

GEORGE GRAHAM and TERENCE HORGAN

MARY MARY, QUITE CONTRARY

When the History of the Twentieth Century Intuition Pump¹ is written, chroniclers will tell of a scientifically omniscient physical scientist named Mary who was confined to a black and white room and of the argument she bore, the so-called Knowledge Argument, designed by her creator Frank Jackson to show that physicalist/materialist² theories of phenomenal consciousness are false. What's more, they are false because, as the thought experiment of Mary purports to show, any adequate account of phenomenal consciousness requires reference to non-physical *qualia* – to self-presented qualities of conscious experience, understood as properties over and above those that appear in the physicalist story of mind and matter.

In our opinion, Jackson's argument is immensely powerful, but alas, currently its force is not sufficiently appreciated. This is because even among those convinced of its power (cf. McConnell, 1994; Nida-Rumelin, 1995), the consensus has emerged that the argument is inadequate as it stands, and that the thought experiment of brilliant Mary is, to borrow two words of a critic, 'poorly conceived' (Thompson, 1995, p. 264). Indeed, Jackson himself has converted to physicalism and joined the consensus (see Jackson, 1998, p. 101).

Our concern in this paper is to redesign the experiment and then to convince the reader that properly conceived, the Mary Intuition Pump poses a serious challenge to physicalism. Through this redesign or reconceptualization we hope to motivate proper appreciation of the experiment and the Knowledge Argument. We don't aim to defend anti-physicalism. Our personal attractions to physicalism are such that it is hard for us to accept anti-physicalism. However it is also hard for us to dismiss Mary. Mary raises problems for materialist stories. They all seem to have a phenomenal residue.



Philosophical Studies 99: 59–87, 2000.

© 2000 Kluwer Academic Publishers. Printed in the Netherlands.

How do we propose to redesign the experiment? Our redesign will consist of building into Mary expertise as a materialistically minded philosopher in addition to her omniscience as a neuro- or physical scientist. Mary's challenge to materialism, which will be explained later in the paper, then goes something like this. Take any physicalist story about phenomenal consciousness of the sort that purports to take Mary's in-her-black-and-white environment/out-of-that-environment epistemic contrast seriously. Then, Mary can be described as being able to understand this story *before* leaving her black and white environment but therein being unexpectedly surprised by the character of her experiential epiphany on leaving the environment. The challenge consists in charging materialism with being unable to explain the rational appropriateness of Mary's surprise upon release; given the materialist story about phenomenal consciousness that she initially accepts, she has no apparent rational basis for such a post-release reaction. There is strong reason, albeit reluctantly acknowledged by us, to doubt whether materialism can meet this challenge, and therein to trust Jackson's original moral for the Knowledge Argument that there are phenomenal properties that escape the physicalist story.

We begin by describing the original thought experiment.

THE ORIGINAL EXPERIMENT

Here is the original experiment in Jackson's own words.

Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room *via* a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wave-length combinations from the sky stimulate the retina, and exactly how this produces *via* the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence "The sky is blue". (Jackson, 1982, p. 130)

Let us suppose that Mary's physical knowledge is even more complete – she actually becomes physically scientifically omniscient – on the basis of her black-and-white books and television lectures, as Jackson writes in a second paper:

In this way she learns *everything* [emphasis added] there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of ‘physical’ which 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, 1984, p. 291)

On the basis of this Mary thought experiment, Jackson argues that physicalism is false, meaning by ‘physicalism’ the doctrine that “all information is physical information”.

What will happen when Mary is released from her black and white room or is given a color television monitor? Will she *learn* anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had *all* the physical information. *Ergo* there is more to have than that, and Physicalism is false. (Jackson, 1982, p. 130)

To use Nagel’s (1974) canonical phrase, according to Jackson, what Mary learns which is not physical information is the “what it’s like” of visual color experience (of sensing redly, bluey, and so on). She discovers, or is presented with, a special sort of information; it is physically indescribable phenomenal information.

CONFINING MARY

Many parties to debate about Mary – both friends and foes of the Knowledge Argument – agree that Jackson’s original thought experiment is not well designed. Sometimes what is said to be missing is more detail about her pre-release situation or black and white environment.³ It has been pointed out, for example, that simply confining her to a black and white environment does not necessarily deprive her of color experience. She may see color when she dreams, rubs her eyes, or has afterimages produced as a consequence of brightness or lightness perception (Thompson, 1995, p. 264). To remedy defects in the description of her confined situation, other suppositions must be introduced into the thought experiment. We may suppose, for example, that she has been completely monochromatic from birth and that her ability to see afterimages in color has atrophied. We may also suppose that she is released from her room,

while unconscious, and then undergoes a medical procedure which permits her to see colors.

Refining the thought experiment in these ways we call the problem of conceiving of Mary's *confined situation*. Proper description of Mary's physical confinement (i.e. of her pre-release situation) insures that, though she is a master of completed physical science, her visual experience is monochromatic.⁴ Of course, deeper, more puzzling design problems characterize debate over the Knowledge Argument's purported anti-physicalist implication. Let's introduce these by reference to materialist resistance to the Argument.

MATERIALIST RESISTANCE TO THE KNOWLEDGE ARGUMENT

Materialist resistance to the anti-physicalist conclusion of Jackson's Knowledge Argument may be classified in different ways, though perhaps the principal division is between those materialists who concede that when Mary leaves her perceptual-physical confinement she acquires new information and those who do not make this concession. The second type of materialist admits that so long as Mary acquires new information physicalism is in trouble, but argues that we should refrain from saying that Mary acquires new information. On this view the difference between Mary confined and Mary unconfined consists in something other than acquiring novel knowledge. This second type of materialism requires a strategy we call *epistemically thinning* the contrast between Mary's pre-release and post-release situation. The brand of materialism we call *thin* materialism.

Materialists have devised, on our reading of the literature, two main ways of developing the second, the thin, type of materialist position and of thinning the contrast between Mary's pre-release and post-release. One consists of endowing confined Mary with stunning powers of imagination to go along with her neuro-omniscience (Dennett, 1991; Churchland, 1989). It consists of claiming that if Mary knows everything physical that there is to know about visual color experience, then she can figure out or imagine what it is like to see color, and if she can figure this out, then she can know what

it is like to experience polychromatically even when she is confined and thus before her release.

The other way of developing the thin position consists of saying that Mary before her release is not ignorant of possible knowledge, but merely lacks the ability or skill to identify visual color experience directly or immediately or 'from the introspective inside' (Nemirow, 1980; Lewis, 1990). Mary's pre-release ability to identify red, blue, and so on is indirect, consisting perhaps of using scientific instruments which tell her whether objects outside the black and white environment are red or blue. What's more, acquiring this skill after confinement does not require acquiring new knowledge; it does not presuppose possessing novel information. Mary already has the needed information in her black and white environment.

Notice that in the case of both the imaginability strategy and the non-epistemic ability strategy, since there is no new information or knowledge gained on release, there is no reason for including non-physical information in Mary's post-release comprehension of color and color experience. There are serious difficulties with thin materialism, however, which have been well documented by others (cf. McConnell, 1994; Lycan, 1996). Neither strategy establishes that Mary fails to secure new (phenomenal) information on release or that she does not require such information to 'inside' (introspectively) understand color and color experience.⁵

Given that thin materialism has been amply criticized by others (though not by that name of course), here we can safely assume that this form of materialism is unpromising, and that it is implausible to deny that Mary makes epistemic progress post-release. Some sort of new information or knowledge comes her way after confinement. Indeed, most materialists readily acknowledge that Mary acquires new knowledge post-release. Such admission deserves to be described as the received physicalist view of the Knowledge Argument.

The received physicalist view consists in saying that Mary makes epistemic progress after her release, polychromatically, and acquires new knowledge. However this does not produce the metaphysical conclusion that there are non-physical properties. The new information is epistemically progressive, but not metaphysically fecund. In the terminology of Horgan (1984), although this newly acquired

information is not *explicitly* physical – i.e., it does not belong to, or follow from, a theoretically adequate physical account of human perceptual processes – it is nonetheless *ontologically* physical – i.e., all the entities and properties involved are physical ones.

This second type of materialism (which concedes new knowledge to post-release Mary) requires a strategy we call *epistemically thickening* Mary's post-release situation and therein preserving *some* epistemic contrast between Mary pre- and post-release. The brand of materialism itself we call *thick* materialism.

How can thick materialism be developed? Thick materialism is sometimes called the 'mode of knowledge', 'manner of presentation', or 'access route' materialist strategy for resisting the anti-physicalist implication of the Knowledge Argument. We will take one line of development as representative for purposes of evaluating how a materialist of the thick sort responds to the Knowledge Argument. This is Michael Tye's (1995) theory that phenomenal content is what he calls PANIC, an acronym for 'Poised Abstract Nonconceptual Intentional Content'.

PANIC THEORY

The most relevant ideas behind Tye's theory are as follows. There is something about conscious experience which deserves to be called its "what it's likeness," i.e., its phenomenal "content" or "character." For Tye, phenomenal content is identical to the state's PANIC – its suitably poised, abstract, nonconceptual, intentional content. Tye says:

Phenomenal content, I maintain, is content that is appropriately poised for use by the cognitive system, content that is abstract and nonconceptual. I call this the PANIC theory of phenomenal character: phenomenal character is one and the same as Poised Abstract Nonconceptual Intentional Content. . . . The claim that the contents relevant to phenomenal character must be *poised* is to be understood as requiring that these contents attach to the (fundamentally) maplike output representations of the relevant sensory modules and stand ready and in position to make a direct impact on the belief/desire system. . . . The claim that the contents relevant to phenomenal character must be *abstract* is to be understood as demanding that no particular concrete objects enter into these contents. . . . Since different concrete objects can look or feel exactly alike phenomenally, one can be substituted for the other without any phenomenal change. . . . The claim that

the contents relevant to phenomenal character must be *nonconceptual* is to be understood as saying that the general features entering into these contents need not be ones for which their subjects possess matching concepts. . . . Consider . . . color. . . . We have names for only a few of the colors we can discriminate, and we also have no stored representations in memory for most colors either. There is simply not enough room. (pp. 137–139)

Tye construes the intentionality of phenomenal states – the ‘I’ in ‘PANIC’ – as a matter of causal covariation between representing state and item represented. “The key idea,” he says, “is that representation is a matter of causal covariation or correlation (*tracking*, as I shall often call it) under optimal conditions” (p. 101). Concerning the representational content of phenomenal states, he says the following. (Red₂₉ is a specific, fine-grained, shade of red.)

Which features involved in bodily and environmental states are elements of phenomenal consciousness? There is no a priori answer. Empirical research is necessary. . . . They are the features our sensory states track in optimal conditions. . . . I conjecture that for perceptual experience, [these] will include properties like being an edge, being a corner, being square, being red₂₉. (pp. 137–141)

As Tye observes, PANIC is a broadly *physical* property, in the sense that there are *conceptually sufficient physical conditions* for its instantiation (pp. 163–164).

Although one can understand the nature of PANIC in a objective way, Tye maintains that there are certain concepts associated with conscious experience that are essentially subjective and perspectival, and can only be possessed and deployed on the basis of having undergone the relevant conscious experiences oneself. These *phenomenal* concepts, as he calls them, are required in order to possess the kind of knowledge of phenomenal character that we call “knowing what it’s like.” He says this about phenomenal concepts:

I call the concepts relevant to knowing the phenomenal character of any state ‘phenomenal concepts.’ Phenomenal concepts are the concepts that are utilized when a person introspects his phenomenal state and forms a conception of what it is like for him at that time. These concepts, in my view, are of two sorts. Some of them are indexical; others are predicative. Suppose, for example, I am having a visual experience of red₂₉. I have no concept red₂₉. So, how do I conceptualize my experience when I introspect it? I bring to bear the phenomenal concepts *shade of red* [a predicative phenomenal concept] and *this* [an indexical phenomenal concept]. Intuitively, possessing the phenomenal concept [*shade of red*] requires that one have experienced red and that one have acquired the ability

to tell, in the appropriate circumstances, which things are red directly on the basis of one's experiences. . . . What about the phenomenal concept *this*? Possessing this concept is a matter of having available a way of singling out, or mentally pointing to, particular features that are represented in sensory experiences *while* they are present in the experiences, without thereby describing those features (in foro interno). . . . What one has [in having the indexical concept] . . . is a way of singling out or discriminating the feature for as long as one attends to it in one's experience (and perhaps for a very short time afterward). (pp. 167–168)

So Tye treats phenomenal concepts as *capacity-based* concepts (as we will call them). For him, possession and deployment of phenomenal concepts is essentially a matter of having certain cognitive capacities: in the case of red, for instance, possession of the relevant predicative phenomenal concept is essentially the ability to distinguish red things from non-red things directly on the basis of one's experiences (and without collateral information); likewise, possession of the relevant indexical phenomenal concept is essentially the ability to indexically pick out, in thought, a shade of red that is currently being represented PANIC-wise in one's experience (e.g., red₂₉).

“Knowing what it's like” to see red is, for Tye, a matter of representing the phenomenal character of one's own experience *under a phenomenal concept*. He says:

Phenomenal concepts, as described, are crucial to knowing phenomenal character. Now, in the case of knowing via phenomenal concepts, knowing what it is like to undergo a phenomenal state type P demands the capacity to represent the phenomenal content of P under those concepts. But one cannot possess a predicative phenomenal concept unless one has actually undergone token states to which it applies. It follows that knowing the phenomenal character of P via predicative concepts requires having experienced tokens of P.

In the case of knowledge via the phenomenal indexical, knowing what it is like to undergo P demands that one mentally point to the content of P while one is experiencing a token of P (or immediately afterward). Again, then, the relevant experience is required. (p. 169)

Just as one can prespectively represent red things, and the property of redness, via the predicative phenomenal concept *red*, so likewise one can perspectively represent – again via the predicative phenomenal concept *red* – the phenomenal character of one's conscious experience itself. One thinks to oneself, “Ah, *red* is what it is like to see red things” – a thought one cannot have without possessing

the phenomenal concept *red*, which in turn requires having undergone the relevant conscious state oneself. Likewise for indexical phenomenal concepts: one thinks to oneself, “Ah, *this* is what it’s like to see red things” – a thought one cannot have without being able to mentally point to a specific shade of red while one is experiencing it visually (or immediately afterward). “Knowing what it’s like” is thus a mental state that one cannot enter into without having the cognitive capacities that constitute possession of phenomenal concepts. And the acquisition of these cognitive capacities, in turn requires actually undergoing the relevant conscious experiential states oneself. So according to Tye’s theory, “knowing what it’s like” is an essentially subjective, essentially perspectival, cognitive state – despite the fact that the phenomenal character of any type of conscious experience is itself a broadly physical property that also can be known in a purely objective way, viz., a PANIC property.

MARY’S EPISTEMIC PROGRESS

The application of this approach to Mary is now straightforward. As Jackson himself maintained in his original paper, Mary does indeed obtain new knowledge upon beginning to have color experiences. For, only then does she acquire capacity-based phenomenal color-concepts; thus, only then does she become able to represent the relevant PANIC properties *via* the deployment of those capacity-based concepts, thereby knowing “what it is like” to undergo color experiences. Nonetheless, the PANIC properties themselves are broadly *physical* properties.

Well then, does Mary learn any new *facts*, after she first has color experiences? Yes and no, says Tye, depending on how we employ the notion of a fact. Mary learns new *fine-grained* facts – facts individuated as contents of distinct propositional attitudes. She knows new fine-grained facts because she acquires new *beliefs* that count as new items of knowledge – beliefs whose content involves her newly acquired, capacity-based, phenomenal concepts. However, on Tye’s account she does not learn any new *coarse-grained* facts – facts identified as “states of affairs that obtain in the objective world, regardless of how those states of affairs are conceived” (p. 173). For, the *properties* referred to by her newly acquired phenomenal

concepts – the properties involved in the new beliefs that constitute her new knowledge – are broadly *physical* properties that are also knowable in a purely objective way (viz., PANIC properties). As Tye himself puts it:

Mary does not know what it is like to experience red. So, on my view, she does not know the phenomenal content of the state of experiencing red (whatever the determinate shade). She does not know this for two reasons. First, she *lacks* the [predicative] phenomenal concept *red*; second, she cannot apply the [indexical] phenomenal concept *this* to the color represented in experiences of red. After all, Mary has never had the experience of red. . . . Still, the state of experiencing red can have a naturalized PANIC essence, as I have argued. And Mary will know that essence (as involving such and such causal correlation, etc.) if she knows all the facts countenanced by physicalism. So there is nothing of a nonconceptual sort not known to Mary. The fact she does not know is a fine-grained one within which there are phenomenal concepts. However, the coarse-grained, nonconceptual fact it contains *is* broadly physical. (p. 174)

Tye's approach is thus a nice example of what we earlier called "thick materialism." His account thickens Mary's post-release epistemic situation without eschewing materialism, since it entails that post-release Mary learns new *fine-grained* facts without learning any new *coarse-grained* facts. So physicalism emerges unscathed under Tye's PANIC theory of phenomenal consciousness, even though Mary obtains genuinely new knowledge upon release from her monochrome environment.

THE MARY MARY CHALLENGE

Is Mary's epistemic progress best described by thick materialism? Does she gain new knowledge of old (i.e., broadly physical) properties or does her progress involve knowledge of new coarse-grained facts involving non-physical properties? In this section of the paper we begin to explore this question by advancing a challenge for Tye's approach in particular, and for thick materialism in general. We call this the Mary Mary Challenge.

In addition to the problem of describing Mary's physical confinement, we suppose that there is a problem about describing Mary's metaphysical convictions (if any) and the contribution which these convictions make to her epistemic progress on release. To reduce

thought-experimental confusion and to set up this problem, suppose that Mary has a daughter, Mary Mary Quite Contrary ('Mary Mary', for short).

Mary Mary, like Mary, has lived in her mother's black and white environment and inherited her mother's physical scientific omniscience as well as monochromaticism. However unlike mother, Mary Mary has read Michael Tye's *Ten Problems of Consciousness*. As a result, she is convinced that PANIC theory is correct. Moreover, her metaphysical commitment has passed through dialectical fire. One of her email correspondents, for example, claimed that he could imagine two possible worlds which are exactly alike PANIC-ly but which differ in some introspective aspect (i.e. in some aspect capturable *only* by phenomenal concepts wielded by those having the relevant experiences). For instance, said the critic, he can imagine a physically possible world which is PANIC-ly just like ours except that persons who are our counterparts in that possible world have color experiences that are systematically inverted with respect to our own experiences; they are victims of what he called 'Metaphysical PANIC Disorder'. If there really are such possible worlds or possible victims, charged the critic, doesn't this mean that PANIC theory is false?

Mary Mary used her exposure to this criticism and others like it to refine her convictions, and therein she heightened her appreciation of Tye's theory. She claimed that her critic was using a metaphysically uninformed concept of first-person experience, and that we can't use such untutored concepts in acts of imagination to extract ontological morals. Although we can indeed *conceive* of the felt quality Q of an experience being instantiated even when the experience lacks the relevant PANIC property P (the broadly physical property to which Q is identical, on Tye's theory), Mary Mary claimed that the inference from conceivability to metaphysical possibility is unwarranted here – and thus that we cannot really *imagine* such a scenario, in the sense of imaginability that entails metaphysical possibility. As Tye himself puts the point,

Consider . . . the claim that Q can be imagined without the specified PANIC. Why do we naturally suppose that this claim is true? The answer surely is that in thinking of a state as having phenomenal character Q, we think of it via the exercise of phenomenal concepts. The conception we have is the one we have when we introspect it. By contrast, in thinking of a state as having PANIC P, we

bring to bear a very different set of concepts. So, we naturally infer that we are thinking of two different things: *Q* and *P*. But this inference is unjustified. The difference in thoughts can be accounted for solely by a difference in concepts. The one thought consists in representing the given PANIC under phenomenal concepts; the other consists in representing the same PANIC under nonphenomenal concepts. . . . Once these points are appreciated, there seems to be no reason left to insist that we really can imagine *Q* without the given PANIC. . . . Conceptual possibility yields apparent imaginability but not imaginability simpliciter. Some things that are conceptually possible are metaphysically impossible and hence unimaginable. (pp. 189–190)

“Just so!” says Mary Mary. She is more than an omniscient physical scientist living monochromatically. She is a True Believer. She has faith in thick materialism, and in Tye’s version thereof.

No thick materialist questions the ability of the physicalist story to accommodate *some* epistemic contrast between Mary pre-release and post-release. However, we see (or think we see) a deep challenge for the thick materialist in giving a plausible account of what happens to Mary Mary – Mary’s daughter – when *she* leaves her black and white environment and becomes polychromatic. Let us explain.

When Mary Mary contemplates a post-monochromatic mental life for herself, what changes can she reasonably expect? She expects to undergo certain visually generated PANIC states she has never undergone before. Since she thoroughly understands the functional-representational role of these states, she expects them to play such a role in herself. In particular, she expects to acquire, on the basis of these PANIC states, certain new discriminatory and recognitional capacities vis-a-vis physical colors – those very capacities which, according to Tye’s theory, constitute the possession of the relevant phenomenal concepts. She expects to have new beliefs and new knowledge-states that employ these capacity-based concepts – including “knowing what it’s like” states, in which the newly tokened PANIC properties themselves are represented under the newly acquired phenomenal concepts.

But should Mary Mary, while still in her monochromatic situation, expect to be *surprised* by the new experiences and new knowledge she would acquire upon beginning to have color experiences? Should she expect unanticipated *delight* at the new experiences and new knowledge, over and above any anticipated delight

that she might expect to arise purely from the acquisition of new discriminatory and recognitional capacities *per se*? No, she should not. For, given Tye's theory, essentially all she is missing is those discriminatory/recognitional capacities *themselves*; and she thoroughly understands these already, even though she does not currently possess them.

To make this clear, let us consider in turn (i) phenomenal properties (PANIC properties, according to Tye), (ii) the cognitive capacities that constitute (for Tye) the possession of the phenomenal concepts, and (iii) the beliefs and knowledge-states that deploy these capacity-based concepts. First, what is psychologically significant about the PANIC properties is just the functional/representational role they play in human cognitive economy – something that Mary Mary thoroughly understands already, by virtue of her scientific omniscience. Their functional/representational role involves the various behavioral, discriminatory, recognitional, and classificational capacities they subserve – including the capacities that constitute (according to Tye) the possession of phenomenal concepts. No expected surprises there.

Second, what is psychologically significant about phenomenal concepts (given Tye's theory) is that that they are *capacity-based* concepts; i.e., they are constituted by the relevant discriminatory/recognitional capacities vis-a-vis red things, vis-a-vis the physical property of redness, and (in the case of introspective application) vis-a-vis redness-representing PANIC states themselves. But she already understands these capacities thoroughly, including how PANIC states subserve them, even though she does not possess the capacities herself. No expected surprises there, either.

Third, the psychological distinctiveness of beliefs and knowledge-states employing phenomenal concepts is completely parasitic (given Tye's theory) upon the capacity-based nature of the phenomenal concepts. So she already understands well the *nature* of these beliefs and knowledge-states, even though she is not yet capable of undergoing them (because she lacks the requisite discriminatory/recognitional capacities, and therefore lacks the capacity-based concepts). Thus, since she should expect no surprises from undergoing the PANIC states themselves, or from acquiring and deploying the PANIC-based discriminatory/recognitional

capacities that allegedly constitute the possession of phenomenal concepts, she also should expect no surprises from the relevant knowledge states that constitute (according to Tye) “knowing what it’s like” to undergo color experiences. These are *capacity-based* knowledge-states, after all, and she already understands the capacities even though she does not yet possess them.

So Mary Mary, as a True Believer in Tye’s PANIC theory of phenomenal consciousness, has no good reason to expect surprise or unanticipated delight upon being released from her monochromatic situation. Let us suppose, then, that since she is rational on this matter, she has no such expectations. But what should be her reaction upon leaving her black and white environment? Surely, we submit, she should be both surprised and delighted. Why? She will not be surprised because she exercises new and direct discriminatory/recognition capacities. This was promised by Tye’s theory. She will not be surprised that some of these capacities are predicative and that some are indexical; that some can be used to sort or classify but others cannot; that her discriminatory capacity with respect to currently-presented colors is far more fine grained than her ability to classify or to use merely predicative phenomenal concepts (for here, again, Tye’s theory was on target).

What will surprise and delight Mary Mary, what will grab her – or should, given her allegiance to thick materialism and to Tye – is (it seems to us) the unanticipated *experiential basis* of her concept-wielding, recognition/discriminatory, capacities, and the acknowledged richness of her experience; she never expected polychromatic experience to be like *this*. It will seem to her that this surprising, and unexpectedly delightful, aspect of her experience is both (i) distinct from the recognition/discriminatory capacities for which it is the experiential basis, and (ii) something that has a quite distinctive character *qua* mental, over and above the representational/functional PANIC role that she knows about already.

So Tye’s theory should now strike Mary Mary as quite mistaken, precisely because it apparently leaves out this distinctive, and distinctively mental, character of her visual experiences. Even though (let us grant) her experiential states do indeed *have* a representational/functional role of the PANIC type, it should now seem clear to her that there is more to them, *qua* mental, than that.

There is also their *phenomenal character*, something evidently quite distinct from the various behavioral, discriminatory, recognitional, and classificational capacities that jointly make up the PANIC role. Although phenomenal states may indeed play a PANIC role in human psychological economy, their phenomenal character is not reducible to that role. It is something more, something surprising and delightful. It's what Mary Mary was not expecting.

Since phenomenal concepts derive from direct experiential acquaintance with phenomenal character, Mary Mary also should conclude that Tye was wrong in claiming that possession of phenomenal *concepts* consists merely in certain recognitional/discriminatory capacities. Although such capacities may well be *part of* what constitutes the possession of phenomenal concepts, they are not the whole. There is also the direct acquaintance with phenomenal character itself, acquaintance that provides the experiential basis for those recognitional/discriminatory capacities. Mary Mary should further conclude that there also is more to "knowing what it's like" to have color-experiences than Tye's theory allows for. Since phenomenal concepts are not mere capacity-based concepts, knowing what it's like is more than a matter of introspectively applying predicative-recognitional capacities and indexical-discriminatory capacities. To wield phenomenal concepts vis-a-vis one's own experiential states, i.e., to know what those states are like, is like ... phenomenal! – where 'phenomenal' stands for the direct experiential acquaintance that is the basis for her new post-release way of knowing. Knowing what it's like is *more* than a merely capacity-based knowledge-state that involves mere capacity-based predicative and indexical concepts. Its greater richness is what is surprising and delightful about it, and Tye's theory leaves this out.

The general line of worry here looks generalizable to us. Take any thick materialist story about phenomenal experience, of the sort that purports to take the progressive contrast between pre-release and post-release Mary seriously. Presumably Mary Mary (Mary's daughter) can understand the materialist story, and can become a believer in this story ahead of time (before release). The story will offer an account of the nature of (i) phenomenal properties, (ii) phenomenal concepts, and (iii) beliefs and knowledge-states employing phenomenal concepts. Pre-release Mary Mary will

thoroughly understand the nature of all three (according to the given story), even though she does not possess the phenomenal concepts herself and hence cannot form beliefs employing these concepts. Given this understanding, it seems likely that acceptance of the materialist story will leave Mary Mary with no rational basis to expect either surprise or unanticipated delight from her post-release experiences. (This is especially likely if the proffered account treats phenomenal concepts as capacity-based concepts of some sort. It is hard to see how a materialist can *avoid* doing so; witness the current popularity, among thick materialists, of the idea that phenomenal concepts are a species of so-called “recognitional” concepts.) But she *will* experience surprise and unanticipated delight, upon release from her monochrome environment – which presumably should lead her to repudiate the materialist theory she previously accepted.

We do not necessarily think that this line of worry shows the falsity of materialism. But it appears powerful to us, and as yet we have not conceived of a materialist rejoinder that we find adequate. It seems that Mary Mary should be surprised on release, which should make it obvious to her that PANIC theory (or whatever materialist theory she happens to accept) leaves something out, viz., knowing what what it’s like *is like*.

THE MARY MARY CHALLENGE: INVERSE VERSION

Essentially the same challenge to PANIC theory, and to materialist accounts generally, can be formulated by an inverted form of the Mary Mary thought experiment. In this version, pre-release Mary Mary attends closely to her monochromatic, gray-tone, color-phenomenology. She forms the belief that what it would be like to have full-fledged color-experiences is something radically different from what her current monochromatic experience is like – so different that she is dying to find out what it is like, and so different that she expects her post-confinement experiences to be both surprising and intrinsically delightful. In addition, let us suppose, she cares very little about acquiring the recognitional/discriminatory capacities that allegedly constitute, according to Tye’s PANIC theory, the possession of phenomenal color-concepts; for, she is already quite adept at using her instruments to make coarse-grained

and fine-grained color discriminations and predications, and at using her autocerbroscope to identify and classify her own brain-states. Well then, it should seem to her that the new knowledge she is dying to acquire cannot simply involve the introspective application of *capacity-based* concepts, since that kind of knowledge is not what she is dying to acquire. Tye's theory of phenomenal consciousness should therefore strike her as wrong, since it reduces knowledge of what color experiences are like to a capacity-based cognitive state that she thoroughly understands already and has no particular interest in acquiring. *Real* "knowing what it's like" is richer and more exciting than that, because real phenomenal content is something over and above PANIC.

Once again, this line of worry looks generalizable to other versions of thick materialism. Mary Mary is dying to know what it's like to experience colors, and she expects this new knowledge to be both surprising and intrinsically delightful. Yet any materialist account of phenomenal properties, phenomenal concepts, and phenomenal knowledge will be fully understandable by Mary Mary in her pre-release state. She will understand well the nature (according to the given account) both of phenomenal concepts themselves, and of the introspective knowledge that involves applying these concepts to broadly physical states and properties. (She will understand these things even though she lacks them.) But then *those* kinds of concepts, and *that* kind of knowledge, will not be concepts and knowledge that she is dying to acquire, precisely because she understands them so well already. So she should conclude, it seems, that the given materialist story leaves out the real phenomenal essence of color experiences, viz., what they are really *like*. That's what she's dying to know.

MATERIALIST REJOINDERS

The default moral of our two versions of the Mary Mary challenge is that phenomenal properties are not broadly physical properties: they are neither identical to, nor conceptually supervenient upon, properties of the kind posited by physics and by the other natural sciences. Rather, they are non-physical properties with which one becomes directly acquainted in experience. Phenomenal concepts

and phenomenal knowledge are grounded in this direct acquaintance. Phenomenal consciousness, and the knowledge of what is like, are surprising and unexpectedly delightful to Mary Mary because they involve acquaintance with genuinely new, non-physical, properties.

Thick materialists will resist this default moral, of course. They will seek to provide a rejoinder to the Mary Mary thought experiment that accommodates its intuitive force without jettisoning materialism. We will consider several potential rejoinders. None of these is persuasive, in our view, but all are philosophically instructive.⁶

One rejoinder consists of a claim and an inference. The claim is that Mary Mary is actually in a position to predict in advance, on the basis of non-subjective scientific information available to her in principle, that she will be surprised by her post-release color experiences and will find them intrinsically delightful. The inference is from this claim to the conclusion that such post-release surprise and delight should not count as grounds for rejecting PANIC theory (or whatever other version of materialism she might initially accept).

The claim itself should be conceded, we think. Grant that Mary Mary's scientific knowledge includes information about how mental state-types like surprise and delight are neurophysically realized in humans. Grant too that in principle, Mary Mary could use her exhaustive neurophysical knowledge to predict in advance any post-release changes in her own neurophysical processes that would be relevant to the realization of psychological states like surprise and delight. If she were to draw upon this *wide* base of non-subjective information that is available to her in principle, rather than confining herself to Tye's theory of phenomenal content and to the workings of the color-vision system alone, then in principle she could indeed predict that she would find her post-release experiences surprising and intrinsically delightful.⁷

The problem, however, lies in the inference from this fact to the conclusion that the Mary Mary challenge poses no genuine problem for PANIC theory, or for materialism more generally. On the contrary, the problem remains. The real challenge is to give a satisfying, phenomenology-respecting, account of why it is *rationally appropriate* for Mary Mary to find her post-release color

experiences surprising and intrinsically delightful. She claims to be delighted by a distinctively *mental* aspect of her new experience – an aspect over and above her new recognitional-predicative and indexical-discriminatory capacities, and over and above any new thought-contents whose novelty consists merely in the exercise of new capacity-based concepts. Auto-phenomenology suggests strongly, *very* strongly, that she is right about this: the intrinsic phenomenal character of color experience is distinct from, and provides the basis for, these recognitional/discriminatory capacities. Yet on Tye's story, there simply *isn't* any distinctive, surprising, intrinsically delightful, mental aspect of her experience, over and above both (i) new cognitive capacities, and (ii) new thought-contents whose novelty consists entirely in the exercise of those capacities. So on his story, Mary Mary's post-release hetero-phenomenological claims evidently must be viewed as rationally inappropriate, and thus as embodying some kind of error or illusion. *That* is the basic problem: the apparent failure to provide adequate theoretical accommodation for the manifest phenomenological facts.

So the in-principle predictability of surprise-reactions and delight-reactions, on the basis of a sufficiently broad base of non-subjective information that is admittedly available in principle to Mary Mary, is beside the point. The real trouble is that such reactions evidently must be treated as illusions or as errors, under Tye's theory – which goes contrary to his own intention of fully accommodating the phenomenological character of conscious experience. In the end, Tye's version of thick materialism is just *too thin*. And this problem threatens to arise for any materialist treatment of phenomenal content.

We turn now to a second rejoinder to the Mary Mary problem, inspired by certain remarks in Tye's book and also by points he has made in discussion. It goes as follows. Three subtle and important points need to be appreciated well. First, contexts like "It is surprising that . . .", "It is amazing that . . .," and ". . . is unexpectedly delightful" are all *intensional* contexts, and hence do not admit free substitution of co-referring singular terms or predicates *salva veritate*. One must beware of intensional fallacies, when arguing for

a metaphysical-dualist conclusion by appeal to notions like surprise, amazement, or unexpected delight.

Second, the capacity-based concepts that are identical, under Tye's theory, to phenomenal concepts are very different from certain non-subjective, theoretical concepts involving the same subject matter. Specifically, phenomenal concepts are very different from the concept of a PANIC state – a concept that Mary Mary herself already possesses and has mastered while in her pre-release monochromatic situation. Phenomenal concepts are also very different from the higher-order, nonsubjective, theoretical concept of a phenomenal *concept*, i.e., the concept of a certain kind of capacity-based concept involving either recognitional-predicative capacities or discriminatory-indexical capacities. Pre-release Mary Mary also possesses these higher-order concepts: she understands the nature of phenomenal concepts, even though she does not possess them herself (because she lacks the requisite capacities).

Third, the differences just mentioned render phenomenal concepts independent of – *conceptually* independent of – these other kinds of concepts. Someone could possess and correctly employ phenomenal concepts without believing that the properties they pick out (when introspectively applied) are PANIC properties, without realizing that these concepts are capacity-based concepts, and indeed without even *possessing* the concept of a PANIC property or the concept of a capacity-based concept. Conversely, someone could possess the latter concepts without possessing phenomenal concepts themselves; this is precisely Mary Mary's own situation, pre-release.

If these three points are kept clearly in mind, the rejoinder goes, then one will realize that Mary Mary's post-release reactions of surprise, amazement, and unexpected delight simply do not entail that she has become acquainted with new, non-physical, *properties*. On the contrary, these reactions are entirely compatible with Tye's theory, given the intensionality of notions like surprise, amazement, and delight, and given also that phenomenal concepts are so different from the theoretical concepts Tye employs that the former are actually conceptually independent of the latter. As Tye himself remarks, using 'fact' for fine-grained facts and 'FACT' for coarse-grained facts (a distinction explained above):

Consider, for example, the feeling of elation. Suppose you have an autocerebro-scope attached to your head at the time at which you feel elated and that you are yourself viewing, in the attached mirror, the particular firing pattern that constitutes your feeling. Is it not going to seem absolutely amazing to you that the brain state you are viewing feels the way it does, indeed that it feels any way at all? . . . Suppose now that you are supplied with the pertinent information about the role of the brain state and what it represents. Imagine that the information is flashed onto a screen placed before you. By reading what is on the screen, you discern the naturalized PANIC of the brain state you have been seeing via the autocerebro-scope. Is it *still* not going to seem absolutely amazing to you that a state with that PANIC feels the way your present state does? Is it *still* not going to seem absolutely amazing that the brain state feels any way at all? Why, yes. But so what? There are facts, and there are FACTS. The former are partly conceptual, the latter are not. . . . The concepts you apply in the two cases are very different. In the one case, the concepts are purely phenomenal; in the other, they are not. The fact you know via introspection is, therefore, very different from fact you know as you read the screen. The one is not deducible from the other. . . . So *of course* it is amazing. . . . But there is still no reason to suppose that there are two different FACTS here rather only one FACT under different modes of presentation. (pp. 178–179).

Furthermore, the rejoinder continues, Tye's PANIC theory is also compatible with statements like the following, made by Mary Mary in her post-release situation:

- (1) What is surprising and unexpectedly delightful is not the fact that I now am undergoing new visually induced PANIC properties, and is not my newly acquired recognitional-discriminatory capacities; rather, it is something that is independent of those capacities themselves.

This statement can actually be *true*, given PANIC theory. It can be true because (i) the notions of surprise and unexpected delight – and also the relevant notion of independence, viz., conceptual independence – create intensional contexts, and (ii) phenomenal concepts are significantly different from, and indeed are conceptually independent of, the concept of a PANIC property and the concept of a capacity-based concept. Given facts (i) and (ii), the following statements are all consistent, under Tye's theory. (Intensional contexts are indicated by braces; and key referring expressions within such contexts that cannot be replaced *salva veritate* by co-referring expressions are indicated by brackets.)

- (2) It is surprising and unexpectedly delightful that {my new experiences exhibit [the property *red*]}
- (3) It is not surprising and unexpectedly delightful that {my new experiences exhibit [such-and-such PANIC-property]}.
- (4) {That my new experiences exhibit [the property *red*]} is independent of my recognitional-discriminatory capacities with respect to red things.
- (5) {That my new experiences exhibit [the property picked out by so-and-so capacity-based concept]} is not independent of my recognitional-discriminatory capacities with respect to red things.

(Here, '*red*' expresses a phenomenal predicative concept; 'such-and-such PANIC property' is shorthand for a theoretical description of the specific PANIC property that is identical, under Tye's theory, to the phenomenal character of visual experiences of the color red; and 'so-and-so capacity-based concept' is shorthand for a theoretical description of the recognitional-discriminatory capacities whose possession constitutes, under Tye's theory, possession of the phenomenal concept *red*.) As the braces and brackets make clear, statements (2)–(5) each can all be true even if the bracketed expressions within them are all co-referential. Hence, Mary Mary's statement (1) too can be true too, given Tye's theory. To maintain that statement (1) must be false, given Tye's theory, is to commit an intensional fallacy.

We are not persuaded by this rejoinder to the Mary-Mary problem. We do grant the three claims that form its basis: that notions like surprise, delight, and amazement create intensional contexts; and that on Tye's account, phenomenal concepts are significantly different from, and are independent of, non-subjective theoretical concepts like the concept of a PANIC property and the concept of a capacity-based concept. We grant, also, that amazement, surprise, and delight about how one's inner states feel is consistent with Tye's theory. We grant too that under Tye's account, statements (1)–(5) are all consistent, and that to claim otherwise is to commit an intensional fallacy. But even with all of this conceded – as indeed it should be – we submit that the basic problem remains.

The real problem is not to demonstrate the mere *consistency* of Mary Mary's post-release state of mind, and of statements like (1)–(5), with Tye's theory. Rather, it is to give a satisfactory, phenomenology-respecting, account of the *rational appropriateness* of her post-release state of mind and of the *truth* of all of state-

ments (1)–(5) – an account that is materialistically kosher, and that does not end up treating her post-release attitude as rationally inappropriate (and thus as some kind of error or illusion).

To make this clear, let us suppose that pre-release Mary Mary is well aware of all the points lately made about the intensionality of notions like surprise, amazement, and unexpected delight, and about the consistency of statements like (1)–(5) with Tye's position. (She *should* be, if she understands Tye's theory sufficiently well.) She certainly understands that the capacity-based phenomenal concepts that she currently lacks – for instance, the predicative and indexical phenomenal concepts associated with visual experiences of red – are different from, and are conceptually independent of, certain theoretical concepts she already possesses – for instance, the concept *such-and-such PANIC property* and the concept *so-and-so capacity-based concept*.

Well, what should Mary Mary, paragon of rationality that she is, rationally expect about her post-release experiences, given all this knowledge? She should expect to acquire new recognitional-discriminative capacities, vis-a-vis red things and also introspectively vis-a-vis her own internal experiences of red things. She should expect to acquire new concepts, constituted by certain such recognitional-discriminatory capacities. She should expect to undergo new kinds of beliefs and knowledge-states that she could not undergo before, whose novelty involves the exercise of her new capacity-based concepts (and whose novelty thus is ultimately a matter of her newly acquired recognitional capacities). She should expect her post-release phenomenal concepts, once acquired, to be conceptually independent of her pre-release theoretical concepts pertaining to color and to phenomenal consciousness. (For instance, if her initial post-release experiences are of bare color patches, without any collateral visual or non-visual information, then she will not yet know which of her newly acquired phenomenal concepts match up with which objective colors, or with which PANIC states, or with which objectively described recognitional-discriminatory capacities.⁸) Given this expected conceptual independence, she also should expect the possibility of being surprised in certain ways. (For instance, she can expect that if her initial post-release experiences are of bare color patches, without any collateral visual or non-visual

information, then she might be surprised when she subsequently finds out how her newly acquired phenomenal concepts match up with which objective colors.)

None of this pre-release knowledge, however, provides Mary Mary with any apparent reason to expect the additional and extreme surprise, the unanticipated delight, or the utter amazement that lie in store for her.⁹ Indeed, insofar as she accepts Tye's theory, she *has* no apparent reasons for such expectations. And that is precisely the problem, the *real* problem. An adequate theory of phenomenal consciousness should provide a theoretical rationale for such expectations, since there can be no serious doubt that those post-release reactions do indeed lie in store. Yet PANIC theory evidently fails to provide such a rationale, and other versions of materialism seem to fare no better.

A third potential rejoinder to the Mary Mary problem, perhaps, is to embrace a variant of the Wages of McGinn (cf. McGinn, 1991). It is to claim that though we may be confident, with Mary Mary, that qualia are physical or PANIC properties, we don't seem to have an explanation of why the relevant experience of color should appropriately give rise to surprise in a confirmed and previously black and white confined materialist. We literally are incapable of understanding as materialists why Mary Mary's reaction on release is rationally appropriate (and Mary Mary qua materialist suffers from this same theoretical incapacity, too, in some sense). Still, *that* isn't enough justification for saying that her reaction reveals the falsity of the physicalist story. *Somehow* it is consistent with materialism.

Physicalists who advocate the epistemic boundedness of much human cognitive capacity may find this Colin-McGinnish reply congenial. However given the attempt of someone like Tye to explicate just how physical or PANIC properties are known in first-person experience, the position that there is something fundamentally inexplicable, materially, about the experience should leave philosophers fretting over whether such an 'explicability gap' is consistent with materialism. Doesn't this mean that materialists such as Tye really don't understand what is going on in phenomenal experience, and doesn't *that* mean that they are not entitled to their materialism?

Other potential materialist replies? Perhaps thick materialism is untenable and should be replaced with the thin variety.¹⁰ We have no quarrel, of course, with trying other replies, but we are skeptical of this one. To embrace thin materialism diminishes the appeal of materialism itself; it abandons an attraction of thick materialism, which is its conviction that Mary (the mother now) (and Mary Mary for that matter) makes genuine epistemic progress after release. Numerous materialists have claimed that Mary makes progress (i.e. learns *something* new). It seems desperate to deny this.

A final ploy. One could aim to respect Mary Mary's epistemic progress in a radically different way. This is to claim that pre-release Mary Mary can't *really* understand thick materialism/PANIC theory, since having direct experience of color is requisite for comprehending materialist metaphysics. Mary Mary's surprise on release is not due to confronting new, non-physical, properties. Rather, hers is semantic startle, in virtue of finding out for the first time what the materialist theory she accepts really means. She realizes post-release that the correct materialist theory of phenomenal consciousness is a theory that one cannot even understand prior to undergoing the relevant phenomenal states oneself.

Well, now we are walking deep into dark metaphilosophical woods, with worries looming about whether materialist metaphysics is something that *could* require taking a first person/colored world perspective so seriously that the metaphysics itself cannot be grasped without some sort of direct encounter with color. This materialist story would reject the notion that direct experience of color reveals some property over and above the physical; but here the story itself cannot be understood in a confined environment. This is a bold but also utterly bizarre idea, since materialism, as traditionally conceived, rejects anything remotely like it. When one adopts physicalism, J. J. C. Smart writes:

One tends to get a certain way of looking at the universe, which is to see it *sub specie aeternitatis*. . . . To see the world *sub specie aeternitatis* . . . is to see it apart from any particular or human perspective. (Smart, 1987, p. 33)¹¹

To see the world under the form of eternity is to see it impersonally (third-personally, non-subjectively). It is precisely *not* to adopt, or feel any need to adopt, a colored world perspective. One can therefore begin to see how embracing materialism is incompatible with

saying that one cannot understand PANIC theory unless one – as it were – lives in a colored environment; unless, that is to say, one lives *here* and not *there* (in pre-release).

CONCLUSION

We may be riding rather hastily over the description of Mary Mary's epistemic reaction on release. Perhaps the relevant reaction is not surprise *per se* but some conceptual kin (e.g. being caught off one's epistemic guard); or perhaps, given possible idiosyncracies in her personality, her reaction to directly recognizing color – to being caught by it, given her expectations – is suffused more with unparalleled delight and unanticipated relish than with surprise. However, the main point of the Mary Mary Challenge is that thick materialism appears to conflict with what is rationally appropriate for her to expect, given her acceptance of a materialist account of conscious experience. Although she does have reason, given the conceptual independence of phenomenal concepts from the relevant theoretical concepts she already possesses, to expect the possibility of being surprised by the specific *correlations* between her newly acquired phenomenal concepts and those theoretical concepts, she has no rational basis to expect to be surprised by the phenomenal correlates *themselves*. But of course she will be.

If the Mary Mary Challenge is indeed the difficulty which we have made it out to be, then the unwelcome phenomenal properties of the traditional Mary thought experiment may not be eliminable, as the thick materialist would have it, by reference to perspectival concepts that allegedly apply to broadly physical properties. Historians of the twentieth century intuition pump may therein better appreciate this thought experiment's force.

Mary Mary Quite Contrary, how *does* your knowledge grow? By experiencing qualia all in a row?

ACKNOWLEDGEMENTS

This is a thoroughly co-authored paper; the order of authorship is alphabetical. Earlier versions were presented at the 1999 Pacific

Division meeting of the American Philosophical Association and the 1999 meeting of the Society for Philosophy and Psychology. Among those who in one form, forum or another offered helpful comments are Fred Dretske (the APA commentator), David Chalmers, Stevan Harnad, Patrick Hayes, Justin Leiber, Victoria McGeer, Andrew Melnyk, John Tienson, Mark Timmons, and Michael Tye. We thank them. We also thank the program chairs, Elizabeth Radcliffe and Guven Guzeldere respectively, for inviting us to present the paper at those meetings, and we thank Michael Tye for organizing the APA session and encouraging us to write on this specific topic.

NOTES

- ¹ Dennett (1984) dubs philosophical thought experiments intuition pumps.
- ² We shall use ‘physicalist’ (‘physicalism’) and ‘materialist’ (‘materialism’) interchangeably.
- ³ No doubt hers may also be a grayish environment, a descriptive wrinkle unnecessary to explicitly accommodate for present purposes.
- ⁴ It is sometimes objected that an adequately redesigned version of the thought experiment needs to be informed by actual scientific work on color and color-experience, and that such work may make it impossible to “conduct” the thought experiment without making so many contrary-to-fact assumptions that the full-fledged conceptual possibility of the scenario becomes doubtful. Those who have such doubts would do well to consult Shepard (1993), a paper that arguably constitutes an existence-proof that everything that science now knows about color and color-experience is consistent with the Mary thought experiment (or ones like it). Shepard independently introduces a thought experiment much like Jackson’s, and he discusses its apparent implications with admirable subtlety and sophistication.
- ⁵ Take the imaginability strategy, for example. Endowing Mary with stunning powers of imagination does not rule out knowledge of color experience as something over and above her neuro-omniscience. Her ‘imagined’ colors may be qualia. Perhaps her stunning powers of imagination consist of being presented with something over and above her neuroscientific knowledge.
- ⁶ For simplicity, we will confine discussion to our original Mary Mary thought experiment, and will ignore the inverse version. Both the rejoinders and our replies can be adapted, *mutatis mutandis*, to the inverse version.
- ⁷ Or, if auto-prediction of epistemic states like surprise presents logical problems insofar as Mary Mary cannot both know that she will be surprised in a certain way and be surprised in that way, then *another* know-it-all confined scientist could have made such predictions about Mary Mary.
- ⁸ By our lights, this kind of lack of knowledge is the appropriate analog of a self-

referential indexical case that Tye uses to illustrate how one can lack knowledge a fact without lacking knowledge of any corresponding FACT. He writes:

Suppose I am looking across the room at a group of people, and I briefly see someone who looks just like me. I take special note of the fact that he is wearing my old school tie. Unknown to me, I have actually caught a glance of myself in a mirror in the vicinity of the group of people. I know that *that* person, the one I am seeing, is wearing my old school tie. But I do not know that *I* am wearing such a tie, for I paid no attention at all to the kind of tie I put on. There really is here a single FACT, the FACT that consists of Michael Tye's wearing a certain sort of tie. . . . [T]here really is a relevant fact here that I do not know. But this fact is not a FACT that I do not know. (pp. 176–177)

Similarly, Mary Mary in her pre-release state already knows that if her initial post-release experiences are of bare color patches, then she will not yet know certain correlational facts linking her new phenomenal concepts with her prior theoretical concepts. As an advocate of Tye's theory, she believes that these not-yet-known facts will involve no new FACTS beyond those she knows about already, but instead will involve conceptually independent concepts that are coreferential.

⁹ As should now be clear, there is not an appropriate analogy between her extreme surprise and the sort of surprise that Michael Tye might express by saying "Good grief, *I* am wearing my old school tie!" Although post-release Mary Mary may indeed be moderately, expectably, surprised by correlational facts linking her newly acquired phenomenal concepts to her previous theoretical concepts, what surprises her extremely and unexpectedly is the intrinsic nature of phenomenal states themselves. Tye's surprise at the tie he is wearing is analogous to the former kind of surprise in Mary Mary, not the latter.

¹⁰ Perhaps thick materialism seems pretty thin anyway – not sufficiently different from the second sort, the Nemirow/Lewis sort, of thin materialism. Parsing Nemirow-Lewis materialism is beyond the scope of the present paper.

¹¹ Lest there be any textual confusion, actually what Smart is talking about in this quote is the eternal (impersonal, non-indexical) perspective of physical science, which he sensibly equates with the perspective taken by a materialist.

REFERENCES

- Churchland, Paul (1989): 'Knowing Qualia: A Reply to Jackson', in *A Neuro-computational Perspective: The Nature of Mind and the Structure of Science*, pp. 67–76, Cambridge, MA: MIT.
- Dennett, Daniel (1984): *Elbow Room*, Cambridge, MA: MIT.
- Dennett, Daniel (1991): *Consciousness Explained*, Boston: Little, Brown, & Co.
- Dretske, Fred (1995): *Naturalizing the Mind*, Cambridge, MA: MIT.
- Harman, Gilbert (1990): 'The Intrinsic Quality of Experience', in J. Tomberlin (ed.), *Philosophical Perspectives, 4: Action Theory and the Philosophy of Mind*, Atascadero, CA: Ridgeview.

- Horgan, Terence (1984): 'Jackson on Physical Information and Qualia', *Philosophical Quarterly* 34, 147–152.
- Jackson, Frank (1982): 'Epiphenomenal Qualia', *Philosophical Quarterly* 32, 127–136.
- Jackson, Frank (1986): 'What Mary Didn't Know', *Journal of Philosophy* 83, 291–295.
- Jackson, Frank (1998): *From Metaphysics to Ethics*, Oxford: Clarendon.
- Levine, Joseph (1997): 'Recent Work on Consciousness', *American Philosophical Quarterly* 34, 379–404.
- Lewis, David (1990): 'What Experience Teaches', in W. Lycan (ed.), *Mind and Cognition*, Oxford: Blackwell.
- Lycan, William (1996): *Consciousness and Experience*, Cambridge, MA: MIT.
- McConnell, Jeff (1994): 'In Defense of the Knowledge Argument', *Philosophical Topics* 22, 157–187.
- McGinn, Colin (1991): *The Problem of Consciousness*, Oxford: Blackwell.
- Nagel, Thomas (1974): 'What Is It Like to be a Bat?', *Philosophical Review* 82, 435–450.
- Nemirow, Lawrence (1980): 'Review of Thomas Nagel, *Moral Questions*', *Philosophical Review* 89, 473–477.
- Nida-Rumelin, Martine (1995): 'What Mary Couldn't Know: Belief about Phenomenal States', in T. Metzinger (ed.), *Conscious Experience* Paderborn: Ferdinand Schöningh.
- Shepard, Roger (1993): 'On the Physical Basis, Linguistic Representation, and Conscious Experience of Colors', in G. Harman (ed.), *Conceptions of the Mind: Essays in honor of George A. Miller*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thompson, Evan (1995): *Colour Vision*, London: Routledge.
- Tye, Michael (1995): *Ten Problems of Consciousness*, Cambridge, MA: MIT.
- White, Stephen (1994): 'Color and Notional Content', *Philosophical Topics* (Spring/Fall), 107–126.

GEORGE GRAHAM
Department of Philosophy
 HUM 414A
University of Alabama at
Birmingham
Birmingham, AL 35294-1260 USA
E-mail: ggraham@uab.edu

TERENCE HORGAN
Department of Philosophy
University of Memphis
Memphis, TN 38152 USA
E-mail: thorgan@memphis.edu

