PHYSICALISM COULD BE TRUE EVEN IF MARY LEARNS SOMETHING NEW

Philosophical Quarterly, April 2007

Barbara Gail Montero

Mary knows all there is to know about physics, chemistry and neurophysiology, yet has never experienced colour. Most philosophers think that if Mary learns something genuinely new upon seeing colour for the first time, then physicalism is false. I argue, however, that physicalism is consistent with Mary’s acquisition of new information. Indeed, even if she has perfect powers of deduction, and higher-level physical facts are a priori deducible from lower-level ones, Mary may still lack concepts which are required in order to deduce from the lower-level physical facts what it is like to see red.

Mary knows all there is to know about physics, chemistry and neurophysiology. Yet she has spent her entire life in a black and white environment. Is there something about the experience of, say, seeing red which she does not know? It is often assumed that if Mary learns something genuinely new upon seeing red for the first time (that is, does not merely come to understand in a different way something she already knows), then physicalism is false.¹ This conclusion, however, does not follow: Mary can acquire genuinely new knowledge upon leaving her achromatic environment even if physicalism is true. Indeed, I shall argue that even if we assume both that all higher-level physical facts (e.g., presumably, facts about rocks and trees, and, if it is physical, about experience) are a priori deducible from the lower-level physical facts (e.g., presumably, facts about neurology, chemistry and physics), and also that Mary has perfect powers of deduction, her coming to learn something genuinely new is still consistent with physicalism.

While there are various criticisms of ‘the Mary argument’, as I shall call it – for example, that no one could learn all the relevant facts of physics,

chemistry and neurophysiology, or that it is impossible to imagine what someone who has learnt all these facts could know, or that Mary would not learn anything upon seeing red for the first time, or that she learns only a new way of understanding something which she already knows – practically all participants in the debate accept that if Mary has perfect powers of deduction, and if she learns a genuinely new fact, then experience cannot be physical. It is this widely accepted view, however, which I deny. In essence, my point is that perfect powers of deduction do not ensure that Mary has the relevant concepts required in order to carry out the deduction. Thus I shall be presenting a version of what David Chalmers has called ‘the under-discussed missing-concept reply to the knowledge argument’.

I. THE LESS THAN IDEAL MARY ARGUMENT

I began with a description of the Mary argument which, while capturing a common way of understanding it, is less than ideal. As a way of leading up to a more careful formulation of the Mary argument, and ultimately to my criticism of it, I shall examine this less than ideal version of the argument, in order to show why it is not valid.

The less than ideal Mary argument

Mary, who has spent her entire life in a black and white environment, knows all the facts of physics, chemistry and neurophysiology relevant to colour vision.

When Mary sees red for the first time, she learns something genuinely new: that seeing red is like this, where ‘this’ refers to the visual experience of seeing red.

Therefore physicalism is false.

---

This less than ideal Mary argument is not valid, because the facts of physics, chemistry and neurophysiology relevant to colour vision need not exhaust all the physical facts relevant to colour vision. When we understand the notion of the physical broadly – that is, so that rocks and trees count as physical, along with the fundamental particles of physics – we can allow for physical phenomena above the level of neurology. And for the non-reductive physicalist, the property of being an experience of red is just such a physical phenomenon: it is broadly physical, since it is determined in the appropriate way by fundamental physical properties, even though it is not a property in physics, chemistry or neurophysiology.\(^7\) Or, to put it slightly differently, the experience of seeing could be a higher-level physical feature of the neural, even if the fact that seeing red is like *this* is not itself a fact of physics, chemistry or neurophysiology. This argument, then, does not refute (non-reductive) physicalism, since it is consistent with non-reductive physicalist views that the mental is a higher-level feature of the neural.\(^8\)

Although this seems to be a rather obvious point, it is one of those odd ‘obvious’ points which often go unnoticed. Part of the reason for this may be that when the argument is laid out, the first premise is sometimes stated in terms of physical facts: Mary is said to know all the physical facts relevant to colour vision. And given in this way, the argument is not obviously invalid. However, we need to ask ‘What is meant here by “physical facts”?’ If the notion of physical facts in the first premise is understood as being facts of the relevant sciences (that is, physics, chemistry or neurophysiology), and if we allow, as the non-reductive physicalist does, that experience can be physical even if it is not countenanced by the relevant sciences, the argument will rest on equivocation with the term ‘physical’. Moreover, avoiding equivocation will not be easy, since understanding ‘physical facts’ in the first premise as ‘facts of the relevant sciences’ is an important step in guiding us how to

---

\(^7\) The general idea of determination is that F properties determine G properties if and only if given the F properties, the G properties must exist with at least nomological necessity. Although many agree with the idea that some sort of determination relation along these lines suffices for physicalism, there is much disagreement about how to formulate the relation precisely. Importantly, the relation needs to be strong enough to formulate a physicalism that rules out such things as psychophysical parallelism and a necessarily existing God. If one thinks that there are no necessary connections between distinct existences, a simple determination relation, such as the one stated above, may suffice for physicalism. However, for an illuminating discussion of how to formulate a determination relation that suffices for physicalism without assuming that there are no necessary connections between distinct existences, see J. Wilson, ‘Causal Powers, Forces and Superdupervenience’, *Grazer Philosophische Studien*, 65 (2002), pp. 117–121.

think about the argument: certainly, it is claimed, someone without colour experience could (at least in principle) learn all the relevant facts of physics, chemistry and neurophysiology. So we are led to imagine Mary as having mastered physics, chemistry and neurophysiology, or at least those parts relevant to colour vision, and then we slip into thinking of these as the physical facts relevant to colour vision. But non-reductive physicalism can allow for further physical facts, such as the fact that seeing red is like this, where ‘this’ refers to the visual experience of seeing red.

One avoids this slip, of course, if one thinks of the relevant physical facts in both the first premise and the conclusion as the facts of the relevant sciences. But this is a mistake, since mental facts need not be facts of the relevant sciences in order to be consistent with physicalism – in order, that is, to be broadly physical. If mental facts are necessitated (in the strongest sense), or determined, by lower-level physical facts, this, as many agree, suffices for the mental facts to be physical in the broad sense, that is, in the sense in which if physicalism is true, everything is physical. Science aims, or at least should aim, to understand the world and everything in it, whether gravity or ghosts. Yet there may be limits to the dominion of science. And if there are, some physical things may lie outside these limits.

Alternatively, one might try to repair this argument by understanding the notion ‘physical’ throughout in the broad sense, that is, as the fundamentally physical, and whatever is determined by this. But if we take the first premise of the argument to state that Mary knows all the broadly physical facts relevant to colour vision, then we have just stipulated that she already knows what it is like to see red, since that seeing red is like this is a broadly physical fact relevant to colour vision.

But what makes something physical? This is a much debated question. However, as a first step one can say that as long as a property is either fundamental and physical, or determined (in the relevant way) by fundamental physical properties, its existence is consistent with physicalism. What, then, are the fundamental physical properties? Since almost everything I shall say is consistent with a variety of interpretations of the fundamental physical (the exception will be clear when I reach it), I shall leave this question aside for now and proceed to explain a few other notions. I use the phrases ‘fundamental physical properties’ and ‘narrowly physical properties’ to refer to those properties which are both fundamental and physical. For example, barring panpsychism, the property of being an up-quark is a narrowly physical property (if it actually is fundamental). Then the broadly

9 Again a simple determination relation may not suffice for physicalism. (See fn. 9 above for further thoughts on this.)
physical properties are either fundamental physical properties or properties appropriately determined by the fundamental physical properties. For example, being a rock is broadly physical because this property is entirely determined by fundamental physical properties. The experiential properties of seeing red, then, are broadly physical if they are entirely determined by fundamental physical properties. (The determination relation is transitive. For example, if experiential properties are entirely determined by neural properties, and the neural is entirely determined by fundamental physical properties, then experience is broadly physical.) Physical facts are facts which are consistent with the truth of physicalism. For example, that carbon has a valence of four, and that there is a pen on my desk, are presumably physical facts. Non-physical facts are inconsistent with the truth of physicalism. For example, if the experience of seeing red is not physical, then the fact that seeing red is like this is a non-physical fact, since the demonstrative ‘this’ then refers to a non-physical property. Since mental properties such as experiencing red, being in pain, believing that the sun will rise, and so forth, are, presumably, not fundamental physical properties, physicalism implies that they are broadly physical and that all the facts about experience, including that what it is like to see red is like this, are physical facts.

II. CAN WE DEDUCE HIGHER-LEVEL PHYSICAL FACTS FROM LOWER-LEVEL ONES?

So far I have pointed out that the less than ideal Mary argument is compatible with the truth of physicalism: given physicalism, Mary can still learn a new fact upon first seeing red, even though she knows all the facts of physics, chemistry and neurophysiology. One cannot repair the argument either by understanding the notion of the physical throughout as the broadly physical, or by understanding it throughout as the narrowly physical. The remedy, however, is to specify that Mary in her achromatic environment knows all the broadly physical facts below some level at or below the level of experience, and to aim for the conclusion that the mental is not broadly physical. So suppose Mary knows all the (broadly) physical facts below the level of experience. Would she be surprised upon seeing red for the first time? The answer to this question depends on what Mary can deduce from lower-level physical facts. If the higher-level physical facts about the world are in general not deducible from lower-level physical facts, then even if the experience of seeing red is physical, Mary will not be able to deduce what it is like to see red from the fundamental physical facts. So if
this is the case, she will learn something new upon seeing red for the first time, even if the experience of seeing red is entirely physical.

Some, however, think that given the appropriate bridge laws, all higher-level physical facts can, in principle, be deduced \textit{a priori} from lower-level, and ultimately fundamental, physical facts.\textsuperscript{10} For example, it is often pointed out that facts about water, such as that it boils at \(100^\circ\) C at sea-level, follow \textit{a priori} from facts about \(H_2O\). Other physical facts might be more complex; but, it is argued, this makes the deduction only more difficult, but not impossible, or at least not in principle impossible. If so, and if what it is like to see red is a physical fact, then Mary should be able to know what it is like to see red in her achromatic environment.

Is it true that all the higher-level physical facts can be \textit{a priori} deduced from the fundamental physical facts? When philosophers discuss the \textit{a priori} deducibility of the higher-level facts from the fundamental physical facts, the fundamental physical facts are typically thought of as the facts of fundamental physics. This is relevant because one line of argument for the conclusion that the higher-level physical facts must be \textit{a priori} deducible from the fundamental physical facts depends on a particular interpretation of ‘physics’. The line of argument I have in mind is, roughly, that all the broadly physical facts are \textit{a priori} deducible from fundamental physical facts, because both the fundamental physical facts (on this view, those given to us by physics) and the higher-level physical facts are structural/relational facts, that is, facts which tell us about the structural/relational features of the world (rather than its intrinsic nature); and it is thought that there is no barrier which in principle blocks deriving structural/relational features of the world from other structural/relational features of the world. Thus on this view, if physicalism is true, then all the facts about higher-level phenomena are in principle deducible from the fundamental physical facts.

This sort of argument is successful, however, only if the physical facts are structural/relational facts. For example, a similar argument cannot be advanced by anyone who understands, as some do, the fundamental physical as the fundamental non-mental, and the broadly physical as whatever is appropriately determined by the fundamental non-mental.\textsuperscript{11} This is because some fundamental non-mental properties may not be structural/relational

\textsuperscript{10} See, for example, Jackson, ‘Armchair Metaphysics’, in M. Michaelis and J. O’Leary-Hawthorne (eds), \textit{Philosophy in Mind} (Dordrecht: Kluwer, \(\text{C}\)). pp. \(\text{C}-\text{C}\). Chalmers and Jackson, ‘Conceptual Analysis and Reductive Explanation’, \textit{Philosophical Review}, \(\text{C}\) \(\text{C}\). pp. \(\text{C}-\text{C}\).

\textsuperscript{11} See my ‘Physicalism in an Infinitely Decomposable World’, \textit{Erkenntnis}, \(\text{C}\) \(\text{C}\). pp. \(\text{C}-\text{C}\), and ‘Post-Physicalism’, \textit{Journal of Consciousness Studies}, \(\text{C}\) \(\text{C}\). pp. \(\text{C}-\text{C}\), where I defend a more precise version of this position. For similar views see J. Levine, \textit{Purple Haze: the Puzzle of Consciousness} (Oxford UP, \(\text{C}\)). pp. \(\text{C}-\text{C}\); J. Wilson, ‘On Characterizing the Physical’, \textit{Philosophical Studies}, forthcoming; Papineau, \textit{Thinking about Consciousness}, pp. \(\text{C}-\text{C}\).
properties: the fundamental non-mental might include the intrinsic nature of matter (if there is such a thing), which, according to many, physics can never reveal. And the higher-level physical properties that are determined by these fundamental properties may also not be structural/relational, and thus the argument will not have shown them to be a priori deducible from the fundamental physical properties. Similarly, we do not arrive at the desired conclusion if, as others do, we take the fundamental physical properties to be, roughly, exhausted by the kinds of fundamental properties found in ordinary inanimate objects. Some of these properties may not be structural/relational properties, for there is no guarantee that they will all be, in principle, accessible to physics (or at least to a physics capable of revealing only structure and relations). And again such properties may determine higher-level properties that are not a priori deducible from the fundamental physical properties. So this reason for accepting the a priori deducibility of the higher-level physical facts from the lower-level ones relies on a certain understanding of the physical. If this understanding is rejected, the argument can be rejected too.

III. THE IMPROVED MARY ARGUMENT

I have shown that one argument for the view that the higher-level physical facts are a priori deducible from the fundamental physical facts falters under certain interpretations of the physical. It would be rash, however, to rely on this alone in denying the claim that the higher-level physical facts are a priori deducible from the fundamental physical ones, since other arguments for this claim may not depend on understanding the physical as the structural/relational. And fortunately I do not need to pursue this approach, since for the purposes of argument I am going to accept that the higher-level physical facts about the world are a priori deducible from lower-level, and ultimately fundamental, physical facts. So I shall assume that if physicalism is true, the higher-level features of the world are deducible, or are at least in principle deducible, from the lower-level and ultimately fundamental

---

12 If there is an intrinsic nature of matter that is non-mental, I think we should not count this as ‘protophenomenal’, since, if I understand the concept, the protophenomenal is neither mental nor non-mental. See, however, Chalmers, ‘Phenomenal Concepts and the Knowledge Argument’, in Ludlow et al. (eds), There's Something About Mary, pp. 179-207, his response to what he calls ‘the incomplete-physical-knowledge reply’.

features of the world. (And, moreover, that if physicalism is true, mentality is a higher-level feature of the world.) Of course, physicalism is still consistent with the idea that all the higher-level physical facts about the world are a priori deducible from the fundamental physical facts, provided that Mary is able to deduce what it is like to see red in her achromatic environment. Would she be able to do this? Again for the purposes of argument, I shall simply accept that she cannot. What now becomes of the Mary argument? I have now arrived at an improved formulation of it:

*The improved Mary argument*

1. Mary, who has perfect powers of deduction, and has spent her entire life in a black and white environment, knows all the physical facts relevant to colour vision below the level of experience
2. The higher-level physical facts are ultimately a priori deducible from the fundamental physical facts
3. When Mary sees red for the first time she learns something genuinely new, that seeing red is like *this*, where ‘this’ refers to the visual experience of seeing red
4. Thus *that seeing red is like this* is not a physical fact
5. Therefore physicalism is false.

**IV. DEDUCTIONS A PRIORI MAY REQUIRE EXPERIENCE**

I claim that even the improved Mary argument is not valid. Mary may have perfect powers of deduction, and the higher-level physical facts may be a priori deducible from the fundamental physical facts; but this does not mean she will be able to deduce all the physical facts about experiencing colour from the fundamental physical facts. And the reason for this is that she may have been deprived of something essential for deducing this particular fact, even if this fact is physical. In other words, it might be that even though an understanding of what it is like to see red is a priori deducible from the lower-level physical facts, Mary, even with her perfect powers of deduction, may not be able to arrive at this understanding because colour experience may be necessary for her to understand the relevant terms involved in the requisite deduction.

I have just said that a type of experience might be necessary to make a particular a priori deduction. Have I misunderstood the a priori? Although it is sometimes thought that a priori deduction is entirely independent of experience, this supposition is not correct. A deduction can count as a priori even if one needs experience in order to grasp the concepts involved in the
deduction. Or, in other words, a deduction can count as *a priori* even if one needs experience in order to have a rudimentary sense of what the relevant terms in the deduction refer to or mean.\(^\text{14}\)

What, then, makes a conclusion deducible *a priori* from a set of premises? It is not that one can perform the deduction without ever having had experience. And it is also not that one can arrive at the conclusion without relying on any particular experience, since many conclusions one arrives at empirically could have been arrived at via a sensory means other than the particular one used. Rather, it is that one can justify the conclusion without testing it in the world, that is, one can justify the conclusion by reasoning in the armchair, as it were.\(^\text{15}\) And if one has the appropriate colour experiences, such as being exposed to red squares which have been labelled ‘red’, it may be possible to justify the view that seeing red is like *this* merely by thinking about the complex physical story about the brain, the physics of light, and so forth.

Kant (*Critique of Pure Reason*, [footnote]) allowed that some knowledge, although rightfully *a priori*, involved concepts which ‘can be derived only from experience’. For example, he says, we know *a priori* that every alteration has a cause, even though we need experience in order to understand the concept of alteration. Similarly, it seems that although we know *a priori* that all sisters are siblings, we cannot understand what this means, and thus cannot know (*a priori* or otherwise) that all sisters are siblings, unless we understand, among other things, what it is to be a sister. Yet understanding this requires experience. Not that you need to have experienced sisterhood – someone may have explained the concept to you – but some type of experience at least was necessary to acquire the building blocks out of which you could

\(^\text{14}\) I have cashed the notion ‘a concept of *p*’ as that which gives us a rudimentary sense of *p*. It is important for my purposes here that our rudimentary sense of *p* does not give us everything, or everything relevant, which follows from *p*, for if it did, then one would not need to deduce anything *a priori*. However, a concept of *red* should give one more than just what some refer to as a ‘phenomenal concept’, which arguably can be acquired merely by being exposed to randomly coloured things, for example, tables, abstract paintings, coloured walls, etc., but no ripe tomatoes; for in order to understand the conclusion of her deduction, Mary needs to understand that the demonstrative *this* is identifying the visual experience of red: that is, she needs to be able to associate the term ‘red’ with a particular visual appearance. For an argument that phenomenal concepts are not sufficient for Mary’s knowing what it is like for others to see, say, ripe tomatoes, see M. Nida-Rümelin, ‘What Mary Couldn’t Know’, in T. Metzinger (ed.), *Conscious Experience* (Paderborn: Schöningh, 2005), pp. 329–33.

construct what it is to be a sister. (In principle, I suppose one could be born knowing what a sister is; however, this does not show that our knowledge that all sisters are siblings is a priori, since one could similarly be born knowing, for example, that the sun is ninety-three million miles from the earth.)

Just as some a priori knowledge involves concepts which one can acquire only through experience, some deductions are a priori, yet involve concepts which one can acquire only through experience. In other words, we should accept that in order to carry out an a priori deduction, rightly so-called, one may need a rudimentary understanding of the terms involved, and that in some cases one may need experience to acquire this rudimentary understanding. For example, most who traffic in the a priori would accept that given the premises that Jane is taller than John, and John is taller than Sally, we can deduce a priori that Jane is taller than Sally. But understanding the idea of one thing’s being taller than another requires experience, and while we may be able to state this conclusion without understanding what it is for one thing to be taller than another (as long as we understand that the relation is transitive), it seems that we cannot understand it without ever having the relevant experiences. Mary also, inside her achromatic environment, may be able to state the conclusion of her investigations thus: ‘Ah, from the following neural patterns I can infer that this is what the experience of seeing red is like; I can see that seeing red is like this’. For she may have figured out that those neural patterns are the neurological basis of experiencing red. However, she may not understand her conclusion; she may not understand what ‘this’ refers to. The experience of seeing red might just be a necessary physical ingredient for this understanding; it might be a necessary ingredient for acquiring what David Chalmers refers to as ‘the missing concept’.

Mary, of course, does not lack all experience, only colour experience. And one might think that while some type of experience may be required in order to understand the relevant terms of one’s a priori deduction, specific sorts of experiences should not be required. The amazing abilities of blind mathematicians might be thought to support this view. There have been numerous blind mathematicians, most notably Euler, who was blind for the last seventeen years of his life, years during which he produced nearly half of his prodigious output. But what this shows is certainly difficult to say, since lack of vision does not always go with lack of visual images. Yet Mary, we are supposed to imagine, has not only been deprived of seeing colours, but has never had any sort of colour experience, not even in dreams or in

mental visualizations. Moreover, blind mathematicians are, from what I have been able to gather, invariably adventitiously blind, that is, they have lost their sight after the development of the visual cortex. So it is possible that some prior visual experiences are even required in order to carry out a priori mathematical deductions. The deduction would count as a priori, since once one has had the experiences, the deduction need not rely on them. And if mathematical deduction requires a background of visual experience, it would not at all be surprising that certain a priori deductions about colour vision, in particular deducing from the relevant facts about neurology, chemistry and physics that seeing red is like this, require a background of colour experience.

One should remember that the type of deduction that Mary would need to accomplish is remarkable on any account, for no one thinks that the a priori deducibility of all the higher-level physical facts, or even all those which are relevant to colour vision, from the fundamental physical facts, or even from the relevant lower-level facts, is something which anyone with our mental capacities could master. Mary, however, is no ordinary person: she has, as has been stipulated, perfect powers of deduction. Thus the a priori deduction of which she is thought to be capable is supposed to be only in principle possible, that is, possible from a God’s eye point of view, as it were. But this still does not show that Mary should be able to deduce that seeing red is like this from the lower-level physical facts without ever having seen anything red. Perhaps, if physicalism is true, God could deduce what it is like to see red from lower-level physical facts; but while God and Mary both have perfect powers of deduction, God, unlike Mary, has never been deprived of colour experience.

To be sure, the a priori does not include knowledge based on introspection. I might, while sitting in an armchair, as it were, be able to justify the statement that I am in pain, even though my knowledge of being in pain is not a priori. Will this sort of introspection necessarily be involved in Mary’s justification of the complex lower-level physical story from which the idea is supposed to follow that the experience of red is like this? It is, of course, difficult to say. However, it seems that for those with the normal range of experiences, it is possible to understand the experiential aspect of what it is like to see red, or what it is like to be in pain, without simultaneously introspecting an experience of red or feeling of pain. And so I would think that as long as she has the concept of red, Mary, similarly, could deduce that seeing

18 Chalmers raises an objection along these lines to the view I am putting forth, in his ‘Phenomenal Concepts and the Knowledge Argument’, p. 142.
red is like this without relying on introspection. This is not to say that she could arrive at the idea without ever having an experience of red, but just that her deduction of the fact that seeing red is like this (once she has the relevant understanding of red) need not rely on introspection.

(One might still think that introspection will be necessarily involved here. However, if it is, it seems also to be involved in mathematical reasoning, or at least it is if my speculative claim above is correct that visual experiences are required in order to carry out a priori mathematical deductions. Yet if one is led to reject the idea that mathematical reasoning is a priori, one might be led to reject the idea of a priori reasoning altogether, which would also collapse the improved Mary argument.)

I have argued that the improved Mary argument is not valid. Even if physicalism is true, and even if the higher-level physical facts are a priori deducible from the lower-level ones, this deduction is only possible from a God’s eye point of view, that is, a point of view that comes complete with the full range of sensory experiences. In particular, in order to deduce that seeing red is like this from the relevant lower-level physical facts, and to understand the conclusion of her deduction, Mary may need to have experienced red, identified as such. Yet this does not show that the property of having a visual sensation of red is non-physical, since it could still be entirely determined by fundamental physical properties.

Of course, the improved argument may not be the best Mary argument, for one could always add as a premise that Mary in her achromatic environment has all the conceptual resources or background knowledge necessary for making the deduction, given that experience is physical. However, I have suggested that this premise is open to question. Mary may have perfect reasoning powers; but without the relevant concepts, the acquisition of which, as I have argued, seems to require experience, she will not be able to deduce from the lower-level physical facts that seeing red is like this (where she understands what this refers to). In any event, that understanding the relevant concepts in Mary’s a priori deduction requires colour experience is a possibility which defenders of the Mary argument have yet to rule out.

V. DISTINGUISHING THIS RESPONSE FROM THE ‘OLD FACT, NEW UNDERSTANDING’ RESPONSE

In presenting this response to the Mary argument, it may seem that I am presenting merely a variation of the increasingly popular view that what Mary learns upon leaving her environment is just a different way of understanding certain physical facts which she already knew. However, the two
responses are significantly different. The old fact, new understanding response denies that experiential facts are a priori deducible from physical facts, and then goes on to argue that even though they are not a priori deducible, the experience of seeing red may be necessarily connected with certain physical properties – in fact, may even be identical with those physical properties. My deductions require experience response, however, accepts that the experiential facts are deducible a priori from physical facts, and then goes on to argue that certain a priori deductions require experience: in particular, the deduction Mary is expected to make requires colour experience.

A further distinction between the two arguments is although the old fact, new understanding response to the Mary argument holds that the experience of seeing red is identical with certain neural processes, and that when Mary sees red for the first time she comes to understand these neural processes in a different way, the response proffered here takes the experience of seeing red to be not identical with, but determined by, certain physical processes. This being so, the fact that seeing red is like this should be understood as a fact distinct from the facts Mary knows about the brain.

If Mary were merely coming to understand old information in a new way, then the proposition she expresses before experiencing red, when she says ‘Seeing red is like undergoing neural process n’, should in a sense be the same as the one she expresses when she says ‘Seeing red is like this’ (where ‘this’ refers to the type of experience she has upon seeing her first ripe tomato). And the sense in which they should be the same is that they should have the same truth-value in any possible world. But a physicalist, in particular a non-reductive physicalist, need not hold that the proposition expressed before her release by ‘Seeing red is like undergoing neural process n’ has the same truth-value in any possible world as the proposition expressed after her release by ‘Seeing red is like this’. Indeed, a non-reductive physicalist can deny that these two propositions have the same truth-value even in our world. While for the non-reductive physicalist it is true that every time neural process n occurs, an experience of red occurs, this does not mean that the non-reductive physicalist holds that seeing red is like undergoing neural process n, since seeing red is experiential, while neural process n is not. Typically, for the non-reductive physicalist, seeing red is a higher-level experiential feature of something that is not itself experiential. According to the non-reductive physicalist, then, the proposition that seeing red is like this is true, while the proposition that seeing red is like being in brain state b is false. Yet because the experiential aspect of seeing red is thought of as being determined by the fundamental physical properties, the fact that seeing red is like this is a physical fact.
VI. WHAT HAS BEEN SHOWN AND WHAT REMAINS

I have argued that it is at least possible that there are genuinely new physical facts about consciousness which Mary cannot know in an achromatic environment, even if she has perfect powers of deduction and all the higher-level physical facts are deducible from the bottom-level physical facts. Learning all the broadly physical facts may not be possible in an achromatic environment. The higher-level physical facts may not in general be \textit{a priori} deducible from the lower-level ones; or it may be that some of the simpler higher-level physical facts are \textit{a priori} deducible, but the complex ones are not; or, as I have argued, it may be that all of them are, but only for an individual who has already had a full range of experiences. This does not imply that ‘what it is like to see red’ could never be fully accounted for by some future science, but only that if this were to come to pass, the science would be one which cannot be fully understood in an achromatic environment. The science which gives a complete understanding of what it is like to see red is not a science Mary in her achromatic environment can learn.

Of course, a critique of the Mary argument is not in itself an argument for physicalism, for there still remains the question of how experience and its neural dependence base are related, the ‘hard problem’: why is it that once you have set the neural base, or whatever you think the relevant dependence base is, you have experience? I have certainly not answered this question, nor have I addressed any of the distinct arguments against physicalism. For example, I have not addressed anti-physicalist arguments intended to show that even for those of us outside black and white environments, facts about colour experience are not \textit{a priori} deducible from fundamental physical facts. Nevertheless, I hope to have shown how a physicalist can avoid the Mary argument, and perhaps even to have convinced some that this solution is obvious.

\textit{City University of New York}