# There's Something About Mary

## Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument

edited by Peter Ludlow, Yujin Nagasawa, and Daniel Stoljar

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

# 6 Motion Blindness and the Knowledge Argument

**Philip Pettit**

## Introduction

In a now famous thought experiment, Frank Jackson asked us to imagine an omniscient scientist, Mary, who is confined in a black-and-white room and then released into the world of color (Jackson 1982; Jackson 1986; cf. Braddon-Mitchell and Jackson 1996). Assuming that she is omniscient in respect of all physical facts—roughly, all the facts available to physics and all the facts that they in turn fix or determine—physicalism would suggest that there is no new fact Mary can discover after emancipation; physicalism holds that all facts are physical in the relevant sense (for a fuller statement see Pettit 1993; Jackson 1998). Yet we cannot help but feel that coming out of that room would be an occasion of dramatic enlightenment and, in particular, an occasion for learning facts to do with how red or yellow or blue looks or, as it is usually said, with what it is like to see red or yellow or blue.

Mary in the black-and-white room knew all the physical facts about the world, where these may be taken to include three sorts of color facts: objectual facts, as to what surface colors different objects have, assuming—as I shall do throughout—that colors are properties of objects; intentional facts, as to which colors different objects or apparent objects are represented as having in the subject's experience, rightly or wrongly; and nonintentional facts, about what such color experiences are like in their effects on subjects—whether they are comforting, or arousing, or whatever. But, according to the argument, Mary didn't know how any color looks or, equivalently, what color experience is like in itself, not just in its effects on subjects. This particular nonintentional fact about the quality of color experience—this phenomenal fact, as it is often described—she did not

know. And so it is said to follow that there are facts—phenomenal color facts—that are not physical in character.

The idea, then, is that for all that Mary's knowledge in the black-and-white room settles, determining which possibilities are realized in the actual world, it leaves open a crucial question. Assume that colors are physical properties in objects—say, spectral reflectances—as it will be natural for physicalists to assume. Mary's knowledge in the black-and-white room, so it is said, does not give her information on whether the actual world is one where red looks phenomenally this way or that way—looks the way red actually looks to us, perhaps, or looks the way green actually looks to us. That issue is resolved for Mary, it appears, only in the light of her color experience: it is only when she actually sees colors that she knows whether one or the other possibility is realized; both possibilities are epistemically in play up to that point. Thus her color experience outside the black-and-white room yields knowledge that is additional to the knowledge of all physical facts that Mary, by hypothesis, had in the room. The phenomenal fact about which it gives her knowledge, therefore, has to be a fact of a nonphysical kind.

One sort of response to this argument holds that Mary could not have known all the physical facts there are to know under the constraints of the black-and-white room: that, not having been exposed to color, she would have lacked the conceptual resources required for knowing all such facts (Harman 1990). The sort of response I favor, however, assumes that she did know all the physical facts in the black-and-white room and maintains instead that what happens when she leaves the room is not that she gains new knowledge but that she changes in some other knowledge-related way: in a broad sense she develops a new mode of knowing. One version of the mode-of-knowing approach, understood in this broad sense, is the know-how response, according to which Mary learns how to discriminate colored things from one another—in experience, in memory, and in imagination—and how to order them in terms of similarity (Lewis 1990). But other, familiar responses to the argument can be seen as versions of the approach too. Mary may be said to learn how to recognize an abstract sort of particular under a new aspect, one we naturally describe as 'the way red looks'; or how to think indexically—that is, in the 'here', 'now', and 'I' way—about a fact previously known nonindexically; or how to form color concepts



proper and to use them in expression of familiar facts; and so on (for a review, see Braddon-Mitchell and Jackson 1996).

My version of the mode-of-knowing approach develops from consideration of a case analogous to Mary's—the case of Eva—where the property to which the subject is initially blind is not color but motion. I try to argue that while Eva could have known all the physical facts in a room where motion-blindness is imposed—say, by stroboscopic lighting—she will undergo a change on leaving that room which is so massive that we are first inclined to say that she learns a new fact: the fact of what motion looks like or, equivalently, what it is like to see motion; and yet that it proves completely implausible, on reflection, to maintain that view. What we should maintain, rather, is that Eva comes to believe or know in what I describe as a practical, nonintellectual mode facts that she previously believed or knew only in a purely intellectual way. I think that this analysis of Eva's case is compelling and it should help us to be persuaded that a similar analysis applies in Mary's. There are differences between the cases, of course, deriving from the fact that whereas motion is a primary property that is available to many senses, color is a secondary property discerned by sight alone. But I hold that these are not sufficient to argue for a substantially different treatment of the two.

The essay is in three sections. In the first I analyze the case of Eva, trying to show how she can undergo a massive change on leaving the stroboscopic room and yet, plausibly, learn no new fact, and so learn no new nonphysical fact, about motion: for example, no new objectual fact about which objects are moving at what rate, in what direction; no new intentional fact about how motion vision represents such objects, rightly or wrongly, as moving; and no new nonintentional fact about what motion vision is like in its effects on subjects or in any other way. In the second, I consider and reject an antiphysicalist objection that, to the contrary, Eva will learn on leaving the stroboscopic room about what motion vision is like in itself, not just in its effects: that is, she will learn about phenomenal—visually phenomenal—motion. Then I go on to argue in the third section that although there is an important difference in the spontaneous intuitions we have about the two cases—this difference may explain why the Mary argument has proved so persuasive—it is not sufficient to block our accepting a similar story in the case of Mary.

## 1 Analyzing the Case of Eva

### The Case of Eva

Recent studies have identified akinetopsia or motion blindness of a kind akin to the phenomenon of color blindness (Zeki 1993, ch. 10). Someone who suffers this ailment sees motion in a jerky series of static images, as if in a stroboscopic light (Gazzaniga et al. 1998, p. 146). Thus a sufferer reports as follows on the difficulty of crossing a road: "When I'm looking at the car first, it seems far away. But then, when I want to cross the road, suddenly the car is very near" (Zeki 1993, p. 82). As our awareness of color blindness makes it easy for us to follow the story about Mary, and to have intuitions about it, so an awareness of motion blindness can serve us in the same way with regard to a parallel scenario.

Imagine that an omniscient scientist, Eva, knows all the physical facts there are to know in just the sense in which Mary does but is confined in a room—call it a stroboscopic room—where she has never been able to see phenomenal motion. For Eva, seeing bodies move has only ever meant seeing them occupy a temporally ordered, gappy succession of distinct positions. And moving her eyes or rotating her body, so we may presume, has only ever meant seeing the scene before her jump from one profile to another; the literature is unclear as to whether this is a feature of those diagnosed with akinetopsia. Eva will understand motion fully: for her, as for everybody else, motion is just temporally continuous change of place (or, if this is thought necessary, temporally continous change of place that has a suitable metaphysical grounding). Eva will be able to perceive motion nonvisually, and to have illusory, nonvisual experiences of motion, in virtue of having kinaesthetic sensations of her limbs and body moving and she will be able to perceive the motion of external objects nonvisually, by touching and tracking them. But she will be visually blind to motion; she will have absolutely no visual sense of what motion, as distinct from discontinuous change of place, involves. Thus, as Mary does not know what seeing red is like—what red looks like—Eva does not know what seeing motion is like or, equivalently, what motion looks like.

So how will things be for Eva when she leaves the stroboscopic room for the first time? Will the shift that occurs within her, as we intuitively imagine it, involve her learning any new fact? Or can we make good sense of it just on the assumption that she develops a new way of knowing old facts?



I shall argue that we can make good sense of the transformation in Eva on this latter assumption, so that there is no reason to think that her new visual access provides evidence of a nonphysical fact. The analogy with Mary suggests that a similar lesson will go through there, and I turn to the case of Mary in the last section.

### Objectual Facts About Motion

There are three candidate sorts of facts that Eva might be thought to learn on leaving the stroboscopic room. First, facts about the objectual properties of motion: facts to do with whether it is slow or fast, accelerating or decelerating, steady or jumpy, going in this direction or that, and so on. Second, facts about the intentional properties of the visual experience of motion. And third, facts about the nonintentional properties of that experience. Visual experience, here and throughout the essay, is meant to refer to visual representaton, which is registered or conceptualized as a representation—it is overtly a representation—for the subject. The intentional properties of such an experience will determine whether it is as of fast or slow motion, as of accelerating or decelerating motion, and so on. The nonintentional properties will determine whether it is soothing or exciting or threatening and so on.

I think it is clear that Eva will not learn any facts about the objectual properties of motion, since they are clearly physical and she would have known about them already in the stroboscopic room. She knows all the physical facts there, by hypothesis, so she knows that this or that object is moving fast or moving slowly, that it is accelerating or decelerating, and so on for the other objectual properties of its motion. Indeed, she has the capacity to register some of those objectual properties perceptually in the stroboscopic room, though perhaps only on the basis of nonvisual perception.

But though Eva will not learn any new facts about the objectual properties of this or that motion, it may well seem that she does. For there is something new, as we might put it, that such objectual facts are going to be like for her; it's just that there being something that those facts are like for her does not mean that she learns a new fact. Consider how things will be for Eva outside the stroboscopic room, as she surveys a moving object. She will not just know, on the basis of prior information about all physical facts, that it is moving and moving in this or that objectual way. Nor will

she be caused on just any old perceptual basis to believe that it is moving, and moving in this or that way. Eva will register the movement visually, and she will be disposed to treat what she sees as evidence of the object moving and of its moving in a certain manner; she will have the visual ability, and the visually based reasoning ability, of a normal human being. And the fact that she has this new visual and reasoning ability means that there is something new that facts about the objectual properties of motion are bound to be like for her.

There is something new that such facts are bound to be like for her because, in registering those facts as such, she will by that very token register them as connected up with other facts in a variety of ways. She will register them as visually salient, of course; that is, she will register the facts in an awareness, explicit or implicit, that the evidence of her eyes—evidence sensitive to distinctively visual obstacles—leads her to do so. And she will register the facts as entailing or supporting other facts, or as being entailed or supported by them. Thus she will not just see the object moving at a certain rate but will see it as moving at such a rate that it is bound to hit a certain obstacle. Or she will register it as moving in such a manner that here or there is the place to reach out and pick it up. Or she will see it as moving in such a way that this or that change of position ought to confirm the appearance. And so on. Seeing motion will involve registering the motion facts in an inferentially rich way, as we might put it; each individual fact will present itself within a network of presumptive implicators and implications.[1]

That Eva registers the familiar motion facts in a manner that displays these connections—those extra facts—means that the familiar facts must present themselves to her in a new way. They are not recorded one by one, in the manner of propositions in an articulated body of prior information, for example; they reveal certain inferential connections more or less immediately. And that means, intuitively, that there is something new that the facts about motion are like for her. It does not mean, however, that there is any new fact she learns. Within the stroboscopic room she already knew all the facts—including all the facts about connections—that she now registers outside. And she was perfectly able to recognize within the room that once outside those facts would present themselves to her in the inferentially rich manner described.



One aspect of the inferential richness with which the familiar facts will present themselves to Eva outside the room is worth mentioning in order to underline the significance of the change involved. With many facts to do with motion, say, with the position of a moving body in her vicinity, she may have registered them previously only in a nonindexical manner; she may have known that the body will occupy such and such a position at such and such a time in objective space, for example, without having any perceptual access to that position. But when she registers the moving body visually before her she will know of the position of the body at the time in question, not just in objective space, but also in egocentric space: she will have knowledge of that very fact in a way that allows her to say 'It's there (at such and such an angle and distance from me)'. This egocentric way of knowing the fact of the body's position facilitates action on the body in a manner that a nonindexical way of knowing it does not; it facilitates relevant practical inferences as to where to reach in order to grasp it—I should reach *there*—and so on. Her knowledge of where she was in objective space would have allowed her in the stroboscopic room to work out that she should reach in a particular way in order to grasp the object. But it would not have facilitated the inference in the same way as the egocentric, visual representation (Perry 1979).

### Intentional Facts About Seeing Motion

As the emergence of Eva's new visual and reasoning ability will not give her access to any new facts about the objectual properties of motion, so it will not give her access to any new facts about the intentional or nonintentional properties of the experience of motion. Take facts about the intentional properties, first: facts to the effect that this visual experience is an experience as of slow or fast motion, as of motion that is steady or jumpy, as of upward or downward motion, and so on. These are all facts about the objectual properties that the experience represents the motion, real or apparent, as having.

What is it for a visual experience to have such an intentional or representational capacity? I shall assume that a visual experience represents motion as having certain properties, in the relevant sense, so far as it overtly disposes the subject to make corresponding judgments about those properties. The visual experience of motion will overtly dispose subjects

to make judgments about its objectual properties such as its direction or velocity or rate of change insofar as two conditions are fulfilled. First, a condition that is grounded in the conception of experience as an overt representation: subjects have to be aware of what is experienced by them as something they experience—something they register and conceptualize in that way. And second, they have to be aware of what is experienced as disposing them, without compelling them, to judge that the direction or velocity or rate of change of the motion is in such and such a range. This notion of an overt disposition of an experience—that is, an overt-to-the-subject disposition—is important for my purposes, and I will be employing it throughout the essay.

Like the objectual facts about the motion itself, intentional or representational facts about motion vision are bound to be known to Eva in the stroboscopic room, given that she knows all the physical facts there are. The Eva argument, like the analogous argument with Mary, is meant to raise a problem for physicalism, without assuming that other, more basic problems get in the way of the doctrine. And if it is not a physical fact that a visual motion experience represents the motion as having these or those objectual properties—if it is not a physical fact that the experience overtly disposes the subject to ascribe these properties—then physicalism is confounded quite independently of the knowledge argument. Thus it must be presumed within the knowledge argument that Eva knows all the intentional facts about various visual experiences of motion while she is in the stroboscopic room. There is no such fact remaining for her to learn when she emerges from that room.

What happens in this case, of course, as with objectual facts about motion, is that Eva will acquire a new perceptual and reasoning ability, on leaving the room, to read off the intentional properties of the visual experiences that lead her to posit motion. Seeing what she takes to be the motion of a fast-moving object, for example, and being aware that that motion is made visually salient to her—being aware, as human beings are able to be aware, that she makes judgments about it on the basis of what appears before her eyes—she will by that very token register an experience in herself as of fast motion. And so on in the other cases: the intentional properties of the experience will be read off from the objectual properties ascribed, rightly or wrongly, to the motion experienced. Thus there will be something new it is like for her to experience motion. There will be



something new, more strictly put, that the intentional facts about her experiences of motion are like for her, as there is something new that the objectual facts about motion itself are like for here. Those facts, known already in the stroboscopic room, will make a new sort of impact as soon as she can visually register motion; they will become available in a new immediate way and, as with the objectual facts, spontaneously display their inferential connections with one another. They will assume a visual life of their own, as we might put it, and cease to be inert items of information recorded in her encyclopedic knowledge of all that there physically is.

Motion vision has certain intentional properties, we assumed, so far as it represents motion as being slow or fast: that is, so far as the experience overtly disposes the subject, in the sense explained, to make corresponding judgments. I shall describe properties that are associated with the effects that an experience overtly has on the subject as overt powers, where the powers are not potential but active. While intentional properties of an experience are overt powers in this sense, being associated with the active effect of the experience in overtly disposing the subject to make certain judgments, they are not the only overt powers an experience may have. Nonintentional properties may be overt powers too, as we shall now see, being associated with the active effect of the experience in overtly disposing the subject toward certain nonjudgmental responses.

### Nonintentional Facts About Seeing Motion

Eva will become aware of nonintentional facts about motion vision as soon as she leaves the stroboscopic room, and finds herself soothed by the gentle motion of a falling leaf, enthralled by the elegant motion of a plane coming into land, recoiling or ducking away from the motion of an object in the direction of her face. Are these also facts that she was in a position to register in the stroboscopic room? And if they are, can we adequately account for how they change in Eva's perspective—for how there is something new they come to be like for her—just on the basis of recognizing that Eva gains new visual and related abilities?

The nonintentional facts mentioned in these examples involve effects that the experience overtly has on the subject. What is experienced is conceptualized by the subject as something experienced, and subjects are aware of what is experienced as disposing them appropriately; they are aware of the motion experienced as affectively soothing or exciting, for

example, or as prompting them to certain motor responses: say, to recoil or duck. In the word I just introduced, the nonintentional facts mentioned all bear on overt, active powers in the experience. Eva will certainly have known about such overt powers just on the basis of knowing all physical facts. If these sorts of facts were not physical in character—and so, being physical, available to her in the stroboscopic room—there would have been no need to invoke the knowledge argument against physicalism; the doctrine would already have been dismissed.

Plausibly, all nonintentional properties of Eva's visual experience of motion are overt powers in that sense. Were there a nonintentional property of which Eva became aware in her visual experience of motion, where that property was not associated with the effects of an experience in the intimate manner of an overt power, then it would have to be a property that she could imagine manifesting itself in visual experiences with quite different effects, intentional and nonintentional. She would have to be able to imagine it being present here in a visual experience as of fast motion, there in an experience as of slow, here in an experience as of an object moving toward her face and causing her to recoil, there in an experience of an object moving away and allowing her to relax, here in an experience of soothing motion, there in an experience of disturbing, staccato motion, and so on. It would have to be a property that was overtly present in the experience but was dissociable in principle from such overt effects; we would have to be able to conceive of it being present without the effects and of the effects being present in its absence. It would have to be a property that made itself manifest, not in any such effects—that is how an overt power makes itself manifest—but in a character that was related only contingently to such effects. It is very hard indeed to imagine what such a property might be in the case of seen motion.

There is a well-established word in the philosophical lexicon, of course, for a qualitative property of experience whose essence or character is overt, not in virtue of overt effects, but just on the basis of inspection. This is the Latin word *quale* (Lewis 1995). That word is sometimes loaded with other connotations. Functionally, for example, *qualia* may be taken by definition to play a certain role in the subject's capacity to reidentify, recognize, and imagine experiences; epistemically, they may be taken by definition to be capable of a foundational role, being properties about which the subject cannot be mistaken and perhaps cannot overlook or miss; and ontologically, they may be taken by definition to be inconsistent with physicalism. I shall set aside all those connotations, however, and think of qualia just as qualitative properties of experience: properties whose essence is inspectionally or effect-independently overt to the subject of the experience.



As overt, experiential qualities, qualia contrast nicely with the overt, experiential powers that we have been invoking. You see the overt, active powers in an experience by feeling as such the effects of the experience—by being overtly empowered by the experience—so that the powers are conceived in a way that makes them incapable of coming apart from the effects.[2] You see the overt qualities in an experience just by inspecting them—that's the theory at any rate—so that the qualities are conceived as capable of coming apart from any effects the experience may happen to have on the subject.

If qualia in this sense are absent from Eva's visual experience of motion, then there is no problem in thinking that she will have known about all the nonintentional properties of that visual experience while still in the stroboscopic room. And equally, there will be no problem in recognizing that when Eva develops the ability, just on the basis of being able to see motion, to recognize nonintentional facts about visual experience—to feel their overt powers—there is bound to be something new that those facts will be like for her. They will be read off spontaneously from the sorts of feelings and inclinations that prove to be occasioned by visual experience, now of this sort of motion, now of that. A further aspect of the inferential richness of the visual experience of motion is that it will spontaneously deliver information about such nonintentional properties of the experience at the same time that it delivers information about the objectual properties of the motion experienced and about the corresponding, intentional properties of the experience itself.

## Old Facts, New Modes of Believing

I say that Eva acquires new perceptual and reasoning abilities on leaving the stroboscopic room, but does not come to learn any new facts. Does that mean that what I am envisaging, then, is that Eva acquires a new causal route to a state she had already reached by virtue of her knowledge of all physical facts: the state of believing in the relevant objectual facts about motion, and the relevant intentional and nonintentional facts about the experience of motion? Not quite. I think that there is a more persuasive

gloss to be given to the connection between the new abilities and Eva's beliefs about those facts, one that makes better sense of why we are inclined to think that she really does enjoy a massive transformation and enlightenment—and why we spontaneously say that there is something new that relevant objectual, intentional, and nonintentional facts will be like for her.

When Eva believes in the relevant facts, whether before or after leaving the room, that is by almost all accounts because—or at least partly because—she has the functional profile of such a believer. Thus, to rehearse the standard account of that profile, she is disposed to remain in the belief state so far and only so far as the evidence supports it, and she is disposed to draw the inferences and perform the actions that being in that state makes rational in the presence of her other states. Or at least that is so, when she is operating under intuitively favorable circumstances and within intuitively feasible limits (Pettit 1999). But if this is something that Eva's believing the relevant facts involves, then the profile may be satisfied in virtue of different causal bases, and that is plausibly what happens as between the period when she is confined to the room and the period after release.

While Eva is inside the room she believes a large number of the relevant facts—she has the profile of such a believer—in virtue of having access to scientific formulations of what they involve, being disposed to assent to those formulas and being disposed to let her responses be organized around them. When Eva escapes from the room, however, with many such facts, she comes to display the profile of a believer in a very different way. She comes to be in the state of believing that an object is moving at such and such a rate, for example, not just in virtue of giving assent to the corresponding proposition and allowing it to influence, say, where she concludes that it would be best for her to reach and try to intercept it. She now instantiates that state in virtue of being spontaneously disposed to reach out to that place as she tries to pick up the object—to reach out *there*, as she will think of it—and being spontaneously disposed to make an indefinite range of similar adjustments. If we describe the profile of a believer in the relevant facts as involving dispositions to act and adjust in certain ways, then Eva in the stroboscopic room is disposed to act and adjust appropriately on a very different basis from the basis on which she comes to be so disposed outside. She believes the same set of facts inside the room as she believes outside, but the shift in her perceptual and inferential abilities ensures that her mode of believing many of those facts will be very different.



This comment connects with the more general claim that believing a given fact can come about in different modes: say, as I have described them elsewhere, relatively more intellectual modes, relatively more practical modes, or modes involving both aspects (Pettit 1998). The logician who can formulate and assent to modus ponens—but who may have to take pains to remain faithful to it in practice—believes the principle in a characteristically intellectual way. The regular reasoner who cannot formulate it but who can recognize instances and is naturally disposed to treat the premises in each case as a reason for endorsing the conclusion believes it in a characteristically practical manner. The anthropologist who believes that all berries of a certain plant are edible—but who may lack the ability to identify such berries reliably—believes that proposition in an intellectual mode. The native who has never reflected on that general claim but who is able to identify those berries and is disposed to think of any such berry he or she identifies that it is edible—a disposition that may appear just in their treating it as edible—believes it in a practical mode.

In the same way I say that Eva in the stroboscopic room believes—indeed knows—various objectual facts about motion, and the intentional and nonintentional facts about the visual experience of motion, in one mode of believing, and that Eva outside believes and knows those facts—or at least believes and knows those to which she has been visually exposed—in quite another. Where she previously knew those facts in a purely intellectual way, as we might put it, she now believes them by virtue of having practical dispositions to draw inferences and make adjustments on the basis of continually monitored visual information. What were previously inert items of information that have to be explicitly connected up with any responses they rationalize now give way to modes of belief that materialize in a network of dispositions that source such responses spontaneously. Not only is there a shift in the perceptual and reasoning abilities that Eva enjoys, then; there is, by that very token, a change in the mode in which she believes many of the physical facts that were previously in her possession.

We can put this in another way if we think about belief from the point of view not of the subject in whom it involves a certain sort of disposition, but from the point of view of the possibilities among which it situates

the subject. When I believe that something is the case, I situate the actual world and my position within the world in relation to various possibilities, taking it to realize this possibility but not that, and so on (Lewis 1983, essay 10). That there are different modes of believing—different ways of being disposed in the way belief involves—amounts to the fact that there are different modes of locating oneself among the possibilities. With some possibilities self-location will be typically effected by the intellectual means afforded by scientific theory; with others, it will be more usual to locate oneself in the perceptually based manner of regular observation and experience.

Now consider Eva inside and outside the stroboscopic room. And think of those possibilities to do with whether an object in Eva's vicinity is moving quickly or slowly, is going to reach a point early enough to avoid colliding with another moving object, is gathering speed as it moves away from her, and so on. Normally sighted people find their way among such possibilities, ruling these in and those out, by relying on continually updated visual information and on the spontaneous reasoning ability that attends this, and that is the way that Eva will come to locate herself in relation to those possibilities when she leaves her room. But within the room things are very different, for here the means of self-location that Eva relies on in relation to those possibilities will often have the very intellectual cast associated with theoretical, scientific knowledge; only occasionally will her kinesthetic and tactile experiences give her any more practical take on the relevant motion possibilities. Thus we have to say that while she may not achieve any more specific level of self-location after leaving the room, she will come to be able to locate herself among the possibilities in a very different manner. She may have nothing to learn in the matter of which possibilities are realized in the actual world and which not, but she does have to learn how to situate herself in a wholly novel way among those possibilities.[3]

## 2 Defending the Analysis

I argued in the last section that though Eva in the stroboscopic room knows all the physical facts there are to know, she does undergo a great shift in the way she can know things on leaving the room. She comes to be



able to believe in a new, practical way facts that she previously believed only in quite a different, purely intellectual manner; there is something new that objectual facts about motion, and intentional and nonintentional facts about seeing motion, come to be like for her. Yet I maintained that nothing in this transformation should lead us to think that Eva learns any new fact. And nothing about Eva, therefore, should suggest that physicalism is false: that there are facts of a nonphysical kind that she did not know in the stroboscopic room and that she learned on leaving that room.

One crucial step in my analysis was the argument that as the intentional properties of a visual experience of motion are all overt powers—as, say, the property of being as of slow motion is an overt power of evoking the judgment that the motion is slow—so too all the plausible nonintentional properties are overt powers. They consist in the experience's being such as to dispose the subject overtly toward this or that response, whether it be a response of being soothed or excited, ducking away from a moving object, or reaching out in its direction. All the intentional and nonintentional properties of the experience, so I argued, should be thought of as overt powers of evoking judgmental or nonjudgmental responses, and powers therefore whose essence is known to the subject only in those very effects. None should be conceived of as an overt quality: a property whose essence is known inspectionally to the subject, not just via the effects that it is found contingently to have.

### An Objection to the Analysis

There is an argument that may be invoked against this step in my analysis, however, and used to support an antiphysicalist, counterintuitive reading of Eva's case. I consider that argument in this section and try to show that it is not compelling. It is important to see why the argument should not be allowed to get a hold on us, for the principle driving it represents a powerful source of mistaken, antiphysicalist intuitions; it is liable to influence our general take on the debate between physicalism and nonphysicalism, and not just to warp our reading of Eva's case, or indeed of Mary's.

The argument can be set out as follows.

(1) If Eva's visual experience outside the room has the intentional property of being as of slow motion, or whatever, then it will overtly dispose her to judge that there is motion and that the motion ascribed is slow.

(2) If it overtly disposes her in this way to make this judgment, then there must be an intrinsic property of the visual process that leads her, as a matter of contingent fact, to make it; otherwise, implausibly, the disposition would be bare and unexplained (see Smith and Stoljar 1998).

(3) One possibility is that the property that leads her to make suitable judgments is a covert—say, subpersonal—property; this property would be causally responsible for the capacity to make the judgments but would be unknown to the subject. This possibility appears to be ruled out, however, by the following argument.

(4) Eva will be able to inspect what she experiences more closely to see if the ascription of slow motion is really well supported: to see if it continues to be reliably prompted and presumptively justified. And so she must be able to inspect a property of the motion experienced that can be identified independently of the effect and judged for whether, contingently, it supports the ascription of slow motion.

(5) Thus the property that operates in disposing Eva to ascribe slow motion, by claim (2), is bound to be a property that is inspectionally overt to Eva—or can become inspectionally overt—in its character or essence, and not just in the effects that it is found contingently to have. It is bound to be an essentially overt, qualitative property of the experience: a quale.

(6) But such a property is not one that Eva could have known in its essence within the stroboscopic room. Lacking the ability to see motion, she could not have had the essence of the property revealed to her there. And none of the physical properties she knew in their essence—certainly none of the causal properties that might underpin a disposition—would be capable of having their essence revealed in that way (Lewis 1995).

(7) Thus there is a new fact that Eva learns, contrary to the analysis offered. She learns what the qualitative property that prompts the ascription of slow motion involves: what it is in its essence. She learns how slow motion looks in the sense in which this involves more that its looking slow: in the sense in which it involves its looking a substantive way that happens as an inductive, contingent matter to prompt the ascription of slow motion.

This argument seems to establish, contrary to the analysis of the last section, that the visual experience as of slow motion will have a property that is accessible in its inspected character only to those enjoying the experience—call this the property of being 'slowish'—and will therefore



escape Eva in the stroboscopic room. Eva in the stroboscopic room will be able to know that there is such a property—she will have access to the argument given, after all—but as between the possibility that it has this inspectable essence or that one she will not know which possibility is realized; she will not even be able to discriminate such possibilities as candidates that she can properly understand. She will not know what slowishness is like in itself.

### Responding to the Argument

How to reply to this line of argument? My intuitive response is to say that there has to be something wrong because, as suggested in the last section, it is really hard to think that motion vision makes salient any qualititative property like that of being slowish or fastish—or there-ish or here-ish, upish or downish, and so on. Being a property known inspectionally in its character or essence to Eva, not just in any of the effects it is found contingently to have, the sort of property in question should be one that she can imagine remaining constant while the associated effects change, and vice versa. Being a qualitative property—a *quale*—slowishness should be one that might conceivably be present in fast motion and that might be associated for people other than herself with fast motion; and similarly upwardishness should be a property that might conceivably be present in downward motion and in other people's visual experience of downward motion; and so on in other cases. But imagination boggles at the attempt to envisage anything Eva comes to see in slow motion or upward motion that might be dissociable in this way from the effects of seeing slow or upward motion.

This response is strongly reinforced by a body of research in empirical psychology (Kohler 1961; Taylor 1962; Kohler 1964; Kagan 1989; see Hurley 1998, 346–349 for a useful discussion). The subjects in a familiar series of experiments were fitted with glasses that turned the visual world upside down or inverted it on the left-right axis: in one sort of experiment what appeared up was really down, what appeared down was really up; in the other what appeared on the right was really on the left, what appeared on the left was really on the right. This would have affected, not just what was statically up and down, or statically on the right and left, but also what moved up and moved down, or what moved right and moved left; it would have affected perception of every such objectual property of

motion. Now if there are qualia of the experiences as of upward movement and as of downward movement, to take that case, then in this sort of experiment what we ought to predict is the following: that while people are likely to learn in time to make correct judgments—and related, motor adjustments—on the basis of the inverted qualia, they should continue to find that upward motion looks the way downward used to look, and vice versa. After all, if the argument is sound, then the people in the experiment have inverted certain qualia and merely learned to change their wonted, judgmental and related responses to those qualia.

But this is not what people in the experiment reported. They did indeed learn to make correct judgments and adjustments after a week or so of wearing the glasses, but as their dispositions changed in this respect so at the same time did the way things looked to them. They gave no report of continuing, unchanged qualia—unchanged qualia playing different roles in producing judgment—but said rather that after the period of correction everything was as it had been for them before they began wearing the glasses: what was up now looked up, once again, and what was down looked down. And equally in the other case, subjects reported that what was to the right now looked to the right, what was to the left now looked to the left. In neither case was there anything that remained constant in the way in which we might have expected qualia to remain constant.

If this is what happened in an attempt to switch the alleged qualia of experiences as of upward and downward motion, and indeed as of rightward and leftward motion, it seems reasonable to suppose that the same sort of thing would happen in any attempted switch of the other qualia that must attend visual motion, according to the antiphysicalist argument. There must be a false premise in that argument, then, or perhaps a mistake of reasoning. But where is the error to be found?

I suggest that the problem is in the fourth claim in the argument, according to which the property that grounds the disposition to ascribe slow motion must have an inspectable, independent character if the experience is to be sensitive to inspection. That claim supposes that if Eva attends closely to the motion seen in the attempt to determine whether she should really ascribe slow motion, then in scutinizing the evidence of her eyes she must be paying attention to a property under an aspect that is independent of the experience's overtly disposing her to ascribe slow motion. The experience may come overtly to dispose her—more or less strongly—



to ascribe slow motion, as she concentrates her gaze. But the idea is that this can only be because her concentration is giving her access to a distinct, qualitative property that causes such strengthening or weakening. It is this idea that lends plausibility to the fourth premise and drives the argument.

But why should we accept that idea? Why shouldn't things go rather as follows? As Eva concentrates her gaze, various things happen in the subpersonal realms of her visual processing, as a result of which the experience that was previously as of slow motion gives way to an experience with a different overt power and a different intentional cast. Perhaps it is now more sharply as of slow motion, as we say, or perhaps it ceases to be as of slow motion. But in either case the only properties that are ever available to her in the experience are the property in the motion ascribed of being more or less slow and—by the same token—the property in the experience of being more or less as of slow motion.

If you and I and Eva are designed by mother nature to operate like this, then in determining whether to trust our eyes as to the speed of observed motion, we will certainly go back to the experience itself and find properties in it—being as of slow motion, looking slow—that we can quote as reasons for the corresponding judgment. Thus there will be something available in experience that supports the ascription of slow motion. But that something—the experience's being as of slow motion—will be an overt power through and through. The experience will dispose me or you overtly to ascribe slow motion but it will have no other aspect than that. It will not be an overt quality of the kind invoked in the antiphysicalist argument.

### Situating This Line of Response

This line of response suggests, in the spirit of Wilfrid Sellars's (1997) attack on the myth of the given, that consistently with becoming better available on the basis of increased attention, a property may be nothing other than an overt power: something that empowers the perceiving subject, without giving him or her access to the property in itself that grounds the power. There is a property given to me, the subject of a visual experience as of slow motion. But what is given is not anything in the nature of a quale, with an inspectable character that is related in only a contingent and inductive way to my becoming disposed to make a certain judgment. What is given is nothing more or less than the character of that experience that is

exhausted in its disposing me in that way; what is given is just an overt power of the experience.

That overt power will have a categorical grounding under any plausible metaphysics, but this can be unmysterious and physical and need not be available to the subject (Lewis 1995). For me the property will reveal itself in the fact that the experience overtly disposes me to ascribe slow motion. But the ground of that disposition will be discernible only within the realms of neuropsychological analysis. Things will conform to the pattern acknowledged as a possibility in the third premise of the argument given.

The position adopted—and, more generally, the rejection of the myth of the given—is supported by a certain class of discoveries in neuroscience. What these show is that perceptual data may elicit in a subject a disposition to behave in a certain way, quite independently of the subject's being aware of the character of those data that explains that disposition; and that with the disposition in place the subject is then positioned to see the data under a corresponding gestalt: that is, as having a suitable character, a suitable overt power. The nature of what is registered, in its subpersonally available profile, prompts a certain response, and, given the tendency to make that response, the subject then sees what is registered as having a corresponding character: in effect, as being the sort of thing that is fitted to that response.

Consider the findings of a recent study in which subjects are asked to gamble with four decks of cards, where two of the decks are stacked against them (Bechara et al. 1997). While the subjects do eventually come to register those two decks as stacked, and so to resist them, it turns out that the resistance response—or at least the disposition toward that response—is present long before any perception of the decks as stacked, or even as suspect, emerges. This is evident from imaging of what is going on in their brains when they are dealing with those decks, as distinct from the fair ones, and from associated skin conductance responses. Unconscious resistance materializes in these subjects on the basis of a subpersonal registering of the fact that things are going wrong with the suspect decks. And the eventual registering of the decks as suspect—'there's something I don't like about them'—takes the form of registering them simply as decks that occasion that resistance.

The myth of the given may be sourced in the idea that the explicitly registered charactered of what is presented—what is inertly given—



prompts the subject's response and so must have a character for the subject that is dissociable from that response.[4] However, the truth may rather be, as this sort of finding suggests, that often things are the other way around. It may be that what perception first of all provides is an unconsciously effective response-guidance system, and that as responses congeal, the percepts get articulated as having a character fit to support them.[5] So far as that is the case, the properties revealed in preception will be overt powers, not qualia; they will be known to the subject in the effects they elicit, not on an inspectional, effect-independent basis.

## 3 The Case of Mary

### From Eva to Mary: A Parallel Analysis

By the analysis offered, Eva gains a practical, visual way of knowing about motion, and hence about the vision of motion, on leaving the room. What she gains is not new factual knowledge but a new mode of registering and believing facts already known. Why shouldn't things be the same for color? And why shouldn't we then think that Mary's enlightenment as she leaves the black-and-white room involves nothing more than Eva's: that Mary comes at that moment to know in a spontaneous, practical way facts that she previously knew in a different mode? Why, more particularly, shouldn't we say that the sense in which Mary learns what it is like to see colors is just the sense in which Eva learns what it is like to see motion? There is something new that the objectual facts about the colors of things—facts to do with hue and saturation and brightness—will be like for Mary. And there is something new that the intentional and nonintentional facts about color experience will be like for Mary. But there will be nothing new that Mary discovers when those facts assume that profile for her.

On this story, Mary will be able on leaving the room to scrutinize the colors of things with a view to making more reliable judgments about their objectual properties; with a view to becoming surer about the properties her vision represents things as having—about the intentional properties of her experience; and perhaps with a view to savoring various nonintentional effects that she enjoys in the experience. But in doing this she will not be eye-balling anything in the way of a quale: a property with a character or essence that is inspectionally revealed to her for the first time. Rather, she will be putting herself under the control of the overt powers of

seen color; she will be putting herself in a position to let the powers of color work on her.

There is no good reason, I believe, why we shouldn't go along with the lesson of Eva's case and endorse these claims about Mary. For the Mary argument to have any point, it has to be assumed that color is a physically analyzable property that is detected by human beings in a physically analyzable way; otherwise physicalism would be undermined before the argument ever got going. But that being so, it seems that the points made in discussion of Eva's case will carry over smoothly to Mary's. Mary in the black-and-white room is bound to know the facts about the objectual properties of color, including the color of this or that object, since she knows all the physical facts. She is bound to know all the corresponding intentional facts to do with which color experiences are experiences as of which colors. And she is bound to know all the facts to do with how far experiences as of this or that color have certain nonintentional properties. Yet such antecedent knowledge is consistent with Mary's enjoying a radical transformation on leaving the black-and-white room; in particular, a transformation that may incline us to think that she learns a new fact. For on leaving that room, by the lesson of our analysis in Eva's case, there will be something new that the objectual, intentional, and nonintentional facts about color are like for her. She will be in a position, given color vision, to believe in a new and practical mode facts that she previously registered only in an intellectual or theoretical manner. She will be able to locate herself among those possibilities on a new basis—to rule this possibility out, that possibility in—but she will continue to locate herself at the same place.

The analogies between the two cases are sufficient in themselves to make it plausible that Mary's case should be analyzed on such parallel lines. But one particular attraction to that strategy is that it suggests a nice account of what it is for something experienced to look red or blue, or to look bright or saturated, or indeed to look warm and comforting or cold and disturbing. We are invited to think about those properties of experienced color as overt powers of what is experienced: as powers that are exhausted for subjects in the experience's overtly disposing them toward relevant responses, judgmental or nonjudgmental. Those powers may wax or wane with attention and focus but the important thing is that they are not qualia. They are not properties whose essence is available in the visual experience on the



basis of inert inspection, as something that just happens to prompt the judgmental and nonjudgmental responses. They are not properties such that we can imagine them varying independently of such effects: remaining present while the effects change, or changing while the effects remain the same.

Experienced color has overt powers, under this story, but no overt character that is independent of the powers it exercises, relating only inductively and contingently to the effects involved. And that being so, we can see why Mary in the black-and-white room will be able to know about those properties and yet not be able to form beliefs about them in the manner of ordinary people. She will know the overt powers of colors, through knowing that color properties—surface reflectance properties, as we are taking them to be—have effects of the required sort. But she will not know them in the manner of ordinary people. While knowing that the properties have such and such effects, none of those effects will materialize in her. She will not undergo those effects, in the manner of ordinary people, as effects overtly prompted by what is seen in the experience of color.

### Objection: There Is a Difference

But while there are many analogies between Eva's case and Mary's, and many advantages in defending a parallel analysis, there is one salient difference of intuition between the cases. Most of us will agree that there really seems to be nothing distinct from the overt powers of seen motion that is identified in seeing objects move; there is no quale that might in principle vary independently of overt effects, being only inductively and contingently associated with such effects. But it is not so clear that most of us will be equally moved by the parallel claim in Mary's case.

On the contrary, so it may be said. When I see or seem to see a color, I don't just know that the color experience has the intentional property of being as of a certain color, nor that it has the nonintentional properties of soothing or boring or exciting me, though I certainly do know those things. Beyond knowing that it has such overt powers—beyond experiencing those powers—I also know that the experience has a certain quale, unexhausted by any particular effect. I know not just that what is experienced looks red or yellow, for example—that is, overtly disposes me to make corresponding judgments—but that it looks reddish or yellowish, where 'reddish' is understood in the same manner as 'slowish' and registers

the presence of an overt quality that is only inductively and contingently associated with disposing subjects in certain ways. If I know this just in virtue of having color experience, then what I know—that red things look this or that substantive, reddish way—is something that Mary will come to know only after emerging from the black-and-white room.

This intuition about what Mary has to learn is associated with the intuitive plausibility, long recorded in the philosophical literature, of an inversion in the space of color experience. The idea is that there is nothing incoherent in imagining that the qualia of different color experiences might be exchanged, or that the qualia of your experience and mine of one and the same color might be different. I can imagine my experience as of red being the overt, substantive way that my experience as of green currently is, and vice versa. And, for parallel reasons, I can imagine my experience as of red being the overt, substantive way that your experience as of green currently is. What looks red to me now is reddish, so it is suggested, whereas what looks red to me later, or to you now, may look greenish. And this may be so without our performance being in any way impaired. The idea is not that this inversion is a real-world possibility, being consistent with the actual laws governing color and the perception of color—that claim has been laid to rest, though it had life for a time (Hardin 1997)—but that it is at least a logical or metaphysical possibility.

We saw in the first section that it is very hard to envisage what the quale of slowishness in seen motion—or any such property—could be, given that it would have to be capable in principle of varying while the effects of the motion seen remained the same, and vice versa. What now appears, however, is that things are very different in the case of color. We do seem to find color qualia intuitive posits, so that we will be naturally disposed to think that Mary, unlike Eva, learns something new on escaping confinement.

### Explaining the Difference

How to explain this difference of intuition? How to explain why we find it almost impossible to imagine what a quale of slowishness in seen motion might be, while having no difficulty in envisaging a quale like the reddishness of a color seen? Why does it require an abstract, specious argument to give life to the idea of qualia in the case of motion but not in the case of color? Happily, I think that there is quite a plausible answer available and



that it allows us to maintain the view that Eva's and Mary's cases should be treated on a par.

The motion of an object concerns me as an ordinary visual observer, not just in its intentional or judgmental effects, and not just in nonintentional effects of an affective kind: its being soothing or exciting and so on. It concerns me above all, and for reasons that reach deep into our nature, in the motor adjustments it requires. If I have an experience as of an object accelerating in the direction of my face, for example, then what I experience had better prompt me to duck and, *qua* experience, overtly prompt me to duck. If I have an experience as of an object beginning to move in the direction of my child then, assuming the object can be stopped, what I experience had better prompt me—and prompt me overtly—to reach out here or there in order to intercept it. And so on. Motion is a primary property that is associated not just with effects on a single sense like that of sight, but also with an indefinite range of other effects. Thus, in learning to see motion, what I see had better be something that overtly prompts me in corresponding ways. Otherwise it will not constitute an experience as of the primary property of motion: an experience as of a property that has the wide range of effects typical of primary properties.

The color of an object concerns me in a very different way, since color is a secondary property that we conceptualize as something that is revealed only to a single sense, sight. There are no responses to color that I clearly must have, in the way in which I must have various responses to motion. And so the color that I see in an object, unlike the motion that I see there, is not something that is so obviously required to dispose me overtly in this or that manner. I will not count as seeing motion unless what I see prompts me overtly to respond with certain bodily adjustments. I will count as seeing color, for all that is obvious to us as ordinary subjects, just so far as what I see prompts me overtly to make suitable judgments. Motion is a primary, multisensory property that concerns me in various ways, so that seen motion must overtly dispose me to respond in that variety of ways. Color is a secondary, unisensory property that concerns me only in its effects on the eyes, so that seen color need only overtly dispose me—so at least it seems—to respond with suitable visual judgments.

Consider the case of motion. The fact that the visual experience of motion is experience that has to overtly prompt a variety of nonvisual effects—in particular, the fact that this is something salient to all of us—

makes it hard to think of anything given in the experience, any quale, that might change while the experience continues to prompt those effects, or vice versa. The quale of a particular visual experience of motion, if there is any, will have to be capable of varying independently of the effects the experience overtly has on the subject. And yet it will also have to mediate—and overtly mediate—those effects; it is supposedly the overt character of the experience in virtue of which it disposes the subject to respond in this or that manner. But it is hard to see how it could simultaneously satisfy these two demands: these constraints of independence, on the one side, and mediation on the other. As the range of effects to be mediated overtly by the experience increases, it is going to seem less and less plausible that the mediating property could be independent enough to vary while those effects remain in place, or vice versa.

For this reason the space in which we can imagine a purely qualititative change occurring in the visual experience of motion is restricted to a vanishing point, as we consider the range of effects, in particular effects of a nonvisual kind, that the experience will overtly have to elicit. The overt powers of the visual experience are so rich and salient that they leave us little room for envisaging the possibility that apart from those powers, the experience might also have a dissociable, overt quality.

Think of the case where people wear glasses that turn the visual world upside down and make upward movement look downward, and downward movement look upward. Imagine those people becoming accustomed to this shift. Imagine them learning to put their heads down, for example—this will be kinesthetically registered as down—when they see an object coming up at them, as it will seem. And imagine this learning becoming second nature, until the way things are visually presented automatically produces required nonvisual responses. It is very hard indeed to think of what it is that might remain the same in the subjective quality of the motion seen, as all of those responses are adjusted. How could the motion continue to look as if it comes from down there, for example—where 'there' will be identified in egocentric space—while the way it looks automatically leads them to put their heads down?

This line of thought explains very well, I think, why it requres an abstract and specious argument to give any life to the idea of qualia in the visual experience of motion. Were there a quale to a visual experience of motion it would have to mediate so many overt effects, that it would be hard to see



how it might have an independent, overt character: a character capable of independent variation.

But now think of color. Things are very different here, since color is a unisensory, secondary property. The objectual properties of color that correspond to properties of motion like speed, acceleration, and direction involve hue, saturation, brightness, and so on. And those are all properties that involve vision, and vision only. That an experience is as of a deeply saturated, brightly illuminated red object, for example, signifies nothing about it—so far as we can tell—that directly impacts on any nonvisual sense. And in this respect it differs greatly from an experience that is as of an accelerating object that is moving in the direction of my face.

The fact that objectual color properties are purely visual means that the demands on experiences as of objects having those properties are not so obviously rich as the demands on visual experiences as of objects having various motion properties. And that in turn makes it intuitively more imaginable that there might be color qualia: properties of seen color that satisfy the dual constraints of mediating certain overt effects and of being able, at least in principle, to vary independently of those effects. It makes it intuitively imaginable that consistently with taking on the greenish quale, for example, the experience as of something red might continue to function perfectly well in overtly disposing us to discriminate red objects from objects of other colors, to see them as similar to objects of the same color, and to judge that they are red; all that would be needed is that a compensating change occur in experiences as of something green.

This being so, it is understandable why we might give credence to color qualia and expect that Mary will learn something when she leaves the black-and-white room. Color qualia are intelligible posits, where it is doubtful if motion qualia—the qualia of seen motion—are. Or at least they are intelligible against the background of our received folk psychology, which identifies much that seen motion has to dispose us to do and little that seen color has to effect.

But the fact that color qualia are folk-psychologically intelligible posits does not imply, of course, that they are posits we are well advised to make, and it does not imply that Mary will actually learn something on leaving the room. There are three points to make in support of sticking with the lesson from the analogy with Eva (for more detail see Pettit 2003). These are: first, that color qualia may be folk-psychologically intelligible without

being intelligible, period; second, that however intelligible, they are not economical posits; and third, that there are some experiments in which they might well have been confirmed but weren't.

**First Point: Color *Qualia* May Not Be Intelligible Posits** What is it for something to look red? In particular, what is it for something to look red, putting aside the possibility that there is a reddish quale present? Our spontaneous folk-psychological intuition—by contrast with intuitions in the motion case—is to think that something rather bare and simple is involved; if the possibility of a *quale* is put aside, it seems that to look red is just to look discriminable in ordinary conditions from orange and yellow and green and so on. It is this intuition that leaves us with the feeling that there has to be more than this occurring when an object looks red: there has to be a distinctive quale on offer also. The intuition is highly suspect, however, and may only survive in the absence of reflection on what color vision has to achieve for us, and in the absence of information about how it does that.

Consider someone who has the sort of discriminatory capacity mentioned. He or she is able to discriminate any two objects on the basis of their having different colors, at least in ordinary conditions: say, where the light is good, the background steady, and the subject's viewpoint unchanging. This ability falls well short of the portfolio of abilities associated with color vision, as a little reflection makes clear, particularly in light of recent scientific investigations. That portfolio must also include these capacities:

- to discriminate two objects of shades of the same color;
- to discriminate two objects of different colors under varying illumination, against changing backgrounds, and from a changing viewpoint;
- to recognize one of the objects as having the same color as a third object which lies at a little distance from it, or even in a different visual context;
- to track an object on the basis of its color—to reidentify it through time as the same-colored object—across different levels of illumination, different backgrounds, and from a changing viewpoint.

Once we take account of the need for these capacities in color vision, we see that to look red, in quale-free terms, involves much more than looking discriminable from other colors in ordinary conditions. Something will look red to a person so far as it looks fit, not just to be discriminated from things of other colors in ordinary circumstances, but to satisfy a much wider range of situation-response connections. It must look fit to stand out in contrast to objects of other shades and colors, under different intensities of incident light; against different backgrounds of landscape and objects; and from different spatial and subjective viewpoints. It must look fit to stand out as similar to other red objects, or to objects of the same shade of red, under parallel variations in lighting, background, and viewpoint. And consistently with looking fit to sustain such contrasts and comparisons, it must also look fit to transform in systematic and distinctive ways, while remaining determinably the same color and determinably reidentifiable, through various changes of lighting, background, and viewpoint.



With this enriched sense of what is involved in something's looking red, it is worth asking whether the quale intuition remains as firm as it seemed to be originally. Can we really imagine something's continuing to have the rich, discriminable, classifiable, reidentifable color profile of an object that looks a certain shade of red, while the supposed, substantively reddish way it looks changes? Can we really be sure that it is reasonable to think that there is a substantive way it looks that is dissociable from the discriminatory, classificatory, and reidentificatory capacities that its looking that way gives us? Isn't it just possible that there is no change conceivable in how the object looks that would not disturb one or another of those capacities? I surmise that this is possible, even plausible, and that a doubt must arise, therefore, as to whether qualia are intelligible posits. It may be that they are intelligible only against a background of an unreflective, uninformed folk psychology of color perception (Pettit 2003).

**Second Point: Color *Qualia* Are Not Economical Posits** Even if we think that color *qualia* are intelligible posits, however, it does not follow that they are posits worthy of credence. And a consideration that suggests they are not is that there is a more economical qualia-free story available to us. We have seen that it is perfectly possible to explain Eva's transformation on leaving her confinement—and so to explain Mary's in parallel—without positing qualia or the learning of a new fact. So why should we posit qualia in the color case and why should we give an antiphysicalist reading to Mary's transformation?

I see no reason why we should do so. Being forced to explain what happens with Eva without invoking qualia—this, because we could not plausibly invoke qualia there—we learned that we could do without qualia. We learned that it makes sense to think of the motion of a body looking slow or fast, upward or downward, without thinking that there is any substantive way it looks: any slowish or upwardish, fastish or downwardish, way it looked. We learned that seen motion can have overt powers, in other words, without having any overt qualities. It would make good sense to invoke the same economical sort of story in the case of Mary, arguing that the ways colors look are not substantive ways that might vary independently of the judgmental and related effects they overtly prompt: that they too are overt powers, not overt qualities.

In particular it would make good sense to do this, given that the status of color as a secondary property explains why there is a quale-intuition in this case, but none in the case of seen motion. That difference in status shows why it is natural in Mary's case, though not in Eva's, to endorse the conclusion that on escaping confinement she learns what certain qualia are. And in explaining away the quale-intuition, it shows why it is not compelling to invoke such learning in explanation of what happens to Mary when she effects her escape.

**Third Point: Color *Qualia* Have Failed Empirical Confirmation** The line taken on motion vision was supported by empirical reseach on how various experiments that should have switched unchanged qualia from one role to another, if there were qualia available to shift, failed to have any such effect. It is worth noting, in conclusion, that the parallel line on color perception is supported in a similar fashion. An experiment in which we might have expected color *qualia* to remain unchanged, while judgmental adaptations shifted, suggests that there was nothing at all that remained constant and so no qualia that remained constant (Kohler 1964; Taylor 1962; for a good discussion see Hurley 1998, pp. 287–288).

The experimenter in this case fitted the subjects with spectacles in which the left side of each lens was tinted blue, the right side yellow—in another variation the top part was red, the bottom green—so that when the wearers looked left things had a blue tinge, when they looked right things had a yellow tinge. But as the subjects learned to adapt in the color judgments they made, and in the associated discriminations and comparisons—this took some weeks—so the way things looked adapted too; all signs of an inappropriate yellow or blue tinge disappeared. Indeed when the subjects removed the lenses, according to the reports, distortions reappeared for an initial period: things had a blue tinge when subjects looked in the direction that the goggles had been yellow, things had a yellow tinge when subject looked in the direction that the goggles had been blue.



These results suggest that as a movement's looking slow or fast just involves the visual experience of the motion having the overt power of disposing the subject to make suitable judgments and adjustments, so something similar holds in how color looks. Let something's looking red or blue consist in its looking fit to satisfy the range of situation-response connections mentioned above; let it consist in its exercising that richly profiled, overt power. When a subject wore glasses of the kind described, then that overt power was disrupted and the subject lost information that was previously available. The information was still there in what reaches the retina, but its encoding was distorted; "the light delivered by the spectacles to the eyes still carried the same information about the environment, but in an altered form. This information had not been destroyed, but only biased" (Gibson 1964, p. 8). What happens, then, when the look of red is restored? On this quale-free story, there is an obvious answer. What happens is that the brain adapts so as to retrieve the customary information from the distorted encoding and, as that adaptation occurs, the things again look exactly as they did without glasses.

If we think that there are color qualia, however, then it is hard to know what to say. Allow that there are qualia of that kind and we should expect that, short of an unexplained shift in quale occurring at the appropriate time, the subjects who began to see normally through a blue or a yellow lens would report that the quale of how things looked in the relevant part of the visual field had changed from what it was without the glasses but that they had learned to make corrected judgments and adjustments in its presence. They reported no such thing, however, giving us serious, additional ground to doubt the reality of color qualia and so to doubt whether there is any reason to treat Mary's case as different from Eva's.

## Conclusion

It may be useful, in conclusion, to sum up the main points argued in the paper.

1. When Eva leaves the stroboscopic room and sees motion for the first time, then she gains a new mode of access to three sorts of facts: facts about the objectual properties—direction, velocity, and so on—of the motion she sees or seems to see; intentional facts about which properties her visual experiences represent such motion, rightly or wrongly, as having; and nonintentional facts about those experiences, to do with how far they are affectively exciting or soothing, or elicit motor responses like ducking or reaching.

2. Eva, knowing all physical facts, would have known all those facts previously. But when she leaves the room there is something new that those facts are like for her, in the sense that she will now see the connections between them in new, spontaneous ways. She will come to believe or know those facts in a practical, engaged way, as distinct from the purely intellectual way she knew them previously. She will situate her world and herself in a new way among the possibilities but the place where she self-locates in this epistemic sense will be just the same place as before.

3. The shift that occurs in Eva is sufficiently significant for it to be understandable why we should be inclined to say, without thinking, that she must learn a new fact. Holding that there is something that such new facts are like for her is close to saying that she becomes acquainted with something new. But we need not, and should not, indulge this inclination. There is no further fact that it is plausible to think she learns when she escapes confinement.

4. Antiphysicalists may resort to a certain argument at this point. Under the story told, when a movement seen or apparently seen looks slow or fast or whatever, then it overtly disposes the subject to judge that it is slow or fast, and when it prompts a response of an affective or motor kind then it overtly disposes the subject toward that response. But the overt power thereby registered in the movement seen or apparently seen—if you like, in the experience of motion—is not a property that has any inspectable character for the subject that is independent of this disposing effect. The antiphysicalist argument I envisage would maintain that apart from such overt powers we also have to acknowledge overt qualities or qualia, and we have to say that Eva learns what the qualia of such motion are on leaving her room.



5. On this usage, qualia are properties whose essence or effect-independent character is something that the subject grasps inspectionally in having an experience like the visual experience of motion. The antiphysicalist argument is that the overt powers of an experience must be grounded in independent properties, and that the fact that subjects can inspect what they experience as they explore how far to let those powers prevail—how far to draw the judgments prompted or make the responses elicited—shows that those independent properties must have a character or essence that is inspectionally overt to the subjects; that is, they must be qualia. The argument implies, then, that Eva will learn new facts on leaving her room: she will learn that when a motion looks slow it has a slowish quale—it looks a slowish way—and so on for other properties seen in motion.

6. This argument is counterintuitive. Qualia have to satisfy two constraints: on the one hand, mediate to the subject in an overt way the judgments and other responses elicited; and second, be independent of those effects and so capable, in principle, of varying while the effects remain the same—and vice versa (Block 1990). But it is highly implausible that anything could fit that bill in the case of seen motion. For being a primary property, any experience that counts as an experience of motion must overtly prompt the subject in a variety of nonvisual ways: the movement of an object toward one's face must prompt one to duck, and so on. And it is all but impossible to see how any property could have all those effects and yet, in principle, be capable of independent variation in an overt dimension. It is almost impossible to imagine, for example, how a fast motion could have a slowish quale, or an upward motion a downish quale. This observation is supported by empirical results to the effect that if people's visual experiences of motion are disturbed by lenses of various kinds—so that upward motion, for example, is presented as downward used to be—then as people learn to adjust and make the correct judgments, they report that motion begins to look appropriate once again; there is no continuing, downwardish quale in motion that they now judge to be upward, and so on.

7. But the argument is not only counterintuitive in Eva's case. It is also uncompelling, since we can plausibly hold, in line with a Sellarsian tradition, that when Eva attends to the seen motion of a body in order to be

clearer about its velocity or rate, she may merely be exposing herself to the overt power of the experience to dispose her judgmentally in this or that way. There is no need to postulate that there is an independent property she is inspecting—slowishness, as it might be—which is only found to be associated contingently with slow motion, and so on. This line fits quite well with recent neuroscientific observations. These show that things we experience may have effects on us without our being aware of what it is about them that elicit those effects and that, with the effects in place, we may see the things we experience as possessed of corresponding properties: the property in each case of overtly disposing us toward the relevant effect.

8. There is every reason why we should take a similar line in Mary's case to that which is defended here for Eva's. The story is economical and plausible. It requires us to acknowledge that when something looks red, there is no substantive way it looks to us—there is no quale on offer—and that it just looks the way that goes with our judging that it is red, making the required discriminations and comparisons. It requires us to think of the looks of color, as we think of the looks of seen motion, as overt powers rather than overt qualities. But this is something that is quite coherent, as argued in the case of motion.

9. All that said, it must be acknowledged that while the antiphysicalistic argument is as uncompelling in the case of color as in the case of seen motion, it is not as counterintuitive in the color case. This is because color is a secondary property that affects vision alone, so that seen color does not have to activate the rich array of effects that seen motion is required to do. It is all but impossible in the case of visually registered motion to see how anything could fill the dual constraint on a quale of mediating such effects but being capable in principle of independent variation: to see how downward motion could continue to have the quale of upward and so on. But no such argument is available with color. And so we can envisage that there is a reddish way that red things look—we can envisage a reddish quale that is registered in seeing red—so that green things might look that way instead of red, and vice versa.

10. We can, but we shouldn't. For it remains possible, as already observed, to hold instead by the economical story motivated by consideration of Eva's case. And two other arguments also offer grounds for rejecting color qualia. The first is that for something to look red is for it to look



fit to sustain a very wide range of situation-response connections, as investigations of color vision have made clear, and that the intuition that there are color qualia is put under pressure by a recognition of this fact. And the second is that as the story told in the case of seen motion was supported by some empirical results, so a parallel story is supported in the same way in the case of color. The earlier results were that people who compensate for lenses that disturb the visual perception of motion do not continue to find the old *qualia* still in place. The results relevant to color show that people who compensate for lenses that disturb the sight of color slightly do not continue to find old qualia in place either. Where their judgment of color goes—and their associated adjustments—so the way color looks to them appears to go also. And that is just what we should expect on the story developed here.[6]

## Notes

1. I discovered in presenting the paper at the Tucson Consciousness Conference 2002 that the views I favor on this topic, and on other matters related to perception—see the discussion of color vision in the last section—are close to the very engaging 'sensorimotor' view defended by O'Regan and Noë (forthcoming). I have adjusted some phrasing in response to their influence. The views are close, more generally, to a range of recent work that emphasizes the role of perception as a skill and the importance of perceptual-practical connections (see Thompson 1995; Hurley 1998; Clark 1999; Myin 2001; Myin and O'Regan 2002). The paper is a companion piece to Pettit 2003.

2. It may be useful to think of the power as a higher-order, disposing property that consists in the experience's having lower-order properties that produce those effects. Why is it said to be a disposing rather than a dispositional property? Because the power is understood in the sense of active, not just potential, power: it exists only so far as the effects are actually materializing. Why is it said to be a higher-order rather than a lower-order property? Because it is overt, being recognized as such by the subject who is exposed to it; the lower-order properties responsible for the effects could not be overt in that way.

3. Perry (1979) argued that a very important shift occurs whenever an agent comes to be able to think indexically, say, in 'I', 'here', or 'now' terms, about themselves or their position; nonindexical thought cannot motivate my acting, or my acting here or now. Such a shift may occur for Eva, as emphasized in the earlier discussion of her being able to think of an object as moving 'there', in egocentric space. I do not think that this sort of transformation reduces the space of epistemic possibilities and represents knowledge of a new fact. But even if that were denied, it should be clear that

the kind of knowledge gained in this indexical shift is not the sort that would vindicate the rejection of physicalism; it does not point us toward a realm of general, physically nonsupervenient facts.

4. Taylor (1962, p. 3) writes: "a great many human actions are preceded by perception, and it is commonly believed that such actions are, by virtue of the preceding perception, essentially different from others, the so-called 'unconsciously motivated' actions. It is assumed that the former are 'guided' by perception while the latter are not. Indeed the truth of this proposition appears self-evident. Nevertheless I propose to doubt it."

5. It may seem strange that perception could prompt responses without first eliciting a conscious recognition that things are such and such, where their being such and such provides a reason for the response. But quite a lot of work suggests that this is so. It can even happen in the extreme cases that subjects are perceptually prompted to respond in a manner that fails to make sense on the basis of their conscious reports. Subjects who prove incapable of distinguishing between letter-boxes with different orientations (north–south, east–west, etc.) can nevertheless adjust their hands appropriately in seeking to insert a letter (Kelly forthcoming). And subjects who are subjected to a visual illusion that makes one disk seem bigger than another of equal size—this is what they report—still adjust their grip as if they were free of that illusion when asked to reach out and grasp one or other of the disks (Milner and Goodale 1995).

6. I was introduced to akinetopsia, and to much else, through my association with the McDonnell Project in Philosophy and the Neurosciences, directed by Kathleen Akins from Simon Fraser University. I am very grateful for helpful observations received in exchanges with David Braddon-Mitchell, Frank Jackson, Sean Kelly, Victoria McGeer, Eric Myin, Alva Noë, and Michael Ridge; for detailed comments given by Laura Schroeter and Daniel Stoljar; and for discussions of presentations made at the University of Queensland, the University of Michigan, and the Tucson Consciousness Conference, 2002.

### References

Bechara, A., H. Damasio, D. Tranel, and A. R. Damasio (1997). "Deciding advantageously before knowing the advantageous strategy." *Science* 275: 1293–1295.

Block, N. (1990). "Inverted Earth." *Philosophical Perspectives* 4: 52–79.

Braddon-Mitchell, D., and F. Jackson (1996). *Philosophy of Mind and Cognition.* Oxford: Blackwell.

Clark, A. (1999). "A Case Where Access Implies Qualia." *Analysis* 60: 30–38.

Gazzaniga, M. S., et al. (1998). *Cognitive Neuroscience.* New York: Norton.



Gibson, J. J. (1964). "Introduction to Ivo Kohler 'The Formation and Transformation of the Perceptual World.'" *Psychological Issues* 3(4): 5–13.

Hardin, C. L. (1997). Reinverting the Spectrum. In *Readings on Color*, volume 1: *The Philosophy of Color.* Eds. A. Byrne and D. R. Hilbert. Cambridge, Mass.: The MIT Press: 289–301.

Harman, G. (1990). "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31–52.

Hurley, S. (1998). *Consciousness in Action.* Cambridge, Mass.: Harvard University Press.

Jackson, F. (1982). "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127–136. Reprinted in this volume.

Jackson, F. (1986). "What Mary Didn't Know." *Journal of Philosophy* 83: 291–295. Reprinted in this volume.

Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis.* Oxford: Oxford University Press.

Kagan, S. (1989). *The Limits of Morality.* Oxford: Oxford University Press.

Kelly, S. D. (forthcoming). "The Logic of Motor Intentional Activity." *Ratio.*

Kohler, I. (1961). "Experiments with Goggles." *Scientific American* 206 (May): 62–72.

Kohler, I. (1964). "The Formation and Transformation of the Perceptual World." *Psychological Issues* 3(4, monograph 12): 1–173.

Lewis, D. (1983). *Philosophical Papers,* Vol 1. Oxford, Oxford University Press.

Lewis, D. (1990). What Experience Teaches. In *Mind and Cognition: A Reader.* Ed. W. G. Lycan. Cambridge, Mass.: Blackwell: 499–519. Reprinted in this volume.

Lewis, D. (1995). "Should a Materialist Believe in Qualia?" *Australasian Journal of Philosophy* 73: 140–144.

Milner, A. D., and M. Goodale (1995). *The Visual Brain in Action.* Oxford: Oxford University Press.

Myin, E. (2001). "Color and the Duplication Assumption." *Synthese* 129: 61–77.

Myin, E., and J. K. O'Regan (2002). "Perceptual Consciousness, Access to Modality and Skill Theories." *Journal of Consciousness Studies* 9: 27–45.

O'Regan, J. K., and A. Noë (forthcoming). "A Sensorimotor Account of Vision and Visual Consciousness." *Behavioral and Brain Sciences* 24.

Perry, J. (1979). "The Problem of the Essential Indexical." *Noûs* 13: 3–21.

Pettit, P. (1993). "A Definition of Physicalism." *Analysis* 53: 213–223.

Pettit, P. (1998). "Practical Belief and Philosophical Theory." *Australasian Journal of Philosophy* 76: 15–33.

Pettit, P. (1999). "A Theory of Normal and Ideal Conditions." *Philosophical Studies* 96: 21–44.

Pettit, P. (2003). "Looks as Powers." *Philosophical Issues* (supp. to *Noûs*) 13: 221–252.

Sellars, W. (1997). *Empiricism and the Philosophy of Mind.* Cambridge, Mass.: Harvard University Press.

Smith, M., and D. Stoljar (1998). "Global Response-dependence and Noumenal Realism." *Monist* 81: 85–111.

Taylor, J. G. (1962). *The Behavioral Basis of Perception.* New Haven, Conn.: Yale University Press.

Thompson, E. (1995). *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception.* London: Routledge.

Zeki, S. (1993). *A Vision of the Brain.* Oxford: Blackwell Scientific.

# 7 Knowing What It Is Like: The Ability Hypothesis and the Knowledge Argument

**Michael Tye**

Mary, as the familiar story goes (Jackson 1982), is imprisoned in a black-and-white room. Never having been permitted to leave it, she acquires information about the world outside from the black-and-white books her captors have made available to her, from the black-and-white television sets attached to external cameras, and from the black-and-white monitor screens hooked up to banks of computers. As time passes, Mary acquires more and more information about the physical aspects of color and color vision. She comes to know all the familiar color names and the objects to which they apply, the physical character of the surfaces of those objects, the way the light is reflected, the changes in the retina and the optic nerve as different colors are perceived, the physical changes in the visual cortex. Eventually, she becomes the world's leading authority on color and color vision. Indeed she comes to know *all* the physical facts pertinent to everday colors and color vision.

Still, as the years go by, she becomes more and more dissatisfied. She wonders to herself: What do people in the outside world *experience* when they see the various colors? *What is it like* for them to see red or green? No matter how often she reads her books or how long she spends examining the printouts from her computers, she still can't answer these questions fully.[1] One day her captors release her. She is free at last to see things with their real colors (and free too to scrub off the awful black-and-white paint that covers her body). She steps outside her room into a garden full of flowers. "So, that is what it is like to experience red," she exclaims, as she sees a red rose. "And that," she adds, looking down at the grass, "is what it is like to experience green."

Mary here seems to make some important discoveries. She seems to find out things she did not know before. How can that be, if, as seems possible