

Of Zombies, Color Scientists, and Floating Iron Bars

Tamler Sommers
Department of Philosophy
Duke University
Durham, North Carolina 27708
U.S.A.

ts7@duke.edu

Copyright (c) Tamler Sommers 2002

PSYCHE, 8(22), November 2002
<http://psyche.cs.monash.edu.au/v8/psyche-8-22-sommers.html>

KEYWORDS: materialism, dualism, logical possibility, Chalmers, zombies, conceivability.

ABSTRACT: In this paper I challenge the core of David Chalmers' argument against materialism-the claim that "there is a logically possible world physically identical to ours, in which the positive facts about consciousness do not hold." First, I analyze the move from conceivability to logical possibility. Following George Seddon, I consider the case of a floating iron bar and argue that even this seemingly conceivable event has implicit logical contradictions in its description. I then show that the distinctions Chalmers employs between primary and secondary intensions, and a priori and a posteriori entailment, break down upon close examination-with iron bars and with consciousness it is impossible to know where primary intensions end and secondary intensions begin. I extend this analysis of logical possibility to the famous zombie thought experiment and conclude not that a zombie world is logically *impossible*, but rather that, at present, the question is open. Finally, I show how a similar line of argument may be used to undermine the "Mary the color scientist" thought experiment as well.

1. Introduction

In his book *The Conscious Mind* David Chalmers presents the following argument against materialism:

1. In our world, there are conscious experiences.

2. There is a logically possible world physically identical to ours, in which the positive facts about consciousness do not hold.

3. Therefore, facts about consciousness are further facts about our world, over and above the physical facts.

4. So materialism is false. (Chalmers 1996^{<1>}, p.123)

To support premise (2) Chalmers asks us to imagine a zombie world, a world physically identical to our own, but lacking consciousness. In this zombie world live our zombie twins, molecule for molecule replicas who behave exactly as we do, but who have no qualitative experiences whatsoever. Chalmers finds that this zombie world is coherently conceivable, and so he concludes that it is logically possible. The rest of his argument follows easily after this. To use Kripke's metaphor, if God still had more work to do after ensuring that the physical facts in the world held, we may safely conclude that materialism is false.

The key premise, then, is premise (2), and in addition to the zombie world, Chalmers offers two other thought experiments to support it. The first claims the logical possibility of an inverted spectrum, and the second features Mary, an omniscient color scientist imprisoned in a black and white room.

In this paper I will address only the Zombie and Mary thought experiments (but for a detailed treatment of the inverted spectrum see Dennett 1991). I will argue that it is an open question whether a zombie world is conceivable, and therefore that Chalmers' modal argument against materialism fails. I will also show that the alleged distinction between a priori and a posteriori entailment from physical to phenomenal facts breaks down upon close analysis. Finally, I will use roughly the same line of reasoning to cast doubt on the first premise of Frank Jackson's 'Mary the Color Scientist' argument against materialism. Along the way I hope to follow Gerald Massey (1991) in discrediting what he calls "off-the-cuff" conceivability claims.

2. What Does it Mean for Something to be Conceivable?

Massey takes Hume's theory of conceivability to be representative of the doctrines in contemporary philosophy. In the *Treatise*, Hume writes:

'Tis an established maxim in metaphysics, *that whatever the mind clearly conceives includes the idea of possible existence*, or in other words, *that nothing we imagine is absolutely impossible*. We can form the idea of a golden mountain, and from thence conclude that such a mountain may actually exist. We can form no idea of a mountain without a valley, and therefore regard it as impossible.

And then later:

'Tis in vain to search for a contradiction in any thing that is distinctly conceiv'd by the mind. Did it imply a contradiction, 'tis impossible it cou'd ever be conceived. (From Massey 1991, p.289)<2>

Chalmers seems to go along with Hume's take on conceivability and the subsequent move to possibility, with this modification: a statement is conceivable (or conceivably true), he says, if it is true in some conceivable world. So in establishing the conceivability of statement, one must be careful to avoid misdescribing the world in which the truth of the statement is being considered. For both worlds and statements conceivability involves a having clear impression *and* discerning no contradiction in the description (it cannot reduce to 'p and not p').

On this view, a married bachelor is clearly inconceivable. There is an obvious contradiction in the statement, and the relevant world in which he might exist is our own. A mile high unicycle is clearly conceivable. Though it may be physically impossible, we can discern no contradiction in its description, or in the description of the world in which a mile high unicycle might exist. Both of these cases are uncontroversial. But there is a third category, one that Chalmers himself brings to our attention using the example of Goldbach's conjecture: the category of statements in which we *don't know* whether or not there lies a contradiction. Take the statement that Goldbach's conjecture is true (S1), and the statement that Goldbach's conjecture is false (S2). One may say that he can conceive of the truth of S1, and another that he can conceive of the truth of S2. One of them is misdescribing the world. Suppose Goldbach's conjecture is true. According to Chalmers, if I say that I can conceive of the falsehood of Goldbach's conjecture (by imagining, say, a world where mathematicians announce that it is false) then I am misdescribing the world. It is really a world where Goldbach's conjecture is true and the mathematicians are mistaken. This way if a mathematician should ever prove the truth of Goldbach's conjecture, he will not suddenly transform S2 from conceivable to inconceivable. He will show that S2 had always been inconceivable and that anyone who thought otherwise was making a mistake.

I believe this third category-the category of current uncertainty-receives surprisingly little attention in discussions over the soundness of modal arguments. Yet the move from conceivability to logical possibility depends on one being at least reasonably sure that there is no contradiction in the description of a statement or world. Off-the-cuff conceivability claims come, I believe, when one mistakes statements and worlds that fall into this category of current uncertainty for those that fall in the clearly conceivable

category. Though perhaps we cannot prove *at this time* that a given statement contains a logical contradiction, there may be one all the same. We simply may not have the wherewithal to know for certain.

3. Establishing the Phenomenon

To Katherine Wilkes, thought experiments can be highly misleading as a philosophical tool. The problems arise when the scientist or philosopher (usually the philosopher) does not adequately give us the background conditions against which he is setting his experiment. She cites James Robert Brown (1986) who presents the following thesis about thought experiments:

Thesis I: The burden of any thought experiment rests on the establishment (in the imagination) of a phenomenon. Once the phenomenon is established, the inference to a theory is fairly unproblematic; that is, the jump from data to theory is relatively small...if we got the phenomenon right then *the theory followed more or less automatically*. (Wilkes, 1988, p.8; her italics).

If there is uncertainty concerning the relevant background conditions, however, then it becomes unclear whether or not the thought experimenter has established the phenomenon. At this point:

...our intuitions run awry, and the inferences are not only problematic, but the 'jump' from the phenomenon to the conclusion is made the larger because of the further need to imagine just what these backing conditions, under the imagined circumstances, would be. The 'possible world' is inadequately described. (Wilkes, 1988, p. 8)

So when is the phenomenon established? Brown (1991) offers as an example the famous thought experiment in which a sixteen year old Einstein imagined himself running along side a light beam at the speed of light. If Maxwell's theory were correct, what Einstein would then observe was a spatially oscillatory electromagnetic field *at rest*: "However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations." (Einstein 1954, p53 from Brown 1991, p.15). Thus the thought experiment showed that there were serious internal problems with Maxwell's theory. Now of course it is impossible for anyone to travel at the speed of light (and survive the experience), but this impossibility is *irrelevant* to the conclusion Einstein wished to make. [3](#) As Wilkes points out, the experiment is not aimed at illustrating human tolerances and capacities (if it were, the thought experiment would be ineffective), but rather at an implication of the properties of light as Maxwell described them. As for the possible world in which the experiment is performed, it is our own, with just one distinguishing difference. "Only one factor is juggled, and its impossibility is not relevant

to the conclusion; the relevant remainder stay constant." (Wilkes, 1988, p. 8). Thus the phenomenon is adequately established.

As for thought experiments with *inadequately* established phenomenon, well, according to Wilkes, they are not difficult to find.<4> Her most relevant example is a thought experiment that asks us to imagine a computer or robot that is as intelligent as a human being:

Although it may seem easy to imagine..., this is again an imaginative possibility that *relies on ignorance* rather than upon a well-based judgment that there is a genuine theoretical possibility here. If the 'how' cannot be imagined, the 'that' thought-experimental conclusion becomes decidedly meagre. We have not 'established a phenomenon' in any possible world. (Wilkes, 1988, p. 34, my italics.)

This is precisely the problem with the zombie world and Mary thought experiments. They rely on our ignorance about consciousness to establish their premises.

4. Are Floating Iron Bars Conceivable?

Before getting into zombies, let us examine the conceivability and logical possibility of a floating iron bar (this example comes from Seddon 1972). At first glance, it seems that we can easily conceive of an iron bar floating on water, and so, by Hume's theory, we may conclude that it is logically possible. But upon closer examination, we see that there is at least what Seddon calls an 'implicit logical contradiction' in this concept, "for we are saying that a mineral with a specific gravity of less than 1 (i.e. it floats), has a specific gravity of 7.3-7.8 (i.e. it is iron)." (Seddon 1972, p. 483) So maybe what we conceived of so clearly and distinctly was not iron floating on water, but something that *looked like* iron floating on water. Maybe floating iron bars are not conceivable after all.

Here Chalmers might bring in his distinction between primary and secondary intensions. The primary intension of a concept, Chalmers writes, is its a priori conception; it fixes the reference. The secondary intension is what we acquire after some a posteriori investigation. So, for example, the primary intension of water is what he calls "watery stuff" (an a priori conception that fixes the reference) while the secondary intension of water is H₂O (a result of a posteriori investigation). Now according to Chalmers, it is the primary intensions that apply that to logical possibilities. Thus, there is no possible world in which water is not watery stuff, its primary intension, but there is a possible world in which water is not H₂O, its secondary intension. So we may say that floating iron bars are conceivable and logically possible as long as we use the primary intensions of iron and water rather than the secondary intensions.

But this seems almost absurd enough for a reductio. Are we to say that "irony bary stuff floating on watery stuff" is conceivable and therefore that floating iron bars are logically

possible? Chalmers, it seems, would have to say yes. Otherwise, the move from conceivability to logical possibility is undermined. But his argument depends on there being a clear distinction between primary and secondary intensions, and between a priori and a posteriori necessity. Let us examine this distinction to see how clear it really is.

5. A Priori Vs. A Posteriori Entailment

"The argument against materialism in *The Conscious Mind* has two parts," Chalmers writes,

...The first part, in Chapter 3 of the book, argues that there is no a priori entailment from physical truths to truths about consciousness. The second part, in Chapter 4 of the book, argues that there is no a posteriori entailment from physical truths to truths about consciousness. Either part of the argument might be resisted. Corresponding to these paths of resistance, there are *two very different* brands of materialism, which I call type-A and type-B materialism. (Chalmers, 1996, p.473, my italics)

So what exactly are these two very different brands of materialism? Well, according to Chalmers, the type-A materialist (who is generally a functionalist, behaviorist, or eliminativist) holds that phenomenal truths, insofar as there are such truths, are necessitated a priori by physical truths. The type-A materialist denies that zombies are conceivable. The type-B materialist, on the other hand, in his quixotic attempt to save materialism without denying qualia altogether, concedes that zombies are conceivable, but argues that zombies are *metaphysically* impossible due to what Kripke calls a posteriori necessity.

According to Chalmers, the type-A materialist holds a position that is so "clearly false," that there is only so far one can take the argument. It is a little like trying to argue with a stubborn child. With a *type-B* materialist, on the other hand, one can talk sense. The difference between their position and his own, Chalmers claims, is a subtle, metaphysical one rather than a difference of basic intuition. So although the type-B theorist is a "materialist," he is actually closer in spirit to Chalmers than he is to a type-A materialist. Why? Because like Chalmers, he takes consciousness seriously.

But let us look more closely at what it means for the type-A materialist to hold that phenomenal truths are necessitated a priori by physical truths. If this is true, in the strictest sense of a priori then there are no type-A materialists. In fact, it is unclear what a priori entailment from physical truths to phenomenal truths would even look like. How could anyone hold an opinion about the entailment of physical truths to phenomenal truths without reference to experience? So in this sense, the claim that there is no a priori entailment is true but trivial beyond belief.

But there is a looser interpretation of what counts as a priori. In his Oxford Dictionary of Philosophy Blackburn writes: "it may be that some experience is required to acquire the concepts involved in an a priori proposition." In other words, for a proposition to count as knowable a priori, we may have just enough (but no more) experience to acquire the concepts therein. This, I think, is the sense of a priori that Chalmers employs, something like 'prior to scientific investigation' rather than 'prior to experience.' But now we have entered ambiguous territory. Exactly how much experience do we need to acquire the concept of the term 'phenomenal truths?' Do we take 'acquire' to mean 'understand,' in which case we still have not acquired it (assuming we do not understand consciousness)? Or do we just have a vague notion of, say, the experience of red? Another question: can this concept change after we have acquired it (as for example, our concept of time has changed since Einstein)?

Chalmers' view, as I understand it, is that the primary intension, the a priori conception, of a notion is fixed for all time. Water is and will always remain "watery stuff," and any change in that notion is inconceivable. Our secondary intensions of notions, however, come from a posteriori investigation, and so are more flexible. It is at least conceivable *in principle* that water could be XYZ rather than H₂O. So, has Chalmers made the distinction clear?

Refer back to the floating iron bar. All would agree that such a concept is impossible. In one sense, the impossibility is a result of Kripkean a posteriori necessity—we measure specific gravity out in the world. In this sense, there is no a priori entailment from the property of 'iron bar' to the property of 'cannot float on water.' (The eminently reasonable type-B materialist might call this concept logically possible but *metaphysically* impossible.) On the other hand, having a specific gravity of between 7.3 and 7.8 can be regarded as a *defining characteristic of iron*. As Seddon points out, there is no doubt that the mineralogist regards it as one. If an unknown sample is determined to have a specific gravity out of this range, then it will not be classified as iron. And if the mineralogist sees the sample floating on water, he will dismiss the possibility of it being iron out of hand, without having to do any tests whatsoever. Furthermore, the only way a mineralogist (or anyone for that matter) can distinguish between a sample of pure iron and, say, an iron alloy, is by appeal to secondary intensions. There is no difference between the primary intensions. Pure iron stuff and impure iron stuff look and feel exactly the same.

Here we can see just how vague the distinction between primary and secondary intensions (as well as the distinction between a priori and a posteriori necessity) can be. Logical possibility thus becomes a blurry concept that allows for a posteriori considerations. As long as we agree that specific gravity is a defining characteristic of iron, then floating iron bars—whether conceivable or not—are logically impossible. Something cannot both have and not have a specific gravity of less than one (Wilkes, 1988). If it can float on water it is not iron. P and not P.

6. Are Zombies Conceivable?

So how does all this relate to zombies? Well, according to Chalmers, "if we can show that there are possible worlds that are physically identical to ours but in which the property introduced by the *primary intension* is lacking, then dualism will follow." (p.132, my italics) And this, he argues, is exactly what he did with the property of consciousness when he established that there is no a priori entailment from physical truths to phenomenal truths.

But there is a sleight of hand being pulled here. Recall that the sense in which there was no a priori entailment was trivial. Once we allow for experience to help us determine a priori truths, we open the door for a priori entailment.<5> Chalmers simply assumes that there is a clear distinction between the primary (a priori) and secondary (a posteriori) intensions, when in fact there is this great ambiguity. What is the primary intension of consciousness? Where, in the case of concepts like consciousness and phenomenal truths, do the primary intensions stop, and secondary intensions begin?

It is true that unlike the case of the iron bar we cannot point to a clear contradiction in the description of the zombie world. Not yet, anyhow. But our knowledge of the brain and the mind is not nearly as developed as our knowledge of metals. It is certainly possible that as we acquire a greater understanding of neuroscientific concepts and processes, we will see the logical contradiction, just as we have in the case of the floating iron bar. Therefore it seems we should place zombie worlds alongside Goldbach's conjecture in the third category of conceivability claims, the category of current uncertainty. This is a 'wait and see' category. Statements therein cannot be used to repudiate time-honored metaphysical positions like materialism. And if my arguments above are correct, the distinction between a posteriori and a priori entailment, blurry as it is, cannot be used to distinguish between Chalmers' conjecture and Goldbach's.

Before defending this claim further, let me make the situation more explicit. My zombie twin, remember, is my exact physical duplicate, down to the last neuron. All the brain processes that occur in me, occur in him as well. Now suppose that someone comes up to my zombie twin and asks him why he looks so sad. He replies: "I'm sad because it is snowing outside. There's something melancholy about the whiteness of snow. I think it reminds me of my childhood in Vermont when I used to sit at the window on winter days. I would sit there for hours, the warm fire at my back, just thinking about how white and pure everything was." Now if I say this, then my zombie twin will say it too, and the speech will be the result of the same neurological processes. But the zombie has no consciousness. He does not experience whiteness or warmth or tangible feelings of sadness. What is it that makes him carry on so sentimentally?

Again, I am not arguing that the zombie's behavior is logically *impossible*, only that it is reasonable to suspect that (to use Chalmers' phrase) a contradiction lurks somewhere in this description. True, we cannot identify the contradiction now, but it could be that in fifty years, or a hundred, or a thousand, we will be able to.

Chalmers gives this kind of objection a fair showing. He writes:

So the only route available to an opponent here is to claim that in describing the zombie world as a zombie world, we are misapplying the concepts, and that in fact there is a conceptual contradiction lurking in the description. Perhaps if we thought about it clearly enough we would realize that by imagining a physically identical world we are thereby *automatically* imagining a world in which there is conscious experience. But then the burden is on the opponent to give us some idea of where the contradiction might lie in the apparently quite coherent description. (Chalmers, 1996, p. 99)

This, then, is the key question: is the description really apparently coherent? Is the zombie world *clearly* conceivable? It is to Chalmers, of course: "I confess that the logical possibility of zombies seems equally obvious [as a mile high unicycle] to me," he writes, "...almost everybody, it seems to me can conceive of this possibility." (p.96) Well, all I can say is: not me. I do not find the conceivability of zombies obvious at all. When I think back to the zombie who talks about the melancholy white snow and the warm fire, I think it is very possible that I am misapplying concepts somewhere and that there is a logical contradiction in my description. What I can do very easily is conjure up a mental picture of a person behaving in such a manner, and then add to myself "and that person is a zombie." (This, I believe, is what "almost everybody" can do.) But conceiving of the two together, the zombie (no consciousness) and the behavior, and coming up with a causal story behind the behavior, is much more difficult. At any rate, I think the matter is just as "open" as is the truth or falsehood of Goldbach's conjecture.

So now we have arrived at the dreaded 'burden of proof' juncture in the debate. On which party does the onus lie to make their case? I agree with Chalmers that there is only so much room for productive argument here, but I would point out the asymmetry in the relative strength of claims. Chalmers claims that the zombie world is clearly conceivable. The materialist argues that zombie worlds are quite possibly inconceivable, that the matter is open. Chalmers arrives at the world-shaking conclusion that "materialism is false." The materialist argues not that materialism is true, but simply that we do not have the wherewithal to determine whether zombies are truly conceivable (or logically possible), and therefore that we cannot use zombies as a tool to evaluate the truth or falsity of materialism—a more modest conclusion, to say the least. Given this asymmetry, I believe it is not the materialist's responsibility to point out the logical contradiction, any more than it would be Goldbach's responsibility to show the contradiction in the logical possibility of the claim that his conjecture is false.

Chalmers, of course, takes the opposite view:

In general, a certain burden of proof lies on those who claim that a given description is logically *impossible*. If someone truly believes that a mile-high unicycle is logically impossible, she must give us some idea of where a contradiction lies, whether explicit or implicit...If no reasonable analysis of the terms in question points to contradiction, then there is a natural assumption in favor of logical possibility (Chalmers, 1996, p. 96)

But what exactly does Chalmers mean here by "in general"? And where does this natural assumption in favor of logical possibility come from? There is no law of nature involved, nor any rule of logic. Perhaps it is a time-honored custom in conceivability arguments (although, if so, I'm not familiar with it). But I do not think there are any general or natural rules regarding burden of proof claims (which is why they tend to be such conversation-stoppers). Furthermore, even if there were such a rule for cases such as mile-high unicycles, it would not apply to zombie worlds for the simple reason that we understand unicycles, and we do not understand consciousness. In the case of zombies, a conclusive analysis of the terms in question is just not possible. And so it is perfectly reasonable to suspect that there may be a contradiction in the description. Finally, recall Brown's remark that the burden of any thought experiment rests on the establishment of the phenomenon. Once this is accomplished, he writes, "the theory flows more or less automatically." If Brown is correct then merely to insist or assume that zombie worlds are conceivable (that is, that the phenomenon is established) is in a very real sense just begging the question.

7. Mary the Color Scientist

Chalmers does not want us to think that *everything* depends on zombies (although one may be forgiven for concentrating heavily on this argument, since it is the one that he refers to over and over again throughout the book). It is true, however, that he presents four other arguments against "a priori entailment." The most effective—at least dialectically—Chalmers writes, is Frank Jackson's Mary argument "because a conceivability claim is easy to deny, a new knowledge claim much harder." (Chalmers, 1999, p.476) In this section I will try to show that one may use a similar strategy to undermine the Mary thought experiment.

The argument (from Jackson, 1982) goes like this. Mary is a brilliant scientist who has spent her entire life in a black and white room, forced to investigate the world through a black and white television monitor:

She specializes in the neurophysiology and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red,' 'blue,' and so on...What will happen when Mary is released from the black and white room or is given a television monitor? Will she *learn* anything or not. It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had *all* the physical information. Ergo, there is more to have than that, and Physicalism is false. (p.130)

The first thing to note, again, is the strength of the conclusion: 'Physicalism is false.' The second is Jackson's use of the phrase "it seems just obvious" (that Mary will learn something when she is released, and sees, say, a rose). This may be rephrased, fairly I

think, as 'it is *unimaginable* that Mary will *not* learn anything from the rose.' So then the argument would run as follows:

1. Mary has (let us suppose) acquired all the physical knowledge there is about color.
2. When Mary sees a rose for the first time, it is unimaginable that she will not learn something from the experience.
3. For physicalism to be true, it is necessary that Mary not learn anything from the rose.
4. Ergo physicalism is false.

For this argument to work, however, we need to do more than "suppose" that Mary has acquired all the physical information about color; we need to *imagine* it. Otherwise the argument will have this general form:

1. Suppose X (X is unimaginable).
2. Given X, *you can't imagine* that Y.
3. Y given X is necessary for the possibility of physicalism.
4. Ergo physicalism is false

Clearly this is not an effective argument. We have two unimaginable states of affairs, only the second of which is used to refute physicalism. But really the two unimaginabilities cancel each other out. Unless we can make sense of what it means for Mary to have *all* the physical information there is to have about color, the phenomenon will remain unestablished in our minds, and any conclusions that follow from "supposing" this fact are open to doubt.

To take an unestablished phenomenon like this one as an acceptable premise will cause a lot of unnecessary trouble, yet many philosophers have gone this route and founded their objections on other considerations. Churchland (1985) and Dennett (1991) have taken the appealing yet implausible position that if Mary truly has all the physical information about color, then we *can* imagine her not learning anything from the red rose. Others grant that Mary discovers *something* when she sees the rose but try to defend physicalism using other means (for example, Lewis' and Nemirow's 'ability hypothesis'). I will not evaluate these arguments in this paper, except to note that many of them may be quite effective in dismantling Jackson's argument even if we allow him the first premise. I do want to suggest, however, that perhaps we should not be so quick to grant him that first premise. Dennett phrases the problem nicely when he says that what we are asked to imagine is "so preposterously immense, you can't even try." He continues:

The crucial premise is that "She has *all* the physical information." That is not readily imaginable, so no one bothers. They just imagine that she knows lots and lots-perhaps they imagine that she knows everything that anyone knows *today* about the neurophysiology of color vision (Dennett, 1991, p.399).

It is not only not readily imaginable, it is not imaginable period. But then why go on? Why claim, as Churchland (1985) does, that "we can indeed imagine how neuroscientific information would give Mary detailed information about the qualia of various sensations," (p. 25) when this just seems false. To me anyway, it is unimaginable that Mary can obtain qualitative experience from physical knowledge, but-and this is the point-it is no more unimaginable than Mary *having* all the physical information in the first place. And because this first premise is unimaginable-that is, the phenomenon is not established-any conclusions that follow are open to doubt.

It may be objected that I am just being stubborn. It is no great philosophical trick to just insist that a thought experiment is unimaginable, that a phenomenon is not adequately established. If one side simply declares that that Mary's omniscience is imaginable and the other that it is not, then the dialogue will just stop and neither side will be convinced. But if what I argue is right, the dialogue *should* stop. If Jackson has not established the phenomenon in his thought experiment, then his argument, like the zombie argument, has nothing to say about the truth or falsehood of materialism (which is, after all, what we are supposed to be investigating). Granting Jackson this or that premise for the sake of argument, or to avoid deadlock, will be at best only tangentially useful. Moreover, if Brown is right that adequately establishing the phenomenon is the difficult part of a thought experiment-that theory flows easily once this is accomplished-then (as was the case with the zombie worlds) Jackson's assumption that Mary's omniscience is imaginable is just question-begging.

8. Conclusion: 'Don't-Have-A-Clue' Materialism

According to Gerald Massey, "one's readiness to accept off-the-cuff conceivability claims is inversely proportional to the importance of the matters at issue." (Massey 1991, p.294) As soon as any really significant mathematical or scientific doctrine is threatened, the standards for true conceivability are immediately elevated, and no off-the-cuff claims are tolerated. I would add to this that casual acceptance of conceivability claims seems also to vary directly with our ignorance of the subject matter at hand. It is perhaps for this reason that the strangest, most uncritically accepted conceivability claims can be found in philosophy, and particularly in philosophy of mind.

In his description of the various materialist positions, Chalmers gives the following, which he says is held widely, though rarely in print:

Don't-have-a-clue materialism. "I don't have a clue about consciousness. It seems utterly mysterious to me. But it must be physical, as materialism must be true. (Chalmers, 1996, p. 162)

Chalmers means this as satire, but I think this is by far the most reasonable position to take, especially if we revise it slightly to: "I don't yet know enough about consciousness to make any real substantive claim about it. But it is *probably* physical, as materialism is *probably* true, given its past success in explaining things that were once deemed 'utterly mysterious.'"

It was once thought that it would be impossible to give a materialist account of the origins of human life on Earth. But with the advance of the biological sciences came a satisfactory materialist theory. Neuroscience, as a discipline, is still in its infancy. Before making premature pronouncements about the truth or falsehood of materialism, why not wait until this scientific discipline matures? Perhaps (as occurred after Einstein in Physics) we will acquire some new concepts that will allow us to make real progress on this question.

Notes

<1>. Unless I note otherwise, all citations from Chalmers will come from Chalmers 1996.

<2>. Stephen Cade Hetherington (1991) offers an interesting objection to this last assertion of Hume's, giving the paradoxical pictures of M.C. Escher as a counterexample. "The pictures are graphic-extremely clear and distinct-in their detail," he writes, "yet they remain representations of impossibilities." (p.320)

<3>. By contrast, the impossibilities are quite relevant in the case of Mary and the Zombies.

<4>. Wilkes primarily targets thought experimenters who focus on personal identity. Her criticisms, however, apply equally well against Jackson and Chalmers.

[5](#). Or, put more precisely, the proposition that phenomenal truths are necessitated by physical truths a priori, becomes one that may in fact be true.

References

Blackburn, S (1994). *A Dictionary of Philosophy*. Oxford University Press.

Brown, J.R. (1991). *The Laboratory of the Mind*. Routledge Press.

Chalmers, D.J. (1996). *The Conscious Mind*. Oxford University Press.

Chalmers, D.J. (1999). Materialism and the Metaphysics of Modality. *Philosophy and Phenomenological Research*, LIX(2), 473-496.

Churchland, P (1985). Reduction, Qualia, and the Direct Introspection of Brain States. *The Journal of Philosophy*, 82, 8-28.

Dennett, D.C. (1991). *Consciousness Explained*. Boston: Little, Brown.

Hetherington, S.C. (1991). 'Conceivability and Modal Knowledge.' *From Thought Experiments in Science and Philosophy*. Horowitz and Massey eds. Rowman & Littlefield Publishers, Inc.

Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly* 32:127-36.

Massey, G. (1991). Backdoor Analyticity. *From Thought Experiments in Science and Philosophy*.

Seddon, G. (1972). Logical Possibility. *Mind* Vol. LXXXI No. 324.

Wilkes, K. (1988). *Real People*. Clarendon Press. Oxford.