

MICHAEL PELCZAR

ENLIGHTENING THE FULLY INFORMED

ABSTRACT. This paper develops a response to the knowledge argument against physicalism. The response is both austere, in that it does not concede the existence of non-physical information (much less non-physical facts), and natural, in that it acknowledges the alethic character of phenomenal knowledge and learning. I argue that such a response has all the advantages and none of the disadvantages of existing objections to the knowledge argument. Throughout, the goal is to develop a response that is polemically effective in addition to theoretically sound.

1. THE KNOWLEDGE ARGUMENT

The knowledge argument runs as follows:

1. Someone who has all the physical information concerning a kind of conscious experience that she has never had could learn something about that kind of experience, by having an experience of that kind.
2. Such learning would be impossible, if the only information concerning that kind of experience were physical information.
3. Therefore *not* all information concerning conscious experience is physical information.

For the purposes of this argument, “physical information” is standardly understood as information expressible in the language of the physical sciences, and that is how I understand it here. It is simply information as to what exists and happens, physically speaking.¹

The first premise of the argument receives support from hypothetical cases like Jackson’s case of the sequestered neuroscientist, Mary. The story is that Mary has lived all her life in an artificially maintained grayscale environment, where she

studies the physical nature of conscious sense experience by means of black and white books, video feeds, etc. We are to imagine that she learns all the physical facts relating to human visual experience in this way – what stimulates it, what goes on in the eyes, optic nerve, and brain of someone who has it, how it ties into motor behavior, etc. In short, we are to suppose that any truth relevant to color experience that can be expressed in the language of physical science is one that she knows. Still, it seems that she would *learn* something if she were to escape her prison to confront a brightly colored object, such as a ripe tomato. Intuitively, this would be an *enlightening* experience for her. That being so, there must be some *information* about color experience that was unknown to her before, notwithstanding her possession of all the physical information. There must then be some *non-physical* information about color experience that she acquires when she sees the tomato.²

There are two main ways to put pressure on this reasoning. One is to question whether someone who really knew all the physical details to the finest degree would remain ignorant of what it was like to experience red. This is not a very popular approach, although it has its adherents.³ It will not be this approach that I take in what follows. The other is to grant that Mary's first experience of color is enlightening, that it is a learning experience, and that it gives her new knowledge, but to deny that this circumstance poses any insurmountable problem for physicalism. This is the approach of most mainstream responses to the knowledge argument, and it is this general approach that I take here.

Ideally, a physicalist response to the knowledge argument that takes this approach should satisfy both of the following desiderata:

Truth Aptness: It should allow that knowing what it is like to have a red experience is a case of *truth apt* knowledge; i.e., knowledge of something that can be assessed for truth value (such as a proposition or item of information), or of something assessable for some equivalent truth-related status (such as a fact, which can be assessed as obtaining or failing to obtain).

Austerity: It should put us in a position to maintain that there is nothing of which Mary has truth apt knowledge only after experiencing red – no fact or proposition that she did not already know, no item of information that she did not already possess, etc.⁴

I say that an *ideal* physicalist response to the argument should satisfy both of these desiderata, because only a truth apt response conforms to the intuition that phenomenal knowledge and learning are of an essentially propositional, factual, or information-oriented character, and only a response that is austere avoids commitment to the claim that one could know all the facts about something without having all the information about it. Some physicalists will consider truth aptness an inappropriate desideratum, on the grounds that there is no compelling reason to regard phenomenal knowledge or learning as truth apt. Some physicalists will consider austerity an inappropriate desideratum, because they believe that conceding the existence of non-physical information leaves the ontological core of physicalism untouched. But an *ideal* response to the knowledge argument should do more than raise an hallelujah from the physicalist choir: it should also address the worries of those whose interests or intuitions genuinely run counter to physicalism. Only a response that satisfies the desiderata given above has any real hope of attaining this polemical objective, and it is a response that is ideal in this sense that I attempt to provide here.

The structure of the paper is as follows. I first consider two popular responses to the knowledge argument. One of these, the acquaintance objection, satisfies the desideratum of austerity, but not the desideratum of truth aptness. I discuss the acquaintance objection in section 2, and explain why its truth inaptness is a problem. In section 3 I consider the old fact/new guise response to the knowledge argument, which is truth apt but not austere, and explain why its lack of austerity is a problem. I then consider an alternative response to the knowledge argument that is both truth apt and austere. This is what I call the “old relatum/new relation” response, which I

develop in section 4 and defend against three anticipated objections in section 5.

2. ACQUAINTANCE

One of the more insightful discussions of the issues surrounding the knowledgement is Earl Conee's. Conee grants that "[w]hen Mary first sees red ripe tomatoes, she learns what it is like to see something red." But according to Conee, learning what it is like to see something red is simply a matter of gaining "acquaintance knowledge" of the *quality* of phenomenal red. Mary learns only in the sense that she becomes acquainted with a new phenomenal quality, and she becomes acquainted with this quality simply by having a conscious experience that has it. Since, according to Conee, it is possible to gain acquaintance knowledge without gaining any new information, there is no need to allow that Mary acquires knowledge of any new fact or piece of information. That being so, we can explain Mary's learning in a way that fully comports with physicalism.⁵

If it is correct, Conee's account yields an austere response to the knowledge argument. For, by this account, to possess knowledge of what a given kind of experience is like is simply to have had that kind of experience consciously, in much the same way as possessing carnal knowledge is simply a matter of having had sex. It is impossible to have carnal knowledge without having had sex, but that is just because it is impossible to have had sex without having had sex. Anyone who has sex gains carnal knowledge thereby, regardless of whether he acquires any information in the process. Likewise, given Conee's account, anyone who has a conscious experience of red gains knowledge of what it is like to experience red, regardless of whether she acquires any information thereby.

Critics of the acquaintance response object that there is *more* to learning what it is like to experience red than simply becoming acquainted with the quality of phenomenal redness, unless indeed becoming acquainted with this quality is something that itself involves the acquisition of new information.⁶

The problem, as these critics see it, is that Conee's account makes the analogy between carnal and phenomenal knowledge all *too* apt. For, while it is true that one may acquire carnal knowledge without acquiring any information, for that very reason the acquisition of carnal knowledge does not seem to count as a properly *enlightening* experience. By failing to allow that Mary's learning is of the sort that involves the cognition of some fact or proposition, or the acquisition of some information, the acquaintance response seems to these critics not to do full justice to the epistemically enlightening character of Mary's learning experience.

Why should we take this complaint seriously? What reasons are there to suppose that when one knows what it is like to see red, or, knows that **this** is what it is like to see red, one possesses some item of information?⁷ One reason is that this supposition does justice to the surface grammar of phenomenal knowledge attributions. When we say that Mary learns that **this** is what it is like to see red, we seem to be saying that there is some *truth* – namely, that *this is what it is like to see red* – that Mary learns. Against this, a proponent of Conee's approach might respond that surface grammar is not everything, and that the acquaintance objection's unfaithfulness to the surface grammar of phenomenal knowledge attributions is a small enough price to pay for a refutation of the knowledge argument. But while a physicalist may be sympathetic to this rejoinder, it is polemically weak.

In fact, the acquaintance theory's unfaithfulness to surface grammar gets it into real trouble. For, it seems that we must regard what Mary learns as truth apt, on pain of making nonsense of statements such as: "If **this** is what it is like to see red, then what it is like to see red is different from what it is like to see blue." Since this statement is not nonsense but indeed true, it seems that the phrase "**this** is what it is like to see red" must possess truth value, and therefore correspond to some fact, proposition, or item of information.⁸

I believe that there is a substantial grain of truth in Conee's idea that knowing what it is like to have a given kind of phenomenal experience is basically a matter of having had a

conscious experience of that kind. But the acquaintance account is open to objections that are not easily met, and that are sure to leave a dualist unconvinced that the knowledge argument has been refuted. In the next section I consider a different objection to the knowledge argument that has certain advantages over the acquaintance objection.

3. OLD FACT/NEW GUISE

The most popular physicalist response to the knowledge argument tries to accommodate the idea that Mary's learning experience is of a truth apt variety. This is the "old fact/new guise" response. According to this, there is a certain fact about human vision that can be expressed in the language of physics (by means of the sentence " Σ ", say), and can also be expressed by Mary's utterance of the sentence, "This is what it is like to see something red," upon seeing the tomato. Before seeing the tomato, Mary knows that Σ , but does not know what it is like to experience red. However, the old fact/new guise account continues, this no more entails a difference between the fact that Σ and the fact that **this** is what it is like to see red than the fact that someone can know that Mark Twain died in 1910 without knowing that Samuel Clemens died in 1910 entails a difference between the fact that Mark Twain died in 1910 and the fact that Samuel Clemens died in 1910. When Mary sees the tomato, she simply becomes reintroduced to a fact that she knew all along under a new "guise" or "mode of presentation," just as someone who knows that Twain died in 1910 can learn that Clemens died in 1910 by becoming reintroduced to a fact he knew all along, under a new guise. That being so, a physicalist may grant that the knowledge argument is sound – i.e., grant that not all of the information about conscious experience is physical information – without giving up the idea that all *facts* about conscious experience are physical facts.

The old fact/new guise response is truth apt, since, unlike the acquaintance response, it allows that Mary learns something of a broadly factual or truth assessable character. For it allows that she gains new information, where each item of information

corresponds to some fact, in the sense that for each item of information, there is some fact upon which that item of information depends for its status *as* information (rather than misinformation, or non-information). What the old fact/new guise theory maintains is that the information that Mary acquires when she leaves her prison corresponds only to a fact she already knew, by virtue of possessing some distinct, physical piece of information. If this is correct, then we can explain Mary's learning in a way that is compatible with physicalism: Mary acquires a new item of information, but since this new information corresponds only to a fact that she already knew by virtue of possessing some distinct item of information, her acquisition of this new information does not require her to learn any new *fact*, in which case we are free to grant that Mary learns something without committing ourselves to the existence of non-physical facts.

Different proponents of the old fact/new guise response have different ways of describing what it is that pre-release Mary does not know, but the general form of this residual unknown is the same in every version of the response. In every version, what pre-release Mary does not know has the logical structure of a ⟨fact, something-or-other⟩ pair. Different versions of the old fact/new guise response differ by reference to what they specify as the relevant something-or-other. For Horgan, the objects of knowledge are ⟨fact, information⟩ pairs, where a single fact corresponds to as many items of information as there are ways of expressing the fact. Thus, while every fact figures in some ⟨fact, information⟩ pair that pre-release Mary knows, there are ⟨fact, information⟩ *pairs* that she does not know prior to her release, since prior to her release the only pairs she knows are ones whose information terms are explicitly physical (corresponding to explicitly physical statements of fact).⁹ For Loar, the something-or-other elements are concepts, so that what one knows in having truth apt knowledge has the structure of a ⟨fact, concept⟩ pair. As for Horgan, for Loar there is no fact that fails to figure in some object of knowledge that Mary knows prior to her release. What pre-release Mary does not know, according to Loar, are certain objects of knowledge

whose conceptual components consist of “phenomenal recognitional” concepts as opposed to the “physical-functional” concepts that Mary possesses even before having her first experiences of color.¹⁰ For Lycan, truth apt knowledge is knowledge of a ⟨fact, mode-of-presentation⟩ pair, where, as according to all old fact/new guise theories, there is no fact that does not figure in *some* pair that pre-release Mary knows. What pre-release Mary does *not* know are certain pairs containing introspective as opposed to third-person modes of presentation.¹¹

Opponents of the old fact/new guise response object to it as follows: “The information that Σ must differ from the information that **this** is what it is like to experience red, since the latter is supposed (by the old fact/new guise theory) to be what Mary gains, despite already possessing the former. But if these are different items of information, then it is a fact relevant to color experience that they *are* factually equivalent. Therefore, if Mary already knew all the relevant facts, she knew the fact that the information that Σ is factually equivalent to the information that **this** is what it is like to see red. But it seems clear she did not know this fact, prior to her release. Therefore, the old fact/new guise response fails.”¹²

There is a reply to this objection. The old fact/new guise theorist can argue that pre-release Mary *did* already know the fact that the information that Σ is factually equivalent to the information that **this** is what it is like to experience red, but only by virtue of possessing some purely physical item of information. What information was this? What the old fact/new guise theory must say, apparently, is that she knew this fact by virtue of possessing the information that the information that Σ is factually equivalent to the information that Σ . Here one might object that this information cannot be factually equivalent to the information that the information that Σ is factually equivalent to the information that **this** is what it is like to experience red, since the former, unlike the latter, is a mere tautology. However, this objection relies on the assumption that no tautology can correspond to the same fact as any non-tautology, which is an assumption that the old fact/new guise

theorist will surely reject. According to him, tautology is not a feature of facts, but of the statements, propositions, or information by which certain facts may be represented or conveyed. As long as it is *necessary* that the information that Σ is factually equivalent to the information that **this** is what it is like to experience red – and a physicalist is likely to maintain that it is – there need be no difficulty in regarding the fact of this factual equivalence as identical to the equally necessary fact that the information that Σ is factually equivalent to itself.

The old fact/new guise response loses some of its intuitive appeal when pushed to this level, and this belies a deeper problem with the account. This is its implausible presupposition that someone could know *all* the relevant facts about something – including all the relevant facts about all the relevant information – and yet lack some relevant item of information. Simply pointing to canonical cases like that of Twain a.k.a. Clemens will not do, since these are not cases in which the person lacking information is supposed to possess knowledge of *all* the relevant facts. Someone who knew *all* the relevant facts about Twain would know all the relevant facts about the sentences “Mark Twain died in 1910” and “Samuel Clemens died in 1910,” all the facts about all other representations of facts having to do with Twain, all the facts about the facts such sentences state, all the facts about all the relevant concepts, modes of presentation, and items of information, and all the facts as to the connections among these. Only if someone who knew all *these* facts could lack the information that Clemens died in 1910 would we have an analogy to support the old fact/new guise account of Mary’s learning.

If, prior to her release, Mary knows all the facts relevant to color experience, then she knows the fact that my statement “This is what it is like to have a red experience” states, as well the fact that my statement states that fact. Here, it is not as if I were peering into cupped hands and saying, “This one is square!” not allowing her to see what, if anything, I held. Rather, according to the old fact/new guise theory, it is as if she knew exactly what fact I was reporting in saying, “This one is square!” (e.g., if what I am holding is a piece of beach glass, a

certain fact having to do with this), but still somehow had less information than what I have, in having the information that (as I might put it) this one is square.¹³

The hope was that by positing a richer field for knowledge, populated not by bare facts but more finely individuated fact-involving cognitive objects, we could grant that pre-release Mary knew all the facts pertaining to color experience without being forced to say that she knew *everything* pertaining to color experience, thus leaving room for her to learn. But when we draw out the implications of comprehensive factual knowledge of the sort pre-release Mary is supposed to possess, we find that it seems to preclude any lack on her part of information pertinent to color experience. The old fact/new guise account fails to constitute a convincing reply to the knowledge argument, because it fails to explain how Mary could know *every* relevant fact without having *all* relevant information.

4. OLD RELATUM/NEW RELATION

I have argued that an ideal physicalist response to the knowledge argument would allow that Mary gains truth apt knowledge upon having her first experience of red, without allowing that there is anything of which she did not already have truth apt knowledge. The question we must ask now is whether an ideal physicalist response to the knowledge argument is at all possible. To see where the possibility for such a response might lie, it will help to chart out the responses we have already explored, together with any related but unexplored possibilities (see Figure 1).

An ideal response to the knowledge argument must locate Mary's learning somewhere under the truth apt branch of the diagram; i.e., as involving the acquisition of knowledge of something that can be assessed for truth value. The primary shortcoming of the acquaintance response is that it construes her learning as truth *inapt*. The truth apt branch of the diagram divides into two further branches. First there is the branch that represents *informative* truth apt knowledge acquisition. This is the kind of truth apt learning that entails an addition to the

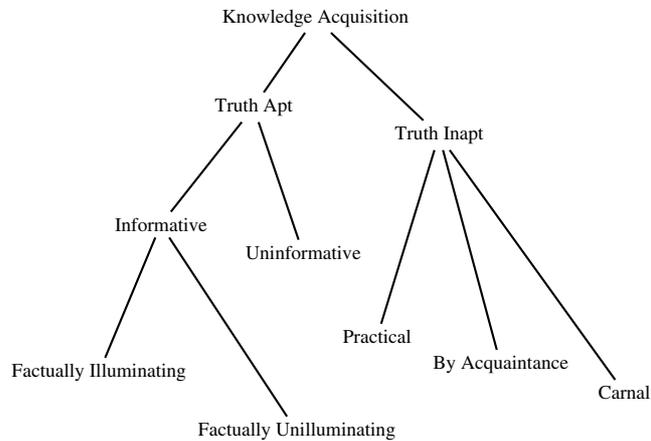


Figure 1. Various Forms of Knowledge Acquisition.

sum total of things that one knows. The old fact/new guise response argues that Mary's learning is informative, but *factually unilluminating*; i.e., that it does involve her coming to know something that she did not know before, but does *not* involve or entail the existence of any *fact* of which she was previously ignorant. However, as we saw in the previous section, the possibility of informative knowledge acquisition on the part of a factually omniscient being such as Mary is quite doubtful. And of course no physicalist can allow that Mary's learning is factually illuminating, since that would entail the existence of a non-physical fact.

This leaves the physicalist with just one option. In order to give an ideal response to the knowledge argument, physicalism must construe Mary's learning as a case of *uninformative* truth apt knowledge acquisition; i.e., as an acquisition of knowledge that she can make even if there is *nothing* – no fact, no proposition, no ⟨fact, something-or-other⟩ pair, no information of any sort – of which she does not already have truth apt knowledge prior to her first encounter with color.

If there is just one relation of truth apt knowledge, so that having truth apt knowledge is always a matter of bearing that relation to some knowable, then uninformative truth apt

learning is impossible, and an ideal physicalist response to the knowledge argument unattainable. For in that case, one cannot gain new truth apt knowledge without gaining truth apt knowledge of something new – without coming to bear the one and only relation of truth apt knowledge to something to which one did not previously bear any relation of truth apt knowledge. Thus, if an ideal physicalist response to the knowledge argument is possible at all, it must be a response according to which there is *not* just one relation of truth apt knowledge.

Call this the “old relatum/new relation” response.¹⁴ What does it look like in detail? In particular, what is the knowledge relation that Mary is supposed to bear to a certain fact or piece of information only after her encounter with the tomato (despite already having borne *some* knowledge relation to it all along)? To answer this question, we can take a page from the acquaintance response. The key insight behind this was that it is a necessary truth – necessary in the strongest sense – that a person knows what it is like to experience red if, and only if, she has had a conscious red experience without forgetting it. Coonee’s argument for this claim invites various objections, but we can defend the claim by means of a more direct argument. First, we show that

- (1) It is impossible for someone to know what it is like to have a red experience without having had a conscious experience of red that he has not forgotten.

This is not to say that one must have a red experience in the normal way (i.e., by looking at a red thing in good light) in order to know what it is like to have a red experience. One might have the experience in a dream, or as a result of being struck hard in the face, or through direct neural stimulation (perhaps administered by the experiencer herself), or in some other way. While (1) implies that you must have a red experience *somehow* in order to know what it is like to have such an experience, it does not imply anything as to *how* you must have the experience, other than that you must have it consciously. This last proviso is to accommodate cases in which a person has

an experience unconsciously, in the way that a driver may experience his environment when driving down a familiar stretch of road. If Mary's encounter with the tomato had been of this sort – for instance, if she had been absorbed by one of her experiments at the time – then perhaps she would not have learned what it was like to have a red experience.

(1) also says that to know what it is like to have a red experience, you must not have *forgotten* your conscious experience of red. I do not want to put much weight on the word “forget,” in this context. Obviously, if Mary experiences serious head trauma after seeing the tomato with the result that she loses much of her memory, including her memory of the tomato, she can no longer be said to know what it is like to see red. At the same time, it may not seem very natural to describe this as a case of “forgetting.” Likewise, it might happen that Mary is so shocked when she sees the tomato that she neurotically suppresses her memory of the experience. In this case too, she cannot, perhaps, be said to know what it is like to experience red. The crucial thing is that Mary not have lost her ability to conjure up mental images with the quality of phenomenal redness, subsequent to her encounter with the tomato. Even if she remembers having seen the tomato, she will have forgotten the experience in the sense of “forget” that is relevant here, if she loses her ability to give herself a phenomenally red memory experience.¹⁵ (1) simply says that to know what it is like to have a red experience, you must have such an experience consciously, without subsequently forgetting it in this sense of “forgetting.” Thus understood, the claim is transparently true, as any dualist will surely agree.

How about the converse? This is the claim that

- (2) It is impossible for someone to have a red experience consciously and not forget it, without knowing what it is like to have a red experience.

With the foregoing clarifications about “forgetting” in place, it seems quite obvious that this claim is also true. If I am now consciously having a red experience, then I now know what it is like to have a red experience. Perhaps dragonflies have

conscious red experiences without knowing what it is like to have them. If so, this simply shows that their minds are too primitive to count as knowing anything beyond what they require to survive and propagate. But for Mary, it is impossible to have a conscious experience of red without at the same time knowing what it is like to have it. Of course, as we have already noted, it is possible to have a red experience and then forget it. But if you have had a red experience consciously and you have *not* forgotten it, then you automatically count as knowing what it is like to have a red experience. If someone who has had a red experience without forgetting it does not count as knowing what it is like to experience red, then no one counts as knowing what it is like to experience red.

So, putting (1) and (2) together: necessarily, one knows what it is like to experience red if and only if one has consciously experienced it without forgetting. Still, we do not want to say that learning what it is like to experience red just *is* consciously and unforgettingly experiencing red, since that would not account for the truth aptness of Mary's post-release knowledge acquisition. What the old relatum/new relation account allows us to say is that Mary's learning what it is like to experience red consists of her coming to bear to the (physical) fact expressed by "**this** is what it is like to see red" a truth apt knowledge relation that a person bears to a given fact just in case she knows that fact in an ordinary way (e.g., by virtue of bearing to it the sort of knowledge relation that Mary bears to all sorts of facts prior to her release) *and* has had a conscious unforbidden experience of red.

More generally, we can specify the new knowledge relation that Mary comes to bear to an old object of knowledge thus: it is the relation that a person x bears to the fact that y is what it is like to have phenomenal experience z , if and only if (i) x knows (in an ordinary, pre-release way) the fact that y is what it is like to have phenomenal experience z , and (ii) x has had z consciously and without forgetting. Call this the relation of *phenomenal knowledge*, and call the knowledge relation (or, relations) that Mary bears to facts prior to her release the relation(s) of *ordinary knowledge*. Then we can say that the new knowledge relation that Mary bears to some old knowledge

relatum is the relation represented by the formula: “ x bears a relation of ordinary knowledge to the fact that y is what it is like to have phenomenal experience z and x has had z consciously without forgetting.”

Exactly what fact *is* the fact that we know in knowing what it is like to see red? The most natural thing for a physicalist to say is that it is whatever fact we know, in knowing of the state or condition we are in when we have a red experience, that it is a state or condition of having a red experience. That is, the fact that I know in knowing that **this** is what it is like to experience red is the same fact as the one I know in knowing that **this** is the state of someone experiencing red. Mary knows this fact all along, since she knows just what (physical) states are states of experiencing red. She also knows precise neural facts about these states of which we laymen are ignorant. However, knowing these facts is not by itself enough to give Mary phenomenal knowledge of what it is like to experience red, and *not* knowing them does not prevent us from knowing that the state we are in when we have an experience like **this** is a state of experiencing red.

Let us consider the other case that Jackson brings forward against physicalism: the case of Fred.¹⁶ Unlike Mary, Fred does not have ordinary human vision. In particular, he is able to have color experiences of a kind that ordinary humans do not. Where the rest of us just see a bushel of red tomatoes, Fred sees a bushel of tomatoes, some of which are red, and the rest of which are some other color, which according to him differs from blue, green, yellow, or any other color that the rest of us can experience. It is not that Fred makes finer discriminations of red than we do – he does not divide the tomatoes into groups according to their possession of different shades of red. Rather, it is that certain objects that we experience as red, he experiences as having a color that we simply do not experience at all.

Fred’s case appears to make trouble for physicalism, because it seems that no amount of purely physical information about Fred – his eyes, his brain, etc. – can tell us what it is like to see the tomatoes in the bushel in the way he sees them. If that is true, the physical information about Fred’s visual experience

would appear not to exhaust the information about his visual experience, in which case physicalism is false.

This argument is subject to the same objection I have raised to the argument involving Mary. We *do* have all the information that there is to have about Fred's visual experience, if we have all the physical information about it. It is just that we lack phenomenal (as opposed to ordinary) knowledge of what Fred knows in knowing what it is like to see the extra color – which is to say that unlike Fred, we do not bear a certain truth apt knowledge relation to an item of information to which Fred himself does bear that relation. Jackson observes that we can imagine that “we find out how to make everyone's physiology like Fred's in the relevant respects,” and that this “would create enormous interest,” since it would put us in a position to “know what it is like to see the extra colour.” (Jackson, 1982, p. 129) This is true, but, again, consistent with physicalism, according to the old relatum/new relation account. There is no need for a physicalist to deny that there are items of truth apt knowledge about Fred that we cannot have unless we ourselves have the kind of experience he has; all he must deny is that there are *objects* of truth apt knowledge that we cannot know unless we have Fred's kind of experience. If there were a truth apt knowledge relation that required for its obtainment that the knower have justifiably believed that *p* while bench pressing five hundred pounds, then most of us would be incapable of having such knowledge too, at least without extensive training. But the fact that we did not know everything about Fred in this way would not suggest the existence of non-physical information about him, even supposing that we had (in one of the usual ways) all the physical information about him. Likewise, even though we do not know what it is like to have Fred's unusual sensations, it does not follow that there is more information to be had about his sensations than what someone with all the physical information about them has.

5. OBJECTIONS AND REPLIES

In this section, I anticipate three objections to the old relatum/new relation response. The first and most serious of these is that

the response is based on an *ad hoc* account of the logic of the concept of knowledge. The other two objections arise from residual doubts as to whether the old relatum/new relation account really satisfies the desiderata of truth aptness and austerity any better than competing physicalist responses to the knowledge argument. I address these objections in turn, beginning with the more serious one.

5.1. *Is the Old Relatum/New Relation Response ad hoc?*

In the previous section, I noted that if there were a relation of knowledge that required the knower to have bench pressed five hundred pounds, few people would be able to bear this knowledge relation to any fact concerning Fred, but that this would not imply the existence of non-physical facts about him. But of course there is *not* a knowledge relation that requires those who stand in it to have bench pressed five hundred pounds, and this might lead us to wonder why there should be a knowledge relation that requires those who stand in it to have had a relevant form of conscious experience. Given that there is no knowledge relation that we cannot stand in unless we have performed some athletic feat, or undergone some unconscious change, why should there be a knowledge relation that we cannot stand in unless we have had a relevant form of conscious experience? Absent an account of why our concept of knowledge should sometimes correspond to the relation of phenomenal knowledge, the old relatum/new relation response may seem *ad hoc*, positing fluctuations in the significance of epistemic concepts and language purely for the purpose of blocking the knowledge argument. Here, of course, it would be polemically weak to point to the posit's success in yielding an ideal physicalist response. What a dualist needs at this juncture is some reassurance that the concept of knowledge really works in the way that the old relatum/new relation response maintains.

Fortunately, such reassurance is not hard to provide. Consider, for example, the logic of *having*, and the corresponding semantics of the verb "to have." The old relatum/new relation

account says that pre-release Mary *has* all the information, but that this does not prevent her from acquiring information, since her having all the information is consistent with there being relations of having that she does not bear to various items of information. A casual consideration of the logic and semantics of having shows that this is not a strained or *ad hoc* thing to say. Analogously, we might say (truly) that a man acquires a kidney that he already had, e.g., if he undergoes a kidney transplant to receive a kidney that he had previously purchased on the black market. The capacity for a verb like “has” (or the concept of having) to express a variety of relations in ways like this is well known, so that the old relatum/new relation account’s requirement that this verb behave in this general sort of way is not at all unnatural.

How about the verb “to know” and the concept of knowledge in particular? Why should the relational content of these representations depend on the phenomenal history or background of those to whom we apply them? The natural and obvious explanation for this is that having had a certain form of conscious experience normally has an important bearing on a person’s thinking and behavior – more important, and much more probable, than the effects of bench pressing five hundred pounds, for instance. However things may be with Mary, who is far from normal, *normally* a person who has seen the world in color is apt to say, think, and do things that a person who has never experienced color is not apt to say, think or do. If I know you have experienced the odor of chlorine and not forgotten the odor, then I can instruct you to turn on the vents if at any time you detect an odor of chlorine, confident that you will be able to follow my instructions should the need arise. If a third party is present and he also knows that you have smelled chlorine and remember the smell, he can reliably predict that you will turn on the vents, if you smell chlorine. He could also predict this if he knew you were omniscient; but real people are not omniscient, which is why our concept of knowledge pays attention to the phenomenal histories of those to whom we apply it.

We learn in many ways, but the most basic way, at least from a survival standpoint, is by having conscious experience of

ourselves and our environment. Phenomenal knowledge is not a marginal or exotic form of truth apt knowledge: it is the most elemental of all forms of truth apt knowledge. Other forms come later in development, along with the acquisition of language and the accumulation of information gathered from an ever increasing variety of sources. A baby may not know that there is water in its tub, and it certainly does not know that water is H₂O, but it knows what it feels like to be held by its mother, and we know that it knows this, and treat it accordingly.

Since the main point of knowledge attributions and other uses of epistemic concepts is to predict, influence, and otherwise make sense of normal people's behavior, it is entirely natural that these concepts would have evolved so as to be sensitive to the presumed experiential status and history of those to whom we apply them, in the way that the old relatum/new relation account maintains. As we have seen, there is a clear and vital role for the concept of knowledge to play in tracking the experiential history of those to whom we apply it, and a deep conceptual connection between knowing what it is like and having had the relevant kind of conscious experience (without forgetting). Far from being *ad hoc*, the old relatum/new relation response affords a very natural account of the way we use epistemic concepts to ascribe phenomenal knowledge, and of how these uses fit into our overall use of the concept of truth apt knowledge.

5.2. *Is the Old Relatum/New Relation Response Really Truth Apt?*

A second objection to the old relatum/new relation response is that it ultimately fares no better than the acquaintance response, when it comes to the desideratum of truth aptness. Now, it is true that Mary does not gain knowledge of any broadly truth-assessable item of which she did not already have knowledge, according to the old relatum/new relation account. But this does not change the fact that the knowledge she gains is truth apt. It is just that she acquires her truth apt knowledge (of what

it is like to experience red) simply by having a conscious experience of red, since, unlike us, she already possessed knowledge of all the relevant facts and information. This does not mean that it would be uninformative for anyone else to learn what it was like to experience red. If Mary had been a lazy teenager instead of a brilliant scientist, then her first encounter with color would no doubt have given her some new information. This is because if she had merely been a lazy teenager, she would not have had all (or perhaps even any) information concerning red experience. The old *relatum/new relation* account can grant that the acquisition of phenomenal knowledge normally involves an acquisition of new information, and that the possession of phenomenal knowledge *always* involves the possession of relevant information, without granting that Mary herself gains new information upon learning what it is like to experience red.

Mary can also be said to acquire her new knowledge on an evolutionary technicality: since, unlike normal people, Mary already knew all the physical facts, her pre-release epistemic condition was already as selectively advantageous as it could well be. This does not make her acquisition of knowledge any less real, however, but only of less practical moment than it would be for someone who was not as knowledgeable as Mary. Again: even though *Mary's* acquisition of the knowledge of what it is like to see red does not involve any acquisition of new information, it does not follow (and is not the case) that learning what it is like to see red never involves an acquisition of new information by the learner.

5.3. *Is the Old Relatum/New Relation Response Really Austere?*

One might think that the old *relatum/new relation* response fares no better than the old *fact/new guise* response, since it concedes that there is knowledge that pre-release Mary lacks, despite her possession of all the (physical) information. But this is not a concession that poses any problem for physicalism. Only if a lack of knowledge entailed the existence of something of which knowledge was lacking would the old *relatum/new*

relation theory's allowance that there is knowledge that pre-release Mary lacks carry the physicalistically undesirable implication that there is some object of knowledge that pre-release Mary does not know. But, as the possibility of uninformative truth apt knowledge acquisition shows, a lack of knowledge does *not* entail the existence of something of which knowledge is lacking. Unlike the old fact/new guise theory, the old relatum/new relation account grants that knowing all the facts entails having all the information, and that pre-release Mary is, therefore, all-knowing. All the old relatum/new relation account concedes is that fully informed does not mean fully enlightened, and this is a concession that it can afford to make.

One might also doubt whether *any* austere account of Mary's learning can really account for all the epistemic changes she undergoes. After Mary has been out and about in the colorful world for a while, she knows that **this** and not **this** is what it is like to experience red. But, thinking back on her grayscale days, can she honestly say that she had this knowledge even then? And if not, doesn't that mean that there are possible worlds that she could not rule out then, but can rule out now?

Before her release, Mary did not know what it was like to experience red, or whether experiencing red was like **this** rather than **this**, since knowing what it is like to experience red, or that experiencing red is like **this** rather than **this**, is a matter of having phenomenal knowledge of circumstances of which she had only ordinary knowledge, prior to her release. But given that she had ordinary knowledge of all the physical facts and all that these facts entail, she had the actual world completely pinned down all along. She simply had not had the relevant experiences, or been able to exercise certain abilities that presuppose having had such experiences (such as remembering the experiences). The impression that her release into the colorful world allows her to rule out certain possible worlds that she was not previously in a position to rule out arises from the fact that a normal person who learned that experiencing red was like **this** rather than **this** *would* thereby be put in a position to rule out certain possible worlds that he was previously unable to rule

out. This is because a normal person is not omniscient as to the relevant facts. But since Mary does know all the relevant facts, for her to learn that **this** rather than **this** is what it is like to experience red requires no more than for her to have the relevant experiences.¹⁷

6. CONCLUSION

In this paper I have developed a physicalist response to the knowledge argument that sees Mary's learning as belonging to one of two possible forms of enlightening experience: the uninformative form. Since it is possible for someone with all the information concerning color vision to have an enlightening experience of the sort Mary has, and to learn as Mary learns, the second premise of the knowledge argument is false. This response to the knowledge argument is ideal, in the sense that it grants the truth apt character of phenomenal knowledge and learning, without granting the existence of more information than what someone with all the physical information has.

At the level of details, the old *relatum/new* relation response is not particularly novel. From Conee, it borrows the insight that acquiring a given item of phenomenal knowledge essentially involves having a relevant form of experience, under suitable conditions. From the old *fact/new* guise theorists, it takes the view that the knowledge argument's fundamental flaw lies in its misconstrual of the distinctive logic of epistemic terms and concepts. And, following Bigelow and Pargetter, it draws on the often overlooked possibility of uninformative truth apt knowledge acquisition. The main accomplishment of the present paper is its assembly of these so far disparate elements into a single, cohesive response to the knowledge argument that is both uncompromising and polemically effective.

NOTES

¹ For Jackson's original statement of the argument, see Jackson (1982). I should note that Jackson himself no longer endorses the knowledge argument in any form: see Jackson (1998).

² In this paper I consider only Jackson's knowledge argument. In particular, I do not consider the argument defended in Chalmers (2004), which (omitting minor technical details) runs as follows:

1. "P \supset Q" is *a posteriori*.
2. If "P \supset Q" is *a posteriori*, then "P \supset Q" is false in some possible world considered as actual.
3. If "P \supset Q" is false in some possible world considered as actual, then "P \supset Q" is false in some possible world.
4. If "P \supset Q" is false in some possible world, then physicalism is false.
5. Therefore, physicalism is false.

Here, "P" is a comprehensive statement of actual physical fact, and "Q" a statement of the instantiation of phenomenal redness. A careful discussion of Chalmers' argument is beyond the scope of this paper, and I shall not address it here, other than to make the following three remarks.

First, Chalmers' argument is quite different from Jackson's, as can be seen from the fact that, unlike Jackson's, Chalmers' calls on the concept of aposteriority and a primitive notion of considering a (*prima facie*) possible world as actual. These differences are important, since the polemical force of Jackson's argument comes largely from the fact that it relies solely on familiar everyday concepts.

Second, although Chalmers describes his argument as a version of the knowledge argument, it is not in fact clear that it is a version of the knowledge argument, as opposed to a version of the modal argument against physicalism. Chalmers appeals to Jackson's original thought experiments only to support the first premise, that the (material) implication of the phenomenal facts by the physical facts can be known only *a posteriori*. But it is difficult to see what the claim of aposteriority comes to, beyond that someone with comprehensive physical knowledge could conceive of the world being just the way she knows it to be physically without being compelled to conceive of it as being this versus that way, phenomenally. In that case, the argument is an attempt to derive the failure of physicalism from the conceivability of worlds just like the actual world physically but differing from it phenomenally; i.e., it is a modal argument, not a version of the knowledge argument.

Third, the broadly Fregean "two-dimensional" semantic framework that forms the basis of Chalmers' analysis of the knowledge argument is not obviously mandatory. There are alternative semantic frameworks more friendly to physicalism that dispense with the apparatus of epistemic intensions, and appear to do at least as good a job as the two-dimensional framework of explaining the distinctive logic of cognitive terms and concepts (see, e.g., Pelczar, 2004). A careful discussion of Chalmers' argument

would require us to weigh the pros and cons of these vying semantic frameworks, which, again, is beyond the compass of the present paper.

³ See Dennett (1991) and Jackson (1998). This approach is not to be confused with the idea that Mary might, with all her knowledge, be able to give herself a red experience whenever she wants. That may be true. But the question is whether she could know what it was like to have a red experience without actually having one (as a result of modifying the structure of her visual cortex through sheer will, or in some more normal way).

⁴ Of course, if Mary had her experience at 0700 GMT on December 26, 2003 and knew this at the time, then the fact that she had it at 0700 GMT December 26, 2003 is one that she knew after but not before her experience. But if, as we are to suppose, Mary already knew all the relevant physical facts even before having her first red experience, then she already knew to be true the proposition that Mary has a red experience at 0700 GMT December 26, 2003. What austerity demands is that Mary already have known every timeless (as opposed to tensed) knowable, prior to her release.

⁵ See Conee (1994, p. 140, 147).

⁶ Torin Alter points out that we can divide episodes of acquaintance-making into two categories: those that do, and those that do not, involve an acquisition of new information on the part of the individual making the acquaintance. If Mary's acquaintance with the quality of phenomenal redness falls into the first category, then (*contra* physicalism) she gains information after all, whereas if her acquaintance with the quality falls into the second category, it does not fully account for her learning (see Alter 1998, pp. 39–40). This criticism is echoed by Brie Gertler, who argues that if knowing what it is like to experience red is simply a matter of becoming acquainted with the quality of phenomenal red, then it is a mystery how becoming acquainted with the quality could issue in any properly *epistemic* change for the better on Mary's part (see Gertler 1999, pp. 327–28).

⁷ The word “**this**” is to be read while having a red color experience, such as the one you just had if you are reading the colored version of this article available at: <http://www.kluweronline.com/issn/0031-8116/>. In the colored version, the word “**this**” appears in red, and the word “**this**” in blue.

⁸ See Loar (1990, pp. 86). One thing that Conee's account *does* allow is the embeddability of complete phenomenal knowledge *attributions*, such as “John knows what it is like to have a red experience.” If (prior to seeing the tomato) Mary says, “If John knows what it is like to have a red experience, then he isn't stuck in a black and white prison,” she speaks truly, and, given Conee's account, what she says is tantamount to this: that if John has had a conscious red experience, then he does not live in a black and white prison. Similarly, if Mary wonders whether everyone who knows what it is like to have a red experience also knows what it is like to have a pink experience,

what she wonders is whether everyone who has had a conscious red experience has had a conscious pink experience. Other logical embeddings of complete phenomenal knowledge attributions may be dealt with similarly, given the resources of the acquaintance account. But the embeddability of phrases like “**this** is what it is like to see red” themselves cannot be explained so easily.

⁹ See Horgan (1984, pp. 149–52). Horgan actually speaks of “ontologically physical information” where I speak of facts, but it comes to the same, since the crucial point for Horgan is that while there is a one-one correspondence between facts and items of ontologically physical information, there is a one-many correspondence between facts and items of information broadly construed.

¹⁰ See Loar (1990, pp. 87–90). It may be more appropriate to read Loar as advocating a view closer to the one I shall defend in the next section than to standard old fact/new guise accounts, since we might be able to construe the differences between the cognitive roles of his “phenomenal recognitional” and “physical-functional” concepts as differences in how such concepts bear on the content of the concept of knowledge as we apply it to pre- versus post-release Mary, rather than as differences between the cognitive content of the phenomenal or physical concepts themselves. Here I simply follow the literature in categorizing Loar’s account as a (sophisticated) version of the old fact/new guise response – see esp. Chalmers (1996, pp. 142–43).

¹¹ See Lycan (1990, pp. 113–14, 117–22).

¹² This is a streamlined version of Chalmers’ primary objection to the old fact/new guise approach: see Chalmers (1996, pp. 141–43). As Alter (1995) points out, it is actually inessential to this objection that what Mary acquires is information, as opposed to misinformation. For even if all she acquires is misinformation (as Pereboom (1994) argues), Mary’s acquisition of this misinformation appears to presuppose the existence of a fact that she did not know prior to leaving her black and white prison; viz., the fact that this misinformation *fails* to correspond to the same fact as her old information. This too is a fact, and one of which Mary was (the dualist maintains) previously ignorant.

¹³ It is true that a non-Chinese speaker could fail to understand a statement made in Chinese despite knowing exactly what fact it stated, perhaps having it on the authority of a fluent bilingual that the statement states that mangos cost two dollars apiece. In this case, the lack of understanding results from a lack of linguistic knowledge. Such an explanation is not available in Mary’s case, however, since unlike someone who has no knowledge of Chinese, and who confronts a Chinese sentence as a black box even if he knows what fact it states, Mary not only knows what fact my utterance of “This is what it is like to experience red” states, but also how this utterance logically relates to my other utterances, as well as to her own utterances.

¹⁴ A germinal form of this response can be found in Bigelow and Pargetter (1990) – see esp. Bigelow and Pargetter (1990, pp. 138–44). Like me, Bigelow and Pargetter recognize that the Truth Apt node of Figure 1 does not divide immediately into the factually illuminating and factually unilluminating branches, but first breaks into informative and uninformative. Unfortunately, while Bigelow and Pargetter recognize the possibility of uninformative truth apt knowledge acquisition, they fail to make it clear what the distinctive merits are of a response to the knowledge argument that construes Mary’s learning as of the uninformative truth apt variety. For example, they maintain that Mary’s first experience of color “enables her to discriminate among new representable possibilities” (op. cit., p. 144). Such remarks obscure the essential point of the old relatum/new relation theory. For the essential point of this theory, and its whole advantage over the old fact/new guise response to the knowledge argument, is that it gives us a way of *denying* that there is any information that pre-release Mary lacks. By conceding that Mary’s learning enables her to discriminate possibilities that she could not discriminate before her release, Bigelow and Pargetter implicitly concede that she gains information; viz., the information that there are these newly discriminated possibilities. Despite these unclarities, however, Bigelow and Pargetter certainly deserve credit for being the first to identify the possibility of an old relatum/new relation response to the knowledge argument.

¹⁵ Here there is a point of contact between my proposal and the ability objection to the knowledge argument, according to which Mary’s learning is of a truth inapt variety that consists of her acquiring certain new abilities, such as the ability to remember red experiences in the sense delineated here; see esp. Nemirow (1980) and Lewis (1990). The ability objection corresponds to the Practical sub-branch of the Truth Inapt branch of Figure 1.

¹⁶ See Jackson (1982, pp. 128–30).

¹⁷ The objection discussed in the preceding two paragraphs corresponds to Chalmers’ objection to what he takes to be Loar’s version of the old fact/new guise response (Chalmers, 1996 pp. 142–43). As Chalmers interprets it, Loar’s account says that there is no difference between what Mary knows before and after her release, but that this does not put her in a position to know *a priori* that a certain physical-functional predicate (or sentence) has the same “primary intension” as a predicate that introduces the property of phenomenal redness (or, a sentence that states the instantiation of this property). Thus, even though the sentence “**this** is what it is like to see red” has the same truth value as a certain physical-functional sentence “ Σ ” relative to every possible (or impossible) world considered as actual, according to Loar this does not imply that Mary’s knowledge of the circumstance that Σ puts her in a position to deduce *a priori* that **this** is what it is like to see red. Chalmers’ objection is that it is implausible to suppose that Mary would not be able to make this deduction, if indeed the two sentences were alike in

truth value in every world that is epistemically possible for her, prior to her release. The old relatum/new relation account is not subject to this objection, because it can simply grant that pre-release Mary's knowledge of the circumstance that Σ puts her in a position to know whatever anyone knows in knowing that **this** is what it is like to see red. Again: all that we must deny, by my account, is that Mary stands in a relation of phenomenal (as opposed to ordinary) knowledge to what one knows in knowing what it is like to see red.

REFERENCES

- Alter, T. (1995): 'Mary's New Perspective', *Australasian Journal of Philosophy* 73, 582–584.
- Alter, T. (1998): 'A Limited Defense of the Knowledge Argument', *Philosophical Studies* 90, 35–56.
- Pelczar, M. (2004): 'The Indispensability of *Farbung*', 138(1), 49–77.
- Bigelow, J. and Pargetter, R. (1990): 'Acquaintance With Qualia', *Theoria* 56, 129–147.
- Chalmers, D. (1996): *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Chalmers, D. (2004): 'Phenomenal Concepts and the Knowledge Argument', in P. Ludlow, Y. Nagasawa and D. Stoljar (eds.), *There's Something about Mary: Essays on Frank Jackson's Knowledge Argument Against Physicalism*, Cambridge: MIT Press.
- Conee, E. (1994): 'Phenomenal Knowledge', *Australasian Journal of Philosophy* 72, 136–150.
- Dennett, D. (1991): *Consciousness Explained*, Boston: Little Brown and Company.
- Gertler, B. (1999): 'A Defense of the Knowledge Argument', *Philosophical Studies* 93, 317–336.
- Horgan, T. (1984): 'Jackson on Physical Information and Qualia', *Philosophy and Phenomenological Research* 34, 147–152.
- Jackson, F. (1982): 'Epiphenomenal Qualia', *Philosophical Quarterly* 32(127), 127–136.
- Jackson, F. (1998): 'Postscript on Qualia', in *Mind, Method, and Conditionals: Selected Essays*. Routledge: London, pp. 76–79.
- Lewis, D. (1990): 'What Experience Teaches', in W. Lycan (ed.), *Mind and Cognition*. Oxford: Blackwell, pp. 499–518.
- Loar, B. (1990): 'Phenomenal States', in J. Tomberlin (ed.), *Philosophical Perspectives IV: Action Theory and Philosophy of Mind*, Ridgeview: Atascadero, pp. 81–108.

- Lycan, W. G. (1990): 'What Is the "Subjectivity" of the Mental?', in J. Tomberlin (ed.), *Philosophical Perspectives IV: Action Theory and Philosophy of Mind*, Ridgeview: Atascadero, pp. 109–130.
- Nemirow, L. (1980): 'Review of Mortal Questions, by Thomas Nagel', *Philosophical Review* 89, 473–477.
- Pereboom, D. (1994): 'Bats, Brain Scientists, and the Limitations of Introspection', *Philosophy and Phenomenological Research* 54(2), 315–329.

Department of Philosophy
National University of Singapore
3 Arts Link, Singapore 117570
Singapore
E-mail: phimwp@nus.edu.sg