

Phenomenal Knowledge Without Experience

Torin Alter

talter@ua.edu

[Forthcoming in *The Case for Qualia*, edited by E. Wright, MIT Press]

Abstract: Phenomenal knowledge usually comes from experience. But it need not. For example, one could know what it's like to see red without seeing red—indeed, without having any color experiences. Daniel Dennett (2007) and Pete Mandik (forthcoming) argue that this and related considerations undermine the knowledge argument against physicalism. If they are right, then this is not only a problem for anti-physicalists. Their argument threatens to undermine any version of phenomenal realism—the view that there are phenomenal properties, or qualia, that are not conceptually reducible to physical or functional properties. I will argue that this threat is illusory. Explaining why will clarify what is and is not at issue in discussions of the knowledge argument and phenomenal realism. This will strengthen the case for physically and functionally irreducible qualia.

Phenomenal knowledge usually comes from experience. For example, I know what it's like to see red because I have done so. Does knowing what it's like to have an experience with phenomenal character X require having an experience with X? No. A famous counterexample is Hume's missing shade of blue, in which one can extrapolate from phenomenally similar experiences (*A Treatise of Human Nature*, Bk. I, Pt. I, Sec. I).¹ One might think that a weaker version of the no-phenomenal-knowledge-without-experience thesis remains tenable, e.g., that knowing what it's like to see in color requires having or having had color experiences. But this thesis also seems doubtful. Peter Unger (1966) devised plausible counterexamples over four decades ago, and since then others (e.g., Lewis 1988, Alter 1998, Stoljar 2005) have done the same. One could have phenomenal knowledge of color experiences without having such experiences. Indeed, one could have such knowledge without having experiences that are remotely like color experiences.

What is the significance of this observation for contemporary debates about consciousness and physicalism? Daniel Dennett (2007) and Pete Mandik (forthcoming) suggest that it undermines the knowledge argument against physicalism.² That is because they take the claim that someone who has never seen in color could not know what it's like to see in color to be the basis of the knowledge argument's main epistemic premise: the premise that (roughly put)

no amount of physical knowledge is sufficient for phenomenal knowledge of color experiences. If they are right, then this is not only a problem for anti-physicalists. Many physicalists (e.g., Loar 1990/97, Papineau 2002, 2007) accept the knowledge argument's epistemic premise. Dennett's and Mandik's arguments threaten all versions of what David Chalmers (2003) calls *phenomenal realism*, the view that there are phenomenal properties, or qualia, that are not conceptually reducible to physical or functional properties.³

I will argue that this threat is illusory. Explaining why will clarify what is and is not at issue in discussions of the knowledge argument and phenomenal realism. The net result will be to strengthen the case for physically inexplicable qualia.^{4,5}

1. The knowledge argument and the no-experience-necessary response

The classic version of the knowledge argument is due to Frank Jackson (1982, 1986). It begins with his famous case of Mary, the super-scientist who is raised in a black-and-white room and has no color experiences. By watching lectures on black-and-white television, she learns all the physical information about seeing in color. This includes "everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles" (Jackson 1986:

291). Then she leaves the room and sees a red rose for the first time. Does she learn any truths? That is, does she gain any information? Intuitively, she does: she learns what it's like to see red.

Jackson takes this intuition to establish the knowledge argument's main epistemic premise, *the non-deducibility claim*: there are phenomenal truths that cannot be a priori deduced from the complete physical truth. He reasons that, if all phenomenal truths were a priori deducible from the complete physical truth, then Mary would learn no truths when she leaves the room; and so, because she does learn truths when she leaves, the non-deducibility claim is true. The knowledge argument's other main premise is the inference from the non-deducibility claim to the metaphysical conclusion that there are non-physical phenomenal truths, truths that are not metaphysically necessitated by the complete physical truth. From that conclusion, most agree, physicalism's falsity follows.⁶

Many phenomenal realists (e.g., Loar 1990/97, Papineau 2002, 2007) reject the knowledge argument's inference from its epistemic premise to its metaphysical conclusion. These philosophers often argue that, when Mary leaves the room, she merely comes to represent information she already knew in new ways. But all phenomenal realists accept both the non-deducibility claim and the claim that Mary's new experiences provide her with information.⁷

Let us turn to the idea that seeing in color is required for knowing what it's like to see in color. Call that *the experience requirement* on phenomenal knowledge of color experiences.⁸ Dennett and Mandik think the non-deducibility claim depends on the experience requirement. But this is not obvious. The non-deducibility claim does not seem to entail the experience requirement. Moreover, Jackson's argument for non-deducibility does not appear to invoke the requirement. His argument is that non-deducibility follows from the premise that Mary learns truths when she leaves the room despite her already knowing the complete physical truth. That reasoning does not appear to assume that the only way to know what it's like to see in color is to have color experiences. The intuition has the form "one cannot learn Q in way W", not "the only way to learn Q is in way W".

Why, then, do Dennett and Mandik think the non-deducibility claim depends on the experience requirement? Their reasoning could be summarized as follows:

The no-experience-necessary response. The claim that Mary makes epistemic progress upon release would make perfect sense if having color experiences were required for knowing what it's like to have them. But if the experience requirement fails—if it is possible to know what it's like to

see in color without having color experiences—then why couldn't Mary put herself in a state that allows her to figure out what it's like to see in color? If there is no logical bar to obtaining this phenomenal knowledge without seeing colors, then there is no reason why Mary could not obtain that knowledge by exploiting her comprehensive physical knowledge.

I will discuss three versions of this response: Dennett's, Mandik's, and another I will construct. If my arguments are sound, then the no-experience-necessary response is fundamentally misguided and no version can succeed.

2. Dennett's argument

2.1 RoboMary

Dennett develops his argument by varying the Mary case in ways intended to weaken the intuition that she gains knowledge when she leaves the room. The variation he emphasizes most is his Locked RoboMary case. RoboMary is a robot with the same physical knowledge as pre-release Mary. In one version of the case, she is equipped with a color-vision system. In another, she is "locked": her color-vision system—"the array of registers that transiently hold the codes for each pixel in Mary's visual field, whether seen or imagined" (Dennett 2007: 28)—is restricted to gray-scale values. According to Dennett, being locked "doesn't

faze her for a minute....” (Dennett 2007: 28). This is because “she builds a model of herself and from the outside, just as she would if she were building a model of some other being’s color vision, she figures out just how she would react in every possible color situation” (Dennett 2007: 28). He explains:

She obtains a ripe tomato and plunks it down in front of her black-and-white-cameras, obtaining some middling gray-scale values, which lead her into a variety of sequel states. ...She looks at the (gray-appearing) tomato and reacts however she does, resulting in, say, thousands of temporary settings of her cognitive machinery. Call that voluminous state of her total response to the locked gray-tomato-viewing *state A*. ...Then she compares state A with the state that her model of herself goes into... [namely] *state B*, the state she would have gone into if her color system hadn’t been locked. RoboMary notes all the differences between state A...and state B...and...makes all the necessary adjustments and *puts herself into state B*. State B is, by definition, *not* an illicit state of color experience; it is the state that such an illicit state of color experience normally causes (in a being just exactly like her). But now she can know just what it’s like for her to see a red tomato, because she has managed to put herself into just such a dispositional state.... (Dennett 2007: 28).

There is no reason why Mary, while in the room, could not do what Locked RoboMary does. Recognizing this, Dennett's contends, should weaken our faith in the intuitive judgment that Mary learns something when she leaves the black-and-white room.

2.2 The problem with RoboMary

Dennett's case threatens the non-deducibility claim only if Locked RoboMary a priori deduces the phenomenology of seeing red from the physical truth. But she does not do that. Rather, she puts herself in a dispositional state that constitutes possessing such knowledge. This is little better than if she had simply unlocked her color-vision system. She uses her physical knowledge to produce the desired effect, but the way she uses that knowledge is not (or not only) a priori deduction. Therefore, her achievement does not threaten the non-deducibility claim.⁹

Sven Walter and I raised the foregoing objection to Dennett while his article was in draft. In response he writes, "I just don't see that this is what matters. ...this objection presupposes an improbable and extravagant distinction between (pure?) deduction and other varieties of knowledgeable self-enlightenment" (Dennett 2007: 29). I take Dennett to mean that the distinction

between a priori deducibility and other sorts of infer-ability is (a) deeply problematic or (b) not relevant to whether physicalism stands or falls. Let me address these charges in turn.

Regarding (a): the notion of a priori deducibility is a straightforward application of the notion of a priori knowledge to reasoning (the latter notion does raise non-trivial issues (Baehr 2006), but none seem particularly relevant here). Let me illustrate the basic idea with a simple example. Contrast two cases of coming to know that the sum of a trapezoid's angles is 360 degrees:

Case 1. You figure out the sum by constructing a proof from Euclid's axioms.

Case 2. A future neuroscientist time travels back to the present and describes a brain state characteristic of someone who knows the sum. She also gives you a device that can be used to put you in that state and explains that the device works only if you contemplate Euclid's axioms for a few seconds. You contemplate the axioms and use the device. It works.

In case 1 you a priori deduce geometrical information from Euclid's axioms. In case 2 you do not, even though the axioms figure essentially in the process you use to acquire that same information.

To see how the same distinction applies to phenomenal knowledge, consider Hume's missing shade of blue. There are different ways of moving from phenomenal knowledge of the surrounding shades to phenomenal knowledge of the missing shade, and only one qualifies as a priori deduction: deducing the missing-shade phenomenology by combining phenomenal information about the surrounding shades, without relying on other phenomenal information. Such an a priori deduction seems possible. But it remains intuitively plausible that Mary cannot deduce what it's like to see red by combining the information she acquires before leaving the room. The same is true of RoboMary. Although she figures out what it's like to see red, her reasoning involves more than a priori deduction from physical information. In this respect, her reasoning is more like case 2 than case 1.

Regarding (b), concerning the relevance to physicalism's truth or falsity of the distinction between a priori deducibility and other sorts of infer-ability: the distinction is relevant because physicalism requires a metaphysically necessary connection from underlying microphysical truths to all truths and, arguably, a corresponding relation of a priori deducibility is needed to ground that necessary

connection (Chalmers 1996, 2004, forthcoming). To be sure, those claims raise hard issues. But those are not relevant here if Dennett wishes to do as he claims, which is to challenge the knowledge argument's epistemic premise. To challenge that premise just is to challenge a claim specifically about a priori deducibility (viz., the claim that there are phenomenal truths that cannot be a priori deduced from the complete physical truth). If Dennett wishes instead to challenge the inference from that premise to the metaphysical conclusion that there are non-physical phenomenal truths (truths that are not metaphysically necessitated by the complete physical truth), then that is another story entirely—and it is hard to see how the RoboMary case even begins to threaten that inference.

Dennett presents the following argument apparently in defense of his skeptical views about a priori deducibility:

Consider Rosemary, another of Mary's daughters, who is entirely normal and free to move around the colored world, and is otherwise her mother's equal in physical knowledge of color. Rosemary has a hard time imagining her mother's epistemic predicament. What must it be like, she wonders, not yet to know what it's like to see red? She is burdened, it seems, with *too much* knowledge.... This is, presumably, a psychological

impediment to her imagination, but not an epistemological lack. (Dennett 2007: 30)

I assume that Dennett does not mean that Rosemary's knowledge *must* interfere with her ability to imagine her mother's epistemic predicament, but only that this can happen. The latter claim is plausible. But how does it show that the distinction between a priori deducibility and other kinds of infer-ability is problematic or irrelevant to materialism? As far as I can tell, it does not.

Perhaps Dennett's idea is this: while RoboMary's inability to imagine in color prevents her from using one comparatively direct way of figuring out what it's like to see red, her inability need not place any limitation on her capacity for using reason to arrive at that knowledge; at best, any such limitation would be a contingent psychological impediment. But the problem is not that RoboMary's inability to imagine in color interferes with her ability to reason. The problem is rather that the method by which she obtains her phenomenal knowledge involves more than a priori deduction from physical information.¹⁰

Why does putting herself in state B enable RoboMary to know what it's like to see red? B is a dispositional and (let us assume) non-phenomenal state; there is nothing it's like to be in B. Nevertheless, B involves color phenomenology in that it contains the relevant phenomenal information. Therein

lies the problem for Dennett's argument. By putting herself in a state that involves color phenomenology, RoboMary cheats. Pre-release Mary should be no less puzzled about B than she is about seeing red. If she lacks phenomenal information about seeing red, then she lacks the phenomenal information that B contains. If there are open epistemic possibilities about the nature of phenomenal redness that she cannot eliminate, then there are open epistemic possibilities about the content of B that she cannot eliminate. RoboMary comes by her phenomenal knowledge of color experience not by a priori deduction from physical information but rather by putting herself in a non-phenomenal dispositional state that contains the relevant phenomenal information.

The problems with the Locked RoboMary case are symptomatic of a more general difficulty with Dennett's strategy. If the states Mary, RoboMary, or another Mary counterpart puts herself in—states that enable her to deduce what it's like to see red—involve color phenomenology, then she cheats: she does not a priori deduce the phenomenology from physical information. In that case, her achievement fails to threaten the non-deducibility claim. If, however, the states she puts herself in do not involve color phenomenology, then it is hard to see how they would enable her to deduce the phenomenology. I see no way of modifying the RoboMary case to evade this dilemma.

3. Mandik's argument

Mandik's version of the no-experience-necessary response consists mostly in showing how empirical considerations count against the thesis that knowing what it's like to see in color requires having color experiences. For example, he discusses a well-grounded prediction, based on a neural model of chromatic information processing, due to Paul Churchland (2005):

...the differential fatiguing and recovery of opponent processing cells gives rise to afterimages with subjective hues and saturations that would never be seen on the reflective surfaces of objects. Such "chimerical colors" include shades of yellow exactly as dark as pitch-black and "hyperbolic orange, an orange that is more 'ostentatiously orange' than any (non-self-luminous) orange you have ever seen...." (Mandik forthcoming: 13)¹¹

Mandik uses Churchland's prediction to construct his own variation of the Mary case. His case involves two characters. One is Larry, who has seen "all the typical colors a normally sighted adult has seen" (Mandik forthcoming: 13) but does not know the theory that predicts chimerical colors. The other is Hyperbolic Mary, who knows all the physical information that Mary knows, including the theory that predicts chimerical colors, and who has, like Larry, seen plenty of

colors. Mandik notes that, upon experiencing chimerical colors for the first time, Larry will be more surprised than Hyperbolic Mary. He concludes:

If it is unreasonable to expect Larry to predict the possibility of hyperbolic orange, pitch-dark yellow, and the like, then it seems unreasonable to predict, on introspection and intuition alone, the impossibility of pre-experiential knowledge of what it is like to see red. (Mandik forthcoming: 14)

Mandik may be right about this. The other considerations he discusses may also count against the thesis that “pre-experiential knowledge of what it is like to see red” is impossible. But as I have noted, the knowledge argument depends on no such thesis. Thus, Mandik’s strategy cannot succeed. If he could show that Hyperbolic Mary can a priori deduce the phenomenology of chimerical colors from physical information, then the knowledge argument would be in trouble. But he provides no reason to suspect that such an a priori deduction is possible.

4. Deviant phenomenal knowledge and the indirect argument

4.1 More cases

Before turning to yet another version of the no-experience-necessary response, let me introduce some terminology and a few more cases. Phenomenal knowledge is *earned* if the experience requirement is satisfied. For example, since I have seen ripe tomatoes, my knowledge of what it's like to see red is earned, whereas RoboMary's phenomenal knowledge is not. To *access* phenomenal knowledge is to exercise closely related abilities, such as the ability to imagine, recognize, or remember relevant experiences. I access my phenomenal knowledge when I visualize a ripe tomato, stop at a traffic light, or have an episodic memory of seeing oxygenated blood. Phenomenal knowledge that is unearned, inaccessible, or both is *deviant*.

Now to the other cases. I begin with another case Dennett describes (which he attributes to Gabriel Love):

Swamp Mary: Just as standard Mary is about to be released from prison, still virginal regarding colors...a bolt of lightning rearranges her brain, putting it by Cosmic Coincidence into exactly the brain state she was just about to go into *after* first seeing a red rose. (Dennett 2007: 24)

As Dennett notes, the lightning bolt does not give Swamp Mary a hallucinatory experience of a red rose but rather "puts Swamp Mary's brain into the

dispositional state, the competence state, that an experience of a red rose would have put her brain into had such an experience (hallucinatory or not) occurred” (Dennett 2007: 24). She has the equivalent of RoboMary’s state B. In relevant respects, Swamp Mary’s epistemic state closely resembles the state I am in while I neither see nor imagine red. I know what it’s like to see red. Therefore, it seems reasonable to say the same of her. She has the same phenomenal knowledge that I have; the only difference is that hers, like RoboMary’s, is unearned.¹²

Further cases of unearned phenomenal knowledge are not hard to devise. Unger imagines that scientists construct a man who is “cell-part for cell-part, cell for cell, nervous structure for nervous structure identical to” a man who both knows what it’s like to see red and has come upon this knowledge the ordinary way (Unger 1966: 50). Others (e.g., Lewis 1988, Alter 1998, Stoljar 2005) imagine cases in which surgeons operate on a person who has never seen red, creating brain structures similar to those found in the brains of people who have seen red. In all such cases, it is plausible that the subject knows what it’s like to see red, despite never having done so.¹³

Regarding inaccessible phenomenal knowledge, consider:

Temporary Impairment. John is epistemically just like pre-release Mary, except he has seen one color: red. Then a brain injury temporarily robs

him of the abilities associated with this knowledge. He can no longer imagine red things, he would not recognize red objects as red if presented with them, and he has no memory of what it's like to see red. As the effects of his injury recede, he recovers these abilities. He does so without seeing red in the intervening time; his hospital room contains no red objects.

Temporary Impairment is modeled on cases of aphasia discussed by Noam Chomsky (1980, 1988). Aphasia is the loss of the ability to use and understand language. Do aphasia patients retain their knowledge of language despite losing the associated abilities? Chomsky argues that sometimes they do. Suppose Maria is raised as a monolingual Spanish speaker and becomes temporarily afflicted with aphasia. As the effects of her injury recede, she recovers the ability to speak Spanish, not Chinese or French. Why would this be unless her knowledge of Spanish was there all along? Our Temporary Impairment case lends itself to a similar analysis. John recovers the abilities associated with seeing red rather than green. Why would this be unless his phenomenal knowledge was there all along? Plausibly, John has inaccessible phenomenal knowledge.¹⁴

Combining Swamp Mary and Temporary Impairment gives us:

Temporarily Impaired Swamp Mary. Before Swamp Mary leaves the room, she injures her brain. As a result, she is temporarily unable to exercise the abilities associated with seeing red. Like John, she eventually recovers those abilities without seeing red in the intervening time.

If, as I have argued, Swamp Mary and John know what it's like to see red, then Temporarily Impaired Swamp Mary does too. Her phenomenal knowledge is both unearned and inaccessible.

4.2 Is the notion of deviant phenomenal knowledge coherent?

Before discussing the implications of these cases, let me discuss a couple of objections. Chomsky's argument, on which I model my argument for the possibility of inaccessible phenomenal knowledge, has been questioned.

Anthony Kenny (1984) argues that linguistic knowledge cannot be detached from abilities in the way Chomsky's argument requires; in Kenny's view, Chomsky's notion of linguistic-knowledge-without-ability is incoherent. One might say the same about the notion of inaccessible phenomenal knowledge. Similarly, one might question the coherence of the notion unearned phenomenal knowledge. And one might invoke a causal theory of knowledge to challenge the attribution

of such knowledge to Swamp Mary and the other deviants whose states are caused by cosmic accidents.

Here such doubts may be set aside, for three reasons. First, regarding the last point, one might argue that a causal theory of knowledge is not appropriate for phenomenal knowledge—that such knowledge “does not consist in a causal relationship to experience, but in another sort of relationship entirely” (Chalmers 1996: 193). Also, the objection may not apply to Unger-style cases, in which the relevant dispositional states are created by neurosurgery rather than cosmic accident. Second, the possibilities of unearned and inaccessible phenomenal knowledge can only help the no-experience-necessary response. Rejecting these notions would decrease the chances of developing a successful version of that response. Third, the doubts may be partly terminological. In his response to Kenny, Chomsky emphasizes that, while impaired, Maria retains something that explains why she later recovers the ability to speak Spanish rather than Chinese or French. It is the possibility of such epistemic states, and not whether they should be described as knowledge, that is most central to his position. Similar considerations apply to deviant phenomenal knowledge. John has states that explain why he recovers the ability to recognize experiences of seeing red, rather than green. Swamp Mary has states that are strikingly similar to states Mary retains after seeing red and returning to the black-and-white room. Here these

similarities, and not whether the states should be described as knowledge, are what matter most. I will continue to describe the states of our deviant subjects as knowledge, but essentially the same issues would arise if I refrained from doing so.

I have discussed the notion of inaccessible phenomenal knowledge elsewhere (Alter 2001). Laurence Nemirow criticizes it, writing “Under Alter’s analysis, knowing what it’s like to see red has been transformed from one thing into something quite different” (Nemirow 2007: 50). This is because phenomenal knowledge “is known for its immediacy” (Nemirow 2007: 50) and inaccessible phenomenal knowledge would lack such immediacy. Nemirow concludes that the notion of inaccessible phenomenal knowledge “is a weird development—one that a proponent of phenomenal information should not welcome” (Nemirow 2007: 50).

But even if phenomenal knowledge “is known for its immediacy,” it does not follow that it can exist only if accessible, let alone immediately accessible. Perhaps the notion of inaccessible phenomenal knowledge seems odd because it is rarely considered. For example, most discussions of the Mary case focus on what happens as soon as she leaves the black-and-white room and becomes immersed in the world of color. But when we consider alternative scenarios, such as our temporary impairment cases, the notion may seem less bizarre. If

phenomenal knowledge consists in (or includes) the possession of information, then presumably that information can be stored in the brain while not being accessed. In that case, why couldn't brain damage block access to that stored information? Far from a weird development, accepting the possibility of inaccessible phenomenal knowledge is a natural consequence of taking seriously the idea that phenomenal knowledge consists (at least in part) in the possession of information.

4.3 The indirect argument

In the cases discussed above, the knower acquires her deviant phenomenal knowledge by means other than a priori deduction. Therefore, these cases do not directly threaten the non-deducibility claim. However, they may seem to create an indirect threat. Consider the following argument:

1. The phenomenal information contained in the deviants' dispositional states is equivalent to that which Mary has after seeing a red rose and returning to the black-and-white room.
2. If the phenomenal information Mary has after seeing a red rose and returning to the black-and-white room is a priori deducible from physical

- information, then the phenomenal information she has while looking at the rose is a priori deducible from physical information.
3. The phenomenal information contained in the deviants' dispositional states is a priori deducible from physical information.
 4. Therefore, the phenomenal information Mary has while looking at the rose is a priori deducible from physical information.

Call that *the indirect argument*. If it is sound, then Mary acquires no information upon leaving the black-and-white room and the non-deducibility claim and phenomenal realism are false.

The argument is valid. Its first two premises are plausible; let us grant them and focus on premise 3, the claim that the deviants' phenomenal knowledge is physically explicable. Why believe this premise? One might argue as follows:

Argument for premise 3. Consider Temporarily Impaired Swamp Mary. Her phenomenal knowledge exists in the form of dispositional states. These states are non-phenomenal; there is nothing it's like to be in them. They have never been triggered and, during her impairment, they cannot be triggered. William Robinson (1996) suggests that the reason phenomenal

information is not physically explicable is that, unlike physical information, it lacks structural expression. This may seem plausible when one reflects on one's own experience of seeing red. One may wonder, "How could *this* be captured in merely structural terms?" But why think that Temporarily Impaired Swamp Mary's unearned, inaccessible phenomenal knowledge includes information that lacks structural expression? The way the lightning bolt endows her with phenomenal knowledge is, after all, by reconfiguring her brain. How could that process add more than structure? There is thus no reason to think that she has knowledge that eludes physical explication. Likewise for the other cases of deviant phenomenal knowledge.

This reasoning may seem cogent. Further, the objection I brought against Dennett's argument does not apply to the indirect argument. That objection was that the method by which RoboMary acquires her phenomenal knowledge of color experience is not a priori deduction. But in the indirect argument, there is no claim that the deviant characters do any a priori deduction. Rather, the claim is that from within the black-and-white room Mary can figure out everything there is to know about their dispositional states. It is Mary, not the deviants, who is said to do the a priori deduction. Finally, unlike Mandik's argument, the

indirect argument challenges the non-deducibility claim, rather than the thesis that knowing what it's like to see in color requires having color experiences.

4.4 The problem with the indirect argument

But the indirect argument fails. The problem concerns premise 3, the claim that the deviants' phenomenal knowledge is physically explicable. Suppose that, before leaving the room, Mary uses her black-and-white television apparatus to study Temporarily Impaired Swamp Mary. Mary asks herself, "What is the content of Temporarily Impaired Swamp Mary's phenomenal knowledge of seeing red?" Intuitively, Mary is in no position to answer that question. At least, she is no better positioned to answer that question than she is to answer the question, "What is it like to see red?" There are many epistemic possibilities that she cannot eliminate. As far as she knows, Temporarily Impaired Swamp Mary's phenomenal knowledge may concern phenomenal greenness rather than phenomenal redness. Mary will be equally perplexed about the content of the other deviants' phenomenal knowledge. Thus, on reflection, the reasons for doubting that ordinary phenomenal knowledge is physically explicable apply equally to deviant phenomenal knowledge.

What, then, of the argument for premise 3? It makes two errors. First, it assumes that brain structures cannot carry physically inexplicable information.

But even non-physicalists typically accept that phenomenal properties naturally supervene (supervene as a matter of contingent, natural law) on physical properties such as neural properties (Chalmers 1996). Given this natural supervenience thesis, there is no clear reason why restructuring a brain could not endow it with physically inexplicable phenomenal information.

Second, the argument for premise 3 assumes that the fact that the deviants' phenomenal knowledge exists in a dispositional form indicates that their knowledge is physically explicable. This is questionable. My earned, accessible phenomenal knowledge exists in a dispositional form while I am neither seeing nor imagining red. Yet my knowledge is no more a priori deducible from physical information than is the phenomenal knowledge I had an hour ago while looking at a red tomato. Why shouldn't we say the same about the deviants' phenomenal knowledge?

5. Phenomenal concepts

5.1 The objection from conceptual parity

I have argued that RoboMary and the other deviants know what it's like to see in color even if pre-release Mary does not. One might raise the following objection to my position:

Objection from conceptual parity. If pre-release Mary does not know what it's like to see in color, then this is because she lacks the relevant phenomenal color concepts, namely, concepts that characterize phenomenal qualities as the qualities they are. But the deviants lack such concepts too. Because they have not had color experiences, they lack direct cognitive access to phenomenal color qualities, just as Mary does. Conceptually, pre-release Mary and the deviants are on a par. So, if pre-release Mary does not know what it's like to see in color, then neither do the deviants.¹⁵

But, arguably, the deviants do have phenomenal color concepts that pre-release Mary lacks. That they possess phenomenal color concepts follows from an assumption I defended in section 4 above: the assumption that the deviants know what it's like to see red. This view may also be confirmed by observing that pre-release Mary should be no less puzzled about the nature of Temporarily Impaired Swamp Mary's phenomenal red concept than she is about what it's like to see red. As far as pre-release Mary knows, Temporarily Impaired Swamp Mary's concept may pick out phenomenal greenness rather than phenomenal redness. The same point applies the other deviants.

5.2 Phenomenal concepts: relational and pure, standing and direct

Chalmers (2003) draws two distinctions that help to explicate the idea that the deviants have phenomenal concepts that pre-release Mary lacks. One is the distinction between pure and relational phenomenal concepts. The other is the distinction between direct and standing phenomenal concepts. I will explain these distinctions and then apply them to Mary and the deviants.

A pure phenomenal concept characterizes a phenomenal quality in terms of its intrinsic, phenomenal nature, i.e., as the quality it is. A relational phenomenal concept characterizes a phenomenal quality in terms of a relation the quality bears to other things. For example, one might have a community relational concept, *red_c*: a concept that characterizes phenomenal redness as the phenomenal quality typically caused in normal subjects in one's community by their seeing paradigmatic red things. Or one might have a demonstrative relational concept, *E*, which picks out whatever phenomenal quality one is internally ostending ("*this* quality"). Pure phenomenal concepts are distinct from relational phenomenal concepts, even when they pick out one and the same phenomenal quality. To see this, consider a belief Mary forms when she first leaves the room and sees a red rose: the belief that seeing red things typically causes those in her community to have experiences with such-and-such phenomenal quality. Her "such-and-such phenomenal quality" concept, *R*, is a

pure phenomenal concept of phenomenal redness. Her belief also involves *red_c*, which also picks out phenomenal redness. Her belief is cognitively significant, and this indicates that *R* and *red_c* are distinct concepts (Chalmers 2003: 225). Similar reasoning shows that *R* is distinct from any relational concept. *E*, *red_c*, and *R* pick out the same phenomenal quality, but only *R* is pure.

A *direct* phenomenal concept is “partly *constituted* by an underlying phenomenal quality” (Chalmers 2003: 235). Typically, one forms such a concept by “attend[ing] to the quality of an experience, and form[ing] a concept wholly based on the attention to the quality, ‘taking up’ the quality into the concept” (Chalmers 2003: 235). Chalmers (2003: 240) remarks that, “The lifetime of a direct phenomenal concept is limited to the lifetime of the experience...that constitutes it.” Presently, I see nothing red but I know what it’s like to see red. Therefore, my knowledge does not involve a direct phenomenal concept. Instead, it involves what Chalmers calls a *standing* phenomenal concept. Standing phenomenal concepts “persist in a way that direct phenomenal concepts do not” (Chalmers 2003: 239). But standing phenomenal concepts may be pure: they may characterize a phenomenal quality not relationally but in terms of the quality’s intrinsic, phenomenal nature. How is their content determined? Chalmers (2003: 239) suggests that “their content is determined by some combination of (1) non-sensory phenomenal states of a cognitive sort, which bear a relevant relation to

the original phenomenal quality in question—e.g. a faint Humean phenomenal ‘idea’ that is relevantly related to the original ‘impression’; (2) dispositions to have such states; and (3) dispositions to recognize instances of the phenomenal quality in question.”

Let us see how Chalmers’ distinctions apply to Mary and the deviants. Before Mary leaves the black-and-white room, she has no pure phenomenal color concepts. By contrast, the deviants do have such concepts. Their pure concepts differ from those I form while looking at a red tomato: only mine are direct. But their standing concepts are pure. Consider Swamp Mary’s standing concept of phenomenal redness. Although she has never seen red, she has a disposition to recognize phenomenal redness as the quality it is; this disposition determines her concept’s content. The content of her standing concept should be no less puzzling to Mary than the content of my direct concept. As far as pre-release Mary knows, either might pick out phenomenal greenness rather than phenomenal redness.

Unlike Swamp Mary, John and Temporarily Impaired Swamp Mary would, while impaired, fail to recognize phenomenal redness as the quality it is. Does this imply that they lack the relevant recognitional disposition? No. As Mark Johnston (1992) points out, a fragile glass cup encased in packing material may retain its fragility—its disposition to shatter when dropped—even if the

packing material prevents it from shattering when dropped. In such cases, a disposition is said to be *masked*. Our temporarily impaired deviants have the same recognitional disposition as Swamp Mary, even though theirs are masked. Arguably, they have pure standing phenomenal redness concepts, even though they are temporarily unable to employ those concepts in their thinking. By contrast, pre-release Mary has only relational phenomenal color concepts. On reflection, the difference between Mary's epistemic situation and that of the deviants is striking.

6. Conclusion

Dennett writes:

Another unargued intuition exploited by the Mary intuition pump...is the idea that the "phenomenality" or "intrinsic phenomenal character"...cannot be constructed or derived out of lesser ingredients. Only actual experience (of color, for instance) can lead to the knowledge of what that experience is like. (Dennett 2007: 22)

Here Dennett misconstrues the relationship between the Mary case and the claim that phenomenality "cannot be constructed or derived out of lesser ingredients":

the case provides support for the claim, rather than the reverse. But put that aside; I wish to draw attention to a different point. The passage illustrates the most basic assumption behind the no-experience-necessary response: the assumption that the knowledge argument depends on theses such as “Only actual experience (of color, for instance) can lead to the knowledge of what that experience is like.” As I have argued, that assumption is false. Associating the knowledge argument with such implausible theses makes the case for physically inexplicable qualia appear weaker than it is.

Acknowledgements: For helpful suggestions, I thank David Chalmers, Amy Kind, Pete Mandik, Jennifer McKittrick, Yujin Nagasawa, and Stuart Rachels. This paper benefited from my participation in John Heil’s 2006 NEH Seminar, *Mind and Metaphysics*.

Notes

1. Another counterexample: one might know about a complex phenomenal quality by knowing about its parts (Alter 2002: 52-53, Tye 1995: 227, n. 2). For example, someone who has seen red, white, and blue separately but never together may be able to deduce what it’s like to see an American flag from a detailed description.

2. Derek Ball (n.d.) gives a related argument. He argues that (i) the knowledge argument requires assuming the existence of phenomenal concepts that satisfy something akin to what I call *the experience requirement* (see section 1) and (ii) there are no such concepts. My main criticisms of the no-experience-necessary response apply to (i), *mutatis mutandis*.
3. For brevity, when referring to physical or functional phenomena, I will henceforth leave off 'or functional'.
4. I use 'physically explicable' to refer to information that is a priori deducible from the complete physical truth. On this usage, phenomenal realists who are physicalists accept that there is physically inexplicable information.
5. Explicit commitment to the thesis that knowing what it's like to see in color requires having color experiences may be rarer than Unger, Dennett, and others suggest. Unger (1966: 49, n. 2) cites three sources: Price 1962: 52, 53, 57-58; Russell 1948: 499; and Ryle 1949: 317. But on none of those pages is the thesis asserted or implied. Dennett quotes passages from Tye 1995 (167, 169) and Lycan 2003 (389) that seem to suggest the thesis. But later in his article (393), Lycan suggests that the thesis may be false. And in Tye 2000 (27), Tye implies that the thesis is false. Locke, Hume, and other early modern empiricists assert principles that seem to suggest the thesis. But as Unger (1966: 49) notes, these principles may be

intended merely as contingent truths about human beings rather than necessary or conceptual truths.

6. This formulation of the knowledge argument is imprecise. For example, the deduction base mentioned in the epistemic premise should include all indexical information, such as the truths expressed by “I am *s*” and “Now is *t*” where *s* and *t* are descriptions specifying unique subjects and times. For a precise formulation (and forceful defense) of the argument, see Chalmers 2004. Jackson (2003, 2007) now rejects the argument. I criticize his basis for so doing in Alter 2007.

7. Despite its popularity, there are good reasons to think the conjunction of phenomenal realism and physicalism is unstable. See Chalmers 1996, 2004, forthcoming.

8. I focus on color experiences because of their role in the Mary case. What I say applies to other kinds of experience, *mutatis mutandis*.

9. Michael Beaton (2005) develops a related objection to Dennett’s argument. But there are notable differences. In particular, Beaton assumes a version of the Lewis-Nemirow ability hypothesis (Lewis 1988, Nemirow 1990, 2007). I make no such assumption. Further, the ability hypothesis is incompatible with phenomenal realism.

10. Walter and I wrote (in correspondence with Dennett) that RoboMary *self-programs* herself into state B instead of a priori deducing the relevant information. In response Dennett (2007: 30) writes, “I didn’t describe RoboMary as ‘programming’ herself....” I regret the slip, but this terminological matter is incidental.

11. The final version of Mandik’s article is pending as I write, and so he might revise. But (in correspondence) he has assured me that attributing the no-experience-necessary response to his article is safe. I quote from the version on his website: <http://www.petemandik.com/philosophy/papers/nos.pdf>.

12. Jackson (1998: 77) makes a similar point: “Seeing red and feeling pain impact on us, leaving a memory trace which sustains our knowledge of what it is like to see red and feel pain on the many occasions where we are neither seeing red nor feeling pain. This is why it was always a mistake to say that someone could not know what seeing red and feeling pain [are] like unless they had actually experienced them: false ‘memory’ traces are enough.”

13. In Alter 1998 (54, n. 15), I use such a case to raise the issue of whether phenomenal knowledge can be innate. There I do not attempt to resolve this issue. I would now suggest that innate phenomenal knowledge is unproblematic for the same reasons that unearned phenomenal knowledge is unproblematic.

The idea that phenomenal knowledge could be produced by cosmic accident is not new. Unger (1966: 54) describes a case in which “the various parts of the duplicate person just (‘by chance’) came together.” And Lewis ((1988: 500) mentions the idea that phenomenal knowledge “could possibly be produced in you by magic.”

14. In Alter 2001, I argue that the possibility of inaccessible phenomenal knowledge creates a problem for the Lewis-Nemirow view, based on their ability hypothesis, on which all Mary gains when she leaves the room are abilities, as opposed to physically inexplicable phenomenal information (Lewis 1988, Nemirow 1990, 2007). The Temporary Impairment case indicates that she gains something that can survive the loss of ability. Lewis and Nemirow could respond by revising their view and arguing that this something else consists in *proto-abilities*—mental states that normally underlie the relevant abilities but do not carry physically inexplicable phenomenal information. I reject a similar response in Alter 2001, but I now think it is probably adequate.

15. The idea that Mary’s epistemic progress derives from her acquiring phenomenal color concepts suggests a natural response for the a priori physicalist (the physicalist who rejects the non-deducibility claim): although the phenomenal truth about color experiences can be a priori deduced from the complete physical truth, pre-release Mary cannot do the deduction because she

lacks the requisite phenomenal concepts. See Tye 2000, Hellie 2004, and Chalmers 2004. But Chalmers (2004) argues that she might not be able to deduce certain phenomenal truths even if she is given the relevant phenomenal concepts. For example, he writes, “once Mary has the relevant phenomenal concept, she will not automatically know whether or not *other* organisms (bats or Martians, say) are having experiences of the relevant sort, even given a complete physical description of them” (Chalmers 2004: 285).

References

- Alter, T. 2007. Does Representationalism Undermine the Knowledge Argument? In *Phenomenal Knowledge and Phenomenal Concepts: New Essays on Consciousness and Physicalism*, T. Alter and S. Walter (eds.). New York: Oxford University Press, pp. 65-76.
- Alter, T. 2002. On Two Alleged Conflicts between Divine Attributes. *Faith and Philosophy*, 19: 47-57.
- Alter, T. 2001. Know-how, Ability, and the Ability Hypothesis. *Theoria*, 67: 229-239.
- Alter, T. 1998. A Limited Defense of the Knowledge Argument. *Philosophical Studies*, 90: 35-56.

- Baehr, J. 2006. A Priori and A Posteriori. In *The Internet Encyclopedia of Philosophy*, J. Fieser and B. Dowden (eds.): <http://www.iep.utm.edu/a/apriori.htm>.
- Ball, D. n.d. Physicalism and Phenomenal Concepts. Unpublished.
- Beaton, M. 2005. What RoboDennett Still Doesn't Know. *Journal of Consciousness Studies*, 12: 3-25.
- Chalmers, D. J. Forthcoming. The Two-Dimensional Argument against Materialism. In his *The Character of Consciousness*. New York: Oxford University Press.
- Chalmers, D. J. 2004. Phenomenal Concepts and the Knowledge Argument. In *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, P. Ludlow, D. Stoljar, and Y. Nagasawa (eds.). Cambridge: MIT Press, pp. 269–298.
- Chalmers, D. J. 2003. The Content and Epistemology of Phenomenal Belief. In *Consciousness: New Philosophical Essays*, Q. Smith and A. Jokic (eds.). Oxford: Oxford University Press, pp. 220-272.
- Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chomsky, N. 1988. Language and Problems of Knowledge. *Synthesis Philosophica*, 5: 1-25. Rpt. in *The Philosophy of Language*, 4th Edition, A. Martinich, (ed.). New York: Oxford University Press, 2001, pp. 558-577.

- Chomsky, N. 1980. *Rules and Representations*. New York: Columbia University Press.
- Churchland, P. 2005. Chimerical Colors: Some Novel Predictions from Cognitive Neuroscience. *Philosophical Psychology*, 18: 527-60.
- Dennett, D. C. 2007. What RoboMary Knows. In Alter and Walter, 2007, pp. 15-31.
- Jackson, F. 2007. The Knowledge Argument, Diaphanousness, Representationalism. In Alter and Walter, 2007, pp. 52-64.
- Jackson, F. 2003. Mind and Illusion. In *Minds and Persons*, A. O'Hear, (ed.). Cambridge, UK: Cambridge University Press, pp. 251-271.
- Jackson, F. 1998. Postscript on Qualia. In his *Mind, Method, and Conditionals: Selected Essays*. New York: Routledge, 76-79.
- Jackson, F. 1986. What Mary Didn't Know. *The Journal of Philosophy*, 83: 291-295.
- Jackson, F. 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32: 127-136.
- Johnston, M. 1992. How to Speak of the Colors. *Philosophical Studies*, 68: 221-63.
- Kenny, A. 1984. *The Legacy of Wittgenstein*. Oxford: Basil Blackwell.
- Hellie, B. 2004. Inexpressible Truths and the Allure of the Knowledge Argument. In Ludlow, Nagasawa, and Stoljar, 2004, pp. 333-364.

- Lewis, D. 1988. What Experience Teaches. In *Proceedings of the Russellian Society*.
Sydney: University of Sydney. Rpt. in *Mind and Cognition: a Reader*, W.
Lycan, (ed.). Cambridge: Basil Blackwell, 1990, pp. 499-518.
- Loar, B. 1990/97. Phenomenal States. In *Philosophical Perspectives 4: Action Theory
and Philosophy of Mind*, J. Tomberlin (ed.), pp. 81-108. Atascadero, CA:
Ridgeview. Revised version in *The Nature of Consciousness: Philosophical
Debates*, N. Block, O. Flanagan, and G. Güzeldere (eds.). Cambridge, MA:
MIT Press, 597-616.
- Lycan, W. G. 2003. Perspectival Representation and the Knowledge Argument. In
Smith and Jokic, 2003, pp. 382-395.
- Mandik, P. Forthcoming. The Neurophilosophy of Subjectivity. In *Oxford
Handbook of Philosophy and Neuroscience*, J. Bickle (ed.). New York: Oxford
University Press.
- Nemirow, L. 2007. So *This* is What it's Like: a Defense of the Ability Hypothesis.
In Alter and Walter, 2007, pp. 32-51.
- Nemirow, L. 1990. Physicalism and the Cognitive Role of Acquaintance. In
Lycan, 1990, pp. 490-99.
- Papineau, D. 2007. Phenomenal and Perceptual Concepts. In Alter and Walter,
2007, pp. 111-45.

- Papineau, D. 2002. *Thinking About Consciousness*. New York: Oxford University Press.
- Price, H. H. 1962. *Thinking and Experience*. Cambridge, MA: Harvard University Press.
- Robinson, W. 1996. The Hardness of the Hard Problem. *Journal of Consciousness Studies*, 3: 14-25.
- Russell, B. 1948. *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson & Co.
- Stoljar, D. 2005. Physicalism and Phenomenal Concepts. *Mind and Language*, 20: 469-494.
- Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.
- Tye, M. 1995. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.
- Unger, P. 1966. On Experience and the Development of the Understanding. *American Philosophical Quarterly*, 3: 48-56.