don’t normally speak of philosophers as materialist (or not) about this or that possible world; we speak of them as materialists (or not) without qualification, i.e. as just materialists. What is it then according to M1 to be a materialist without qualification? To be a materialist without qualification is from this point of view to be materialism about one possible world *in particular*, viz., the actual world. This has the benefit of portraying materialism as if true a contingent truth, because it is a contingent truth (if it is a truth) that the fundamental properties instantiated at the actual world are physical properties. The dualist or the vitalist, for example, will say that at least some of these properties are psychological or biological. Neither dualism nor vitalism is impossible, it is simply that neither is true or at any rate not if materialism is true.

Turning to the supervenience definition, this focuses on the idea of two possible worlds being, as Lewis says, duplicates of one another (1983, 27). Lewis starts his discussion here by portraying materialism as the thesis that, for any two possible worlds, if they are duplicates with respect to the physical, they are duplicates simpliciter. But, as he notes, this again wrongly treats materialism as necessary if true. As before, a world-relative account is better:

M2. Materialism is true at a possible world w iff for any possible world w^* if w and w^* are physical duplicates, they are duplicates simpliciter.

Like M1, M2 permits materialism to be a contingent truth, if it is true at all. If dualism is true at the actual world, there will be at least one world that is a physical duplicate of the actual world but is not a duplicate simpliciter, i.e. because it is different psychologically.

As Lewis notes, M2 faces an important counterexample. Consider a putatively possible world ϵ , which is physically exactly like our world but which contains some additional epiphenomenal ectoplasm: some stuff that is non-physical and yet does not in anyway disrupt the physical goings on in the world. Such a world seems possible; in particular, the materialist should not deny its possibility in view of the contingency of materialism. But M2 has the consequence that if materialism is true, ϵ is impossible. For if materialism is true, and M2 is the definition of materialism, any world physically exactly like the actual world is exactly like it in all respects. But ϵ is a counterexample to this: it is a world that is physically like the actual world but which is not exactly like it. Hence, M2 is mistaken.

To deal with this problem Lewis appeals to the notion of an alien property, where a “property is alien to a world iff (1) it is not instantiated by an inhabitant of that world; and (2) it is not analysable as a conjunction of, or as a structural property constructed out of, natural properties all of which are instantiated by inhabitants of that world” (1983, 37).² In the light of this, it is possible to modify M2 as follows:

M3. Materialism is true at w iff for any possible world w^* , if (a) w^* and w are physical duplicates, and (b) w^* contains no natural properties alien to w ; then w^* and w are duplicates simpliciter.

² As Paolo Santorio pointed out to me, there is at least the following problem with this definition of an alien property: it apparently counts as alien any property which meets (a) and (b) and would be instantiated in worlds that have the same laws as ours. I will not try to deal with this problem here.

Unlike M2, M3 does not have the truth of materialism entailing that ϵ is impossible, since ϵ is a world that includes properties alien to the actual world, at least if we suppose that epiphenomenal ectoplasm is or instantiates alien properties.

What is the difference between the fundamental property and supervenience definitions, i.e. between M3 and M1? One might think they are different in that one makes use of the notion of a fundamental property while the other does not. But this is not so. First, fundamental or at least natural properties are implicit in Lewis's definition of duplication and so in his definition of supervenience (Lewis 1983, 27). Second, as we have seen, Lewis appeals to alien properties in M3, and these are defined in terms of natural properties.

One might think that the two definitions are different in that one makes use of modal notions and in particular supervenience while the other does not. But this is not so either. Fundamental properties, as Lewis conceives them, are properties "which figure in a minimal basis on which all else supervenes" (2009, 205), so there is clearly a modal element in M1 too. Indeed, both M1 and M3 include at some level both supervenience and fundamentality.

One might also think that the two definitions are equivalent or at least that Lewis thinks that they are. But Lewis does not say that they are equivalent. In fact, in 'New Work for a Theory of Universals', the 1983 paper in which he explicitly advances the supervenience definition, he considers the fundamental properties definition³ and rejects it in the following passage:

Couldn't there be a natural property X (in the nature of the case, it is hard to name an example!) which is shared by the physical brains in worlds like ours and the immaterial spirits that inhabit other worlds? Or by this worldly quarks and certain otherworldly particles that cannot exist under our physics? Physics could quite properly make no mention of a natural property of this sort. It is enough to recognize the special case applicable to our world, X-cum-physicality, brainhood or quarkhood as it might be. Then, if by physical properties we mean those that are mentioned in the language of physics, a Materialist ought not to hold that all natural properties are instantiated in our world are physical properties. (1983, 34)

The suggestion here seems to be that the fundamental properties definition—M1, in our terms—is mistaken because it erroneously requires materialism to be false in the situation in which X is instantiated. After making this suggestion, Lewis goes on immediately to consider the supervenience definition (i.e. M3 in our terms), which suggests that this is his reason for abandoning M1 in favour of M3.

But it is very difficult to see this passage as containing a good reason for abandoning M1 in favour of M3. First, and this is merely *ad hominem*, in other papers (e.g. 1994) Lewis operates with the fundamental properties formulation and there is no suggestion that it is open to an objection that the supervenience definition is not. Second, if it worked against M1, this objection would also work against M3. For suppose Lewis's property X is instantiated. Presumably X is either necessitated by the physical or it is not. Suppose it is not necessitated; then it follows that M3 is false too. For in that case it follows that there is a world which (a) is a physical duplicate of our world; (b) instantiates no further alien properties; but (c) is not a duplicate simpliciter, i.e., because it does not instantiate X. Now suppose it is necessitated by the physical; then X presents no problem for M3, but likewise it presents no problem for M1—for a proponent of M1 might reasonably reply that X is a property like many others which is necessitated by the fundamental physical properties but is not itself a fundamental physical property. Finally, waiving these issues of necessitation, it is not clear that the

³ Lewis in fact discusses the view that all natural properties are physical properties, rather than all fundamental properties, but I will take him as talking about the fundamental properties definition too—the reference to 'quarks' in the passage at least is suggestive of it.

objection *does* work. For suppose there were a fundamental property unrecognized by physics but which imposed a genuine objective similarity on physical brains and immaterial spirits. It is quite natural to think that in this case that materialism *is* false, just as both M1 and M3 predict; hence there is no problem for either here.

Part of what makes this objection difficult is the reference in it to Lewis's account of what it is for a property to be a physical property. No definition of materialism, and in particular none of M1-M3, is complete without an account of what a physical property is. As Lewis says in the passage just quoted, "by physical properties we mean those mentioned in the language of physics", and in the same paper he says which physics he has in mind: something very like the one used by contemporary, actual, physicists, though "presumably somewhat improved" (1983, 33-4). But such an account faces serious problems. For one thing, could there not be properties that are both physical *and* alien, properties instantiated at worlds whose physics is not definable in terms of ours, and yet is in other respects recognizably like ours?⁴ Moreover, consider what Lewis in a later paper (2009, 205) calls *idlers*, i.e., fundamental properties which are instantiated at our world but which play no role in physical, or indeed any, theory. If there are such properties, then no matter what the intrinsic nature of such properties, it immediately follows from M1 and M3, together with Lewis's account of a physical property, that materialism is false, for then there will be fundamental properties that are not physical by his definition. Lewis argues that there is no reason to deny that idlers are instantiated (2009, 214). But then by his own account of what physical properties he has no reason to believe physicalism!

In view of the fact that is unclear both what the relation is between M1 and M3, and what account of a physical property they presuppose, there are clearly some open questions for Lewis's approach of materialism.

Element #2: Experience

In ordinary English, an experience is an event of some sort that in a certain hard to specify way engages our (or some salient subject's) psychology in certain distinctive ways. Take the Royal Wedding between Prince Dullsville and Princess Whatever-her-name-is. That was an experience, or at least was so for the 2 billion people that watched it on television. The formation of the solar system, by contrast, was an event but not an experience because it did not have the right sort of engagement with our (or anyone's) psychology. Obviously spelling out the right sort of engagement is a difficult matter. But however this is done, an experience need not in this ordinary sense be a psychological event like a judgement or the onset or persistence of pain. The Royal wedding is not a psychological event though it is an experience. One can say, for example, 'the experience left the Princess a changed person', meaning that the wedding left her changed.

In the dialect of English spoken by philosophers of mind, 'experience' is usually restricted to psychological events, and in particular to psychological events associated with sensory perception and bodily sensation, such as tasting the champagne, catching a glimpse of the Abbey, feeling the ring in one's pocket, and so on. In some parts of his writing, Lewis uses 'experience' in this way: "I will say *experiences* when I mean particular events of experiencing" (1995, 326), and it is clear in the context that he has in mind psychological events. Lewis also talks of experiential states, which are types of events of experiencing, or perhaps properties of these events. The two ideas are not unrelated since an experiencing event is naturally thought of as consisting (at least in part) in the onset or having of an experiential state, and indeed the word 'experience' might legitimately be used for both. But they are also different: an event is something that happens at a particular time and place and which cannot be repeated; an experiential state, by contrast, is a type of thing or a property,

⁴ Examples like this may be constructed by considering 'twin-physics' cases. See Stoljar 2010, ch.4 and the references therein.

something that two people can have or be in, and one person at different times can have or be in.

If experience (in the restricted sense that operates in philosophy of mind) is an event that consists in the onset or the persistence of an experiential state, what is it that makes something an experiential state? For Lewis, experiential states are implicitly defined by folk psychology. Folk psychology is a tacitly known theory or body of information about how people in general act and think; so folk psychology might say that experiences have this sort of feature, and are typically brought about this way, and so on. Lewis thinks that if we can isolate the key principles or theses of folk psychology that concern experience, then in principle one could use them to construct an explicit definition of what experience is, using the famous Ramsey-Carnap technique (see Lewis 1972, 1994, 1997). We will not go into that technique here; for us the key point is that experience is that thing which satisfies all or most of the relevant principles of folk psychology.

What then are the relevant principles of folk psychology? Lewis did not give a detailed analysis of what folk psychology says about experience; at one point he says he offers ‘recipes for analysis’ rather than fully-fledged accounts (1994, 298). Nevertheless, it is possible to draw out from his writings some central principles, some of which are more controversial than others, some are more general than others, some of which overlap, and some of which might turn out to be explained in terms of others. These are:

Causal Thesis: when you have an experience⁵ you are in some inner state which typically causes you to behave in various ways, to form other mental states, and which is typically caused by certain things in the world. (Lewis 1966, 1972, 1978)

Informational Thesis: when you have an experience, you are in a state with a certain informational content. (Lewis 1983a)

Belief Thesis: when you have an experience with the informational content p you are disposed to form a belief that p (or a similar content). (Lewis 1983a, 1994)

Something it is Like Thesis: when you have an experience, there is something it is like for you to have that experience. (Lewis 1988)

Knowing what it is like Thesis: when you have an experience, you know what it is like to have that experience. (Lewis 1988, 1995)

Ability Thesis: when you have an experience you have an ability to imagine the experience, to remember the experience, and to recognize the experience. (Lewis 1988, 1994, 1995)

Identification Thesis: when you have an experience you know the essence of the experience, i.e. exactly what it is in an ‘uncommonly demanding sense’. (Lewis 1995, 1997)

Self-Intimation Thesis: when you have an experience you know that you are having the experience. (Lewis 1972, 1996)⁶

Putting these theses together we arrive at the view that an experiential state is a state that satisfies all or most of them. As we will see, ‘all or most’ part is important. Consider the view that, by definition, an experiential state satisfies all (and not merely most) of these. And suppose now we discover that nothing that exists in a physical world—i.e. a world at which materialism is true—could satisfy all of them. It would follow that nothing that exists in a physical world could be an experience; that is, materialism and the existence of experience are

⁵ Some of these principles are more naturally expressed in terms of experiencings, and some are more naturally expressed in terms of experiential states. I will leave that unresolved in the text.

⁶ In (1972) Lewis says that self-intimation is an open question, whereas in (1996) is committed to it, as we will see below.

not co-tenable. On the other hand, if we require only that an experiential state satisfies most of these this result may well be avoided.

Element #3: Materialism and Experience in Tension

We have considered what materialism is for Lewis, and what experience is; I now turn to the two main arguments he discusses against the co-tenability of materialism and the existence of experience.

The first—the Identification Argument, I will call it—starts from an inconsistency between materialism, the identification thesis, and some agreed-on facts. Consider the following claims:

- (1) If Fred is in an experiential state E, then Fred knows or is in a position to know the essence of E.
- (2) Fred is in experiential state E.
- (3) The following is at least part of the essence of E: the experiential state E is identical to the physical state P.
- (4) Fred does not know, and is not in a position to know, that experiential state E is identical to P.

These claims are inconsistent. If (1) and (2) are true, Fred knows the essence of E. But if (3) and (4) are true, Fred does not know the essence of E, i.e. because to know the essence of E would be to know that it is identical to P, and Fred does not know that. On the other hand, (1) is an instance of the identification thesis; (3) follows from materialism as Lewis understands it; and (2) and (4) are the agreed-on facts I mentioned. We can all agree, after all, that there are agents (or at least could perfectly well be) that both have experiences and are ignorant of their physical basis—Fred is just a stand in for such an agent. If (1-4) are inconsistent, one of them is false. But (2) and (4) are obvious, and (1) follows from what folk psychology says an experience is. Hence (3), and materialism generally, is false.

One might wonder whether (3) *does* follow from materialism as Lewis understands it. For example, one reason some philosophers find supervenience definitions of materialism (such as M3) attractive is that they apparently permit one to be materialist and deny an identity between physical states and psychological states. Be that as it may, for Lewis supervenience and identity naturally go together. Indeed, for him the idea that experience is defined in terms of its folk psychological role provides the starting point of a powerful argument for the identity theory; this is Lewis's famous argument for the identity theory set out in Lewis 1966 (see also Lewis 1972, 1994; and Armstrong 1968). The first premise of the argument is that, where 'E' some experiential state, E is the state that satisfies all or most the features above. The second premise is that, where 'P' is some physical state, P is the state that satisfies all or most of the features above. The conclusion is that E is, i.e., is identical to, P. The first premise here is supported by the approach to experience we reviewed earlier. The second premise is supported on empirical grounds—for example, on the ground that there is some state of the brain that is the state that satisfies most of the features above. And the conclusion of the argument follows, Lewis says (1972), "by the transitivity of identity".

The second argument against the co-tenability of materialism and the existence of experience is the (better-known) Knowledge Argument. This argument may be set out in various ways, but a simple version has it as proceeding from two main premises. The first premise concerns what it is possible for a person to know; in particular, it is possible for a person to know all the physical facts as well as every fact that follows a priori from the physical facts, and yet not know what it is like to have an experience of certain type. Jackson's (1982) Mary is the best-known illustration of this possibility, and the one that

Lewis mostly focuses on.⁷ The second premise of the knowledge argument is that if this is possible then materialism is false. The conclusion is that materialism is false, or anyway it is false if there are facts about what it is like to have certain experiences and if people know these facts.

One might wonder again whether this argument threatens materialism in the form that Lewis advances it. Another reason some philosophers find supervenience definitions of materialism (such as M3) attractive is that they apparently permit one to be materialist and deny the a priori deducibility of the mental by the physical. Such philosophers are a posteriori materialists rather than a priori materialists, as it is often put, and on the surface, the knowledge argument targets only the a priori materialist. Be that as it may, for Lewis supervenience and a priori deducibility naturally go together; he is an a priori materialist. The reason for this is his commitment to (and understanding of) the ‘two-dimensional’ analysis of the necessary a priori. We will not go into the details of that position here.⁸ Suffice it to say that, if materialism is true, and if M3 is the definition of materialism, there is a necessary conditional of the form ‘If S then S*’, where S gives precise information about the instantiation of every physical property in the world, and also says that no alien properties were instantiated, and S* gives precise information about the instantiation of every psychological property. The issue which divides the a priori from the a posteriori materialist is whether this conditional is a priori or not. Lewis thinks that the ‘two-dimensional analysis’ of the necessary a posteriori entails that the conditional ‘If S then S*’ is a priori if materialism is true (1994, 297; see also Chalmers 1996 and Jackson 1998). Conclusion: a posteriori materialism is not a possible position.

Element #4: Distinct Conceptions of Experience

If he does not respond to the Identification Argument and the Knowledge Argument by rejecting identity or a priori deducibility, how does Lewis respond? In both cases, his strategy is, first, to distinguish two different conceptions of experience (or a related notion), and second to suggest that while the arguments rely on one conception, the materialist may employ another. The materialist need not insist that his own is the best one, or the perfect one, or the one that best answers to ordinary or philosophical thought; but he does need to argue that it is a legitimate one.

This strategy is particularly clear in the case of the Identification Argument. As we saw above, if you operate with a conception of experience according to which something is an experience only if it satisfies all the requirements of folk psychology, it will follow that the Identification Argument is sound and materialism is false. But Lewis argues that a conception of experience that dispenses with the identification thesis is a legitimate one: something which satisfies the other features of experience listed above would still deserve the name ‘experience’, even if one can imagine a more perfect deserver of the name. It is in this context that Lewis makes the remark about eliminativism that we quoted right at the beginning. Does Lewis deny the existence of experience? Yes and no; that is, he denies it if you mean something that satisfies all the requirements of folk psychology and the identification thesis in particular; he does not if you mean something that satisfies most of the requirements.

In the case of the Knowledge Argument, the ‘two legitimate conceptions’ strategy assumes a slightly different form. Here Lewis focuses on two conceptions, not of experience,

⁷ Here is Lewis’s own description of the case: “Mary, a brilliant scientist, has lived from birth in a cell where everything is black and white. (Even she herself is painted all over.) She views the world on black-and-white television. By television she reads books, she joins in discussion, she watches the results of experiments done under her direction. In this way she becomes the world’s leading expert on color and color vision and the brain states produced by exposure to colors. But she doesn’t know what it is like to see color. And she never will, unless she escapes from her cell” (1988, 263).

⁸ For some discussion of this issue, see Schwarz this volume

but of knowing what an experience is like. On the first, to know what it is like is to know a phenomenal fact, where a phenomenal fact is defined as a kind of fact which does not supervene on the physical facts, and which (remote cases aside) requires experience in order to be known—Lewis calls this the “hypothesis of phenomenal information”. On the second, to know what it is like to see red is (merely) to know how to do something, or to have certain ability, viz., to recollect, recognize, and imagine certain experience—Lewis calls this “the ability hypothesis”. Lewis first argues that the Mary example could be described in either way: on the first, she learns a phenomenal fact; on the second she learns some know-how. He then argues that if the example is described in the second way there is nothing in it to threaten materialism, since materialism does not rule out the possibility that someone might know all the physical facts, and yet gain some abilities.

Lewis’s responses to these two arguments are different in an important way. In the case of the Identification Argument, both conceptions he operates with are present in folk psychology, or at any rate so Lewis claims. In the case of the Knowledge Argument, however, no such claim is made. In particular, the hypothesis of phenomenal information and the conception of knowing what it is like which goes along with it is an overt piece of philosophy—a “gratuitous metaphysical gloss” (1988, 290)—and indeed one which is question begging, since one can only know what an experience is like in this sense if materialism is false. This seems to me a weakness in Lewis’s discussion of the knowledge argument in his 1988 paper. The psychological plausibility of the Mary example and the argument founded on it is very striking; it is implausible that only those who are in the grip of some tendentious metaphysics find them compelling. It would be preferable here therefore if the two conceptions of ‘knowing what it is like’ Lewis’s response requires are plausibly part of ordinary thought.

In “Can a Materialist Believe in Qualia”, the 1995 paper that came out seven years after Lewis’s main discussion of the knowledge argument, we find the materials to deal with this problem. In that paper, Lewis argues that ‘knowing what it is like’ is a matter of knowing the essence of the experience; in other words, knowing what an experience is like is explicated in terms of the identification thesis. Moreover he argues explicitly that this idea is built into folk psychology, and is connected to the knowledge how that is involved in the ability hypothesis because according to folk psychology one has these abilities because knows the essence. He also says that a commitment to the identification thesis is central to dualist thinking, and mentions Kripke (1980) in particular in this connection. While he does not explicitly draw the connection the knowledge argument, it is natural to read his remarks as applying there too. If so, we may adjust his response to the knowledge argument in a way that avoids the problem about folk psychology just mentioned. According to this revised response, there are two conceptions of knowing what it is like: the first is merely to have the ability to imagine (etc.) the relevant experience, the second is to know the essence of that experience. The Mary example can be described in either way, but the materialist can explicitly adopt the first.

Know-how and The Ability Hypothesis

Turning now to objections to Lewis’s account, the first I will consider focuses on Lewis’s response to the knowledge argument. As we have seen, Lewis adopts the ability hypothesis here, and argues that Mary gains knowledge-how or abilities rather than knowledge-that when she emerges from her cell. However, a number of philosophers (e.g. Loar 1990, Lycan 1996, Stanley and Williamson 2001, Stanley 2010; see also Alter 2001 and Cath 2009) have objected that knowing how to do something is simply a kind of knowledge that, and that in consequence the ability hypothesis is mistaken. The premise of this objection—that know-how is a species of know-that—is certainly plausible. And no one can deny that Lewis formulates the ability hypothesis in terms of know-how which is not know-that. The question is whether this part of his account might be dropped while key elements of the ability

hypothesis are retained—whether there is a ‘fallback,’ as Stanley and Williamson (2001) put it.

The problem is that there is an apparently decisive argument that no fallback is possible. For suppose that knowing how to imagine (etc.) an experience is a kind of knowledge that—for example, knowledge of some proposition *p*. Either Mary knows *p* in her room or she does not. Suppose she does. In that case, a proponent of the ability hypothesis may still say that Mary gains some ability on coming out, for it is possible to separate the ability from the knowledge-how.⁹ But what the proponent can’t say is that in gaining this ability Mary gains some knowledge-how—for by assumption she already has that knowledge. And this means that the ability hypothesis loses a lot of its appeal. Central to the Mary case is what might be called ‘the knowledge intuition’, the idea that Mary learns something on coming out, (i.e. gains some knowledge on coming out). If the ability hypothesis can explain this by saying Mary gains some knowledge-how, well and good; but if the ability hypothesis cannot explain it, it provides no reasonable description of the case. Suppose then that she does not know *p* in her room. In that case, a proponent of the ability hypothesis can say that Mary gains some propositional knowledge on coming out, and may appeal to this to explain the knowledge intuition. But the problem now is that the propositional knowledge in question cannot be (if materialism is true) the sort of propositional knowledge that Lewis is mainly interested in, i.e. the kind that distinguishes one possible world from another. It is true that one can define various notions of propositional knowledge that do not distinguish one possible world from another. But Lewis is scathing on the idea that doing so is connected to the Mary example. Indeed of the six ways he sets out of missing the point of the knowledge argument, no fewer than four are related to this idea!

Yuri Cath (2009) has suggested nevertheless that the second option here is the best one. His idea is that pre-release Mary might have the propositional knowledge but not under the right mode of presentation. Cath makes a good case that this is plausible for some proponents of the ability hypothesis, but it is hard to see it as plausible for Lewis, as indeed Cath points out at the end of his paper. If we are out to find a plausible fallback for Lewis, it is better to focus on the first horn of the dilemma just outlined, and to try defend the view that the ability Mary gains is genuinely a kind of knowledge even if it is not know-how.

I think such a defence emerges when we observe that a sentence such as ‘Mary knows what it is like to see red’ is ambiguous in a way that many ‘know what’ sentences are. On the one hand, it might be used to say that Mary knows some fact that (in the context) answers the embedded question ‘what is it like to see red?’ Since that question seems intuitively to ask ‘what type of experience does one have when one sees a red thing?’ the sentence on this reading means that Mary knows some fact that (in the context) answers this question. Call this the ‘interrogative’ reading of the sentence. On the other hand the sentence can also be used to say that Mary knows the thing or property denoted by the referring expression ‘what it is like to see red’. Since the referring expression intuitively denotes a type of experience, on this reading the sentence says that Mary knows a type of experience, i.e., the one you have when you see red. Call this the ‘free relative’ reading of the sentence. It is the interrogative reading we use when we say of pre-release Mary ‘she wonders what it is like to see red’—what she wonders is not a type of experience, but what fact answers a certain question. It is the free relative reading we use when we say of post-release Mary, ‘she loves what it is like to see red’—what she loves is a type of experience, not an answer to a certain question.¹⁰

⁹ It is possible for someone to know how to do something without having the ability to do it; witness the unfortunate pianist who knows how to play but lacks the ability because she has just lost her arms in a car crash (cf. Ginet 1975 Chomsky 1988, Stanley and Williamson 2001, Cath 2009) Similarly, it is possible that Mary knows to imagine but does not have the ability, i.e., because she has not had the experience.

¹⁰ For the distinction (plus the labels), though not the application to ‘know what it is like,’ I am indebted to Schaffer 2010.

Which of these readings is Lewis interested in? While he does not explicitly draw this distinction, I think it is plausible to interpret him as being concerned mainly with the free relative reading. One data point is that Lewis explicitly says (1988, 286, fn.12) that the phrase ‘what experience E is like’ in the larger phrase ‘know what experience E is like’ denotes E itself; this strongly suggests he is setting aside the interrogative reading of the sentence. Another data point is Lewis’s second way to miss the point of the knowledge argument (1988, 266). The second way to miss the point, according to Lewis, is to suppose that what Mary doesn’t know is an answer to a certain question, viz., ‘what type of experiences she will have when sees red for the first time’. Lewis argues that she does know an answer to this question, and so ‘knowing what it is like’ should not be understood in this interrogative way.

Suppose then that Lewis does indeed employ ‘know what it is like to see red’ on its free relative reading; in that case, we have a response to the first horn of the dilemma outlined earlier. Mary may well know how to imagine red but she does not have the ability to do so until she has an experience. But why is the ability that she gains properly called a state of knowledge? The answer is that it is the state of knowing a type of experience, i.e. ‘knowing what it is like’ on the free relative sense.

Contextualism and the Identification Thesis

The second objection I want to consider focuses on the identification hypothesis. We have seen that Lewis rejects this, but a number of philosophers (Schwarz 2007, Stalnaker 2008) have pointed out that doing so seems inconsistent with Lewis’s own contextualist account of propositional knowledge.

According to this account, “subject S knows proposition P iff P holds in every possibility left uneliminated by S’s evidence; equivalently, iff S’s evidence eliminates every possibility in which non-P” (1996, 422). This account raises a number of questions, but the connection to the identification thesis has to do with Lewis’s account of what it is for evidence to eliminate every possibility.¹¹ Here is what he says:

A possibility W is uneliminated iff the subject’s perceptual experience and memory in W exactly match is perceptual experience and memory in actuality. (1996, 424)

Pairing this account of the elimination of possibilities with the main part of Lewis’s contextualism already has an important consequence for Lewis’s account of experience. In particular, it entails the self-intimation thesis about experience mentioned earlier: that if I am in a certain experiential state, I know that I am in that state. Some might think that this is bad enough¹², but Stalnaker and Schwarz argue that the problem goes deeper. As Stalnaker puts it:

...on Lewis’s account, we will at least know, in any context, that the possibilities excluded by our experience—possible situations in which our experience does not match our actual experience—are possibilities that are incompatible with our knowledge. The problem is that all of these possibilities will be possibilities in which our experience has whatever essential properties our actual experience has. That is, Lewis’s account of knowledge implies that even in our most sceptical context, we will know the essential nature of our experiences....Lewis’s account of knowledge entails the identification Thesis that he rejects...” (2008, 100)

¹¹ As Wolfgang Schwarz pointed out to me, it is Lewis’s account of what it is to eliminate possibilities that generates this problem not the contextualism strictly speaking. I will ignore this however in the text.

¹² As Hawthorne (2004) points out, “if I have a visual array with 137 red dots, it does not seem that I automatically know that (60, n.26), though Lewis’s account apparently implies that do.

To bring this out, suppose I am in an experiential state E; and suppose in accordance with materialism of the kind Lewis defends, that this state is identical to some physical state P. As we have just seen, it will immediately follow that I will know that I am in E. Does it follow also that I know that I am in P? According to the objection, the answer is yes. For in all the worlds that remain uneliminated by my evidence I will be in P; hence I will know that I am in P. And this seems tantamount to the identification thesis.¹³

It is hard to overstate what a disaster it would be for Lewis's account of materialism and experience if this objection were correct. As we have seen, the rejection of the identification thesis is crucial to his response to the identification argument, and perhaps too the knowledge argument. But is the objection correct? I think this depends on what it is for a subject's perceptual experience in one world to 'exactly match' that subject's perceptual experience in another. The first thing to say is that what Lewis has in mind here are events of experiencing, and as we saw an event is an experience if consists in the onset or persistence of an experiential state. Presumably, therefore, experience *e* in possible world *W* will match experience *e** in possible world *W** if and only if they consist in the onset or persistence of the same experiential state; for short, if and only if they consist in the same experiential state. But what it is for two experiences (in distinct possible worlds) to consist of the same experiential state? I think there are two possibilities here. According to the first, two experiences consist of the same experiential state (and so match each other) if and only if they consist of experiential states with the same functional role. According to the second, two experiences consist of the same experiential state (and so match each other) if and only if they consist of experiential states with the same essence. Two experiences could match each other in the first sense and not in the second. For example, contrast a dualist world *D* and a materialist world *M*, and suppose that at both worlds I am in pain. Do the two experiences I have in these different worlds (i.e., the one I have in *D* and the one I have in *M*) match? According to the first account of matching, they do; but according to the second, they do not.

The distinction between these two accounts of matching provides an answer to the objection made by Stalnaker and Schwarz.¹⁴ It is true that if the second account of matching is in play, it will follow that I know that I am in P, and this is objectionable. But if only the first account is in play, nothing similar is true: in some of the possible worlds at which I have an experience which matches the actual experience I will not be in P. Is it reasonable to think that Lewis is operating only with the first account? It would seem so. First, as we saw earlier, an experiential state is something defined in terms of its folk psychological role, which suggests that it is not defined in terms of its essence. Second, the contextualist program in epistemology he outlines could proceed unhindered if only the first account is in play. If so, the objection fails and disaster is avoided.

Concluding Remarks

I have defended Lewis against these two objections, but I should say for the record that I have a number of sharp disagreements with Lewis's account of experience and materialism, which I have set out elsewhere. I will conclude with a brief statement of what these are.

Disagreement #1 concerns the ability hypothesis. Lewis's ability hypothesis focuses on the idea of a novel experience—that is, an experience one has not had before—and this focus is shared by many contemporary accounts of the knowledge argument, in particular acquaintance views, phenomenal concept views and so on. Now the Mary example *is* an

¹³ Seems tantamount, but perhaps is not quite identical; the identification thesis requires not simply that I know that I am P but that I know that E is P. However, the knowledge described in the text is bad enough.

¹⁴ In his (2007), Schwarz suggests something along these lines, but objects that, on that view, one will always know that various causal facts involved in having an experience. I am not sure if that is an objection or not, but even if it is it is clearly preferable to knowing the essence of the experience from Lewis's point of view.

example in which someone has a novel experience. But this feature of the example is incidental to the persuasiveness of the argument founded on it; other examples can be constructed that support that key possibility just as well but which do not involve novel experiences. The ability hypothesis has nothing to say to examples of this kind, and neither does any response to the knowledge argument founded on novel experiences (see Stoljar 2005, 2006, ch. 10). So I think a completely different approach is required.

Disagreement #2 concerns the identification thesis. I agree with Lewis that the identification thesis is false, and in that sense with his response to the identification argument. But I doubt that the identification thesis plays the role in ordinary and philosophical thought that Lewis thinks it does. For one thing, what is built into folk psychology is an empirical question, logically like the question of whether folk physics is a version of medieval impetus physics. But even so, the suggestion that the identification thesis is built into folk psychology is something about which we should be sceptical. It is plausible that folk psychology asserts some connection between experience and belief about or knowledge of experience. But there are many such theses, and the identification thesis is only one (and a fairly extreme one at that); there is so far as I know no good reason to believe that this thesis *in particular* is built into folk psychology (see Stoljar 2009). Nor is it credible to suppose that the identification thesis is a presupposition of philosophical thought that is sympathetic to dualism. Take Kripke's discussion of the modal argument in *Naming and Necessity*, something that Lewis mentions. Kripke is concerned there to emphasize that there is no distinction between feeling pain and being in pain, but I think one can emphasize that without being committed to the identification thesis (Stoljar 2006, ch. 11, Stoljar 2009).

Disagreement #3 concerns Lewis's optimism about current science. We saw earlier that, according to Lewis, a property is a physical property if and only if it is expressed by a physical theory that is a 'presumably somewhat improved' version of what we currently have. I think, as I said, that this is subject to persuasive counterexamples (see Stoljar 2010, ch. 4). But there is more important issue at stake here. For Lewis's discussion of physical properties and indeed of materialism itself embodies a quite radical optimism about current science, an optimism that is routinely underappreciated, as Jackson noted in his first discussion of the knowledge argument (see Jackson 1983).¹⁵ Lewis himself thought that physics was very nearly complete, i.e. that every type of physical truth was known or nearly so, and that every contingent truth follows a priori from physical truths. Now the question of whether optimism or pessimism is true is an example of a disagreement in philosophy that will not easily be resolvable, if at all. Presumably, the dedicated optimist will remain unmoved by reminders about the fallibility of humans, and about the history of wrong turns and mistakes in our attempts at understanding the world (see Stoljar 2010, ch.5). Speaking personally, however, I find this sort of optimism intellectually alien. I don't think scientists believe it, even if they occasionally say things that suggest that they do. And if they do believe it, I think they are wrong; I don't think it is worthy of belief. Moreover, even if this sort of optimism is by some miracle true, I don't think arguments that presuppose it are persuasive.

Disagreement #4 concerns what the knowledge argument and similar arguments do presuppose. As I have said, in my view the knowledge argument does not presuppose the identification thesis, nor is it essentially focused on novel experiences. However, I do think it presupposes something like the optimism just mentioned, in particular, I think it presupposes that we have complete knowledge of the physical or non-experiential world if not in detail then at least in outline (Stoljar 2006, 2009a). But as I have just indicated, that presupposition is in my view false (in fact I think the plausibility of these arguments is evidence that it is false). And if it is false, I think it is quite clear where the arguments go wrong. If we assume we have incomplete rather than complete knowledge of the relevant facts, what the knowledge argument and related arguments show is the unremarkable truth that experience

¹⁵ This optimism is distinct from Lewis's (2009) argument for Ramseyan Humility, as Lewis notes in fn.5 of that paper.

comes apart from some physical facts (i.e. the known ones), not the remarkable truth that it comes apart from all.

I know these points are controversial; my aim in this last section was to state my disagreements with Lewis, not defend them.

References

- ARMSTRONG, D.M. 1968. *A Materialist Theory of the Mind* (London: Routledge)
- ALTER, T. 2001. 'Know-how, ability and the ability hypothesis' *THEORIA* 67 (3):229-39
- CATH, Y. 2009. 'The Ability Hypothesis and the New Knowledge How' *NOUS* 43:1 (2009) 137-156.
- CHOMSKY, N. 1998. *Language and Problems of Knowledge* (Cambridge, MA: MIT Press)
- GINET, C. 1975. *Knowledge, Perception and Memory* (Dordrecht: Reidel)
- HAWTHORNE, J. 2004. *Knowledge and Lotteries* (Oxford: Oxford University Press)
- JACKSON, F. 1983. 'Epiphenomenal Qualia' *Philosophical Quarterly* 32:127-136
- KRIPKE, S. 1980. *Naming and Necessity* (Cambridge, MA: Harvard University Press)
- LEWIS, D. 1966. 'An Argument for the Identity Theory' *The Journal of Philosophy*, 63: 17-25, repr. in his *Philosophical Papers Vol.I* (Oxford: Oxford University Press, 1983), ppppp. All references are to the reprinted version.
- LEWIS, D. 1972. 'Psychophysical and Theoretical Identifications' *Australasian Journal of Philosophy*, 50: 249-258, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 248-261. All references are to the reprinted version.
- LEWIS, D. 1978. 'Mad Pain and Martian Pain' in Block, N. (1978) *Readings in the Philosophy of Psychology Vol I*, pppppp. Repr. in his *Philosophical Papers Vol I. apers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 325-31. All references are to the reprinted version.
- LEWIS, D. 1983. 'New Work for a Theory of Universals' *Australasian Journal of Philosophy*, 61: 343-377, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 8-55. All references are to the reprinted version.
- LEWIS, D. 1983a. 'Individuation by Acquaintance and by Stipulation' *Philosophical Review Vol. 92, No. 1*. pp. 3-32, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 373-402. All references are to the reprinted version.
- LEWIS, D. 1988. 'What Experience Teaches', *Proceedings of the Russellian Society*, 13: 29-57, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 325-31. All references are to the reprinted version. References are to the reprinted version.
- LEWIS, D. 1994. 'Reduction of Mind', in S. Guttenplan ed., *A Companion to the Philosophy of Mind* (Oxford: Blackwell, 1994), 412-31. repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 291-324. All references are to the reprinted version.
- LEWIS, D. 1995. 'Should a Materialist Believe in Qualia?' *Australasian Journal of Philosophy*, 73: 140-144, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 325-31. All references are to the reprinted version.

- LEWIS, D. 1996. 'Elusive Knowledge' *Australasian Journal of Philosophy*, 74: 549-567, repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 418-445. All references are to the reprinted version.
- LEWIS, D. 1997. 'Naming the Colours', *Australasian Journal of Philosophy*, 75: 325-342. repr. in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 332-358. All references are to the reprinted version.
- LEWIS, D. 2009. 'Ramseyan Humility' in Robert Nola and David Braddon Mitchell (eds) *Conceptual Analysis and Philosophical Naturalism* (Cambridge, MA: MIT Press) pp. 203-222.
- LOAR, B. 1997. 'Phenomenal States', in Block N, Flanagan O, and Guzeldere, G (eds) *The Nature of Consciousness* (Cambridge, MA: MIT Press), pp. 597-616
- LYCAN, W.G. 1996. *Consciousness and Experience* (Cambridge, MA: MIT Press)
- SCHAFFER, J, 2010. 'Knowing the Answer Redux' *Philosophy and Phenomenological Research*, vol. 78, no. 2, pp. 477-500.
- SCHWARTZ, W. 2007. 'Lewis in Knowing one's Evidence'. Blogpost at: <http://www.umsu.de/wo/2007/500>
- STANLEY, J AND WILLIAMSON T, 2001. 'Knowing How' *Journal of Philosophy* 98:411-444
- STANLEY, J. 2011. *Know How* (Oxford: Oxford University Press)
- STALNAKER, R. 2008. *Our Knowledge of the Internal World* (Oxford: Oxford University Press)
- STOLJAR, D. 2005. 'Physicalism and Phenomenal Concepts' *Mind and Language* 20 (5) 469-494
- STOLJAR, D. 2006. *Ignorance and Imagination* (New York: Oxford University Press)
- STOLJAR, D. 2009. 'The Argument from Revelation' in Robert Nola and David Braddon Mitchell (eds) *Conceptual Analysis and Philosophical Naturalism* (Cambridge, MA: MIT Press) pp. 113-138
- STOLJAR, D. 2010. *Physicalism* (London: Routledge)