

## Mary's Powers of Imagination

Amy Kind  
Claremont McKenna College  
akind@cmc.edu

*To be published in Sam Coleman, ed., Cambridge Classic Arguments series  
volume on the Knowledge Argument*

In contemporary discussions of philosophy of mind, Frank Jackson's story of Mary is a familiar one (Jackson 1982, 1986). Though Mary has spent her entire life locked inside a black and white room and has never seen color, she has nonetheless mastered color science. When she is released from her room and sees a ripe tomato for the first time, it seems that she learns something – that she comes to know what seeing red is like. But if Mary did not know all there was to know about color while in the room, and thus if there are facts over and above the physical facts about color, then physicalism is threatened. The standard physicalist responses to this argument, typically referred to as the *knowledge argument*, are also by now familiar. While some deny that Mary learns anything at all about color and color experiences upon her release from the room – how could she, given that she already has all the physical information there is to have about color and color experiences? – others acknowledge the power of the intuition that Mary has an “Aha!” moment when she exits the room and are slightly more concessionary. Such concessionary responses typically grant that Mary learns something upon her release but deny that this threatens physicalism. Rather, her learning should be understood to consist in the acquisition of abilities, or in her newfound acquaintance with color, or in her apprehension of an old fact under a new guise.

In this paper I want to focus on the first of these more concessionary responses – what's now generally known as *the ability hypothesis*. As originally proposed by Laurence Nemirow, the ability hypothesis claims that what Mary gains when she leaves the room is the ability to imagine seeing red (1980, 1990). A subsequent modification by David Lewis (1983, 1988) gives us what is now considered to be the standard version of the view: What Mary gains when she leaves the room is a cluster of abilities including not only the ability to imagine seeing red but also the abilities to recognize and recall seeing red. For proponents of the ability hypothesis, knowing what an experience is like is not propositional in nature but rather consists in the possession of the relevant ability or abilities.

Criticism of the ability hypothesis has tended to focus on this last claim, and considerable attention has been devoted to the question of whether the possession of such

abilities are either necessary or sufficient for knowledge of what an experience is like.<sup>1</sup> To my mind, however, this critical strategy grants too much. Focusing specifically on imaginative ability, I argue that Mary does not gain this ability when she leaves the room for she already had the ability to imagine red while she was inside it. Moreover, despite what some have thought, the ability hypothesis cannot be easily rescued by recasting it in terms of a more restrictive imaginative ability. My purpose here is not to take sides in the debate about physicalism, i.e., my criticism of the ability hypothesis is not offered in an attempt to defend the anti-physicalist conclusion of the knowledge argument. Rather, my purpose is to redeem the imagination from the misleading picture of it that discussion of the knowledge argument has fostered.

## I. What Can Mary Imagine?

David Hume famously said that nowhere are we more free than in the imagination, and the human capacity for imagination is typically taken to be almost without limit. We readily imagine all sorts of variations on things that we've encountered – from talking mice and golden geese to flying carpets and magic beans. And we also readily imagine all sorts of things that we've never actually encountered and that we don't even believe to exist, from ghosts and goblins to light sabers and time machines. Insofar as our imaginative capacities are in any way limited, those limits stem solely from the limits of possibility. On this point we might again look to Hume and, in particular, to what he referred to as “an established maxim in metaphysics,” namely, that “*that nothing we imagine is absolutely impossible.*” (1739/1985, I.ii.2) Since what we imagine contains within it “the idea of possible existence,” we are unable to imagine impossible entities like round squares.<sup>2</sup>

The uses to which we put our imagination also seem to be almost without limit.<sup>3</sup> We engage in imaginative exercises for purposes both playful and serious – in our engagement with fiction and games of pretense on the one hand, and in our decision-making and future planning on the other. In the more fanciful contexts, we tend to let our imagination run wild; in the more serious ones, we tend to rein it in. Sometimes our imagining is spontaneous, while sometimes it's deliberate. Sometimes our imagining is effortless, while sometimes it's hard work.

---

<sup>1</sup> See, e.g., Conee 1994, Alter 1998. Nida-Rümelin 2009 provides a useful summary of criticisms in this vein.

<sup>2</sup> That imagination has even this limit is sometimes disputed. To give just one example, Tamar Gendler's story of the Tower of Goldbach may suggest a way to imagine that  $7 + 5$  does not equal 12 (Gendler 2000). See also Kung 2016 and Tidman 1994.

<sup>3</sup> For a survey of some of the many uses of imagination see Kind 2016a, 7-10. See also Colin McGinn's argument that “imagination is a faculty that runs through the most diverse of mental phenomena” and that humans should be understood as *Homo imaginans* (2004, 5).

Surprisingly, many of these basic and uncontroversial claims about imagination seem to have been forgotten in philosophical discussion of Mary, and in particular, in philosophical discussion of the ability hypothesis. In arguing that Mary *gains* the ability to imagine the experience of seeing red when she leaves the room, proponents of the ability hypothesis presuppose that Mary *lacked* the ability to imagine the experience of seeing red while she was in her black-and-white room. Even opponents of the ability hypothesis tend to accept that Mary gains the ability to imagine the experience of seeing red when she leaves the room; what they question is whether this ability-gain adequately accounts for Mary's newfound knowledge of what seeing red is like. It is also generally assumed that there is nothing that Mary can do, while in the room, to gain the ability to imagine the experience of seeing red.<sup>4</sup> No matter what she does, and no matter how hard she tries, she will be unable to engage in this kind of imaginative exercise. What happened to the almost limitless human capacity for imagination?

We see a similar phenomenon – a phenomenon in which basic and uncontroversial claims about the imagination seem to have been forgotten – in connection with a different familiar example employed against physicalism: Thomas Nagel's bat. In developing his case that we can't know what it's like to be a bat, Nagel explicitly argues that our imagination will be no help. Bats navigate the world through echolocation, and they thus have perceptual experiences that are completely different in subjective character from any of the perceptual experiences that we have. For Nagel, the fact that echolocation is closed off to us experientially means that it is also closed off to us imaginatively:

Our own experience provides the basic material for our imagination, whose range is therefore limited. It will not help to try to imagine that one has webbing on one's arms, which enables one to fly around at dusk and dawn catching insects in one's mouth; that one has very poor vision, and perceives the surrounding world by a system of reflected high-frequency sound signals; and that one spends the day hanging upside down by one's feet in an attic. In so far as I can imagine this (which is not very far), it tells me only what it would be like for me to behave as a bat behaves. But that is not the question. I want to know what it is like for a bat to be a bat. Yet if I try to imagine this, I am restricted to the resources of my own mind, and those resources are inadequate to the task. I cannot perform it either by imagining additions to my present experience, or by imagining segments gradually subtracted from it, or by imagining some combination of additions, subtractions, and modifications. (Nagel 1974, 439)

These claims – that pre-release Mary can't imagine the experience of seeing red, that humans can't imagine the bat's experience of echolocation – have become deeply entrenched

---

<sup>4</sup> Though the agreement on this point is widespread it is not without exception. I discuss a couple of philosophers who take an alternative view of Mary's powers of imagination in Section II below.

in philosophical discussion. In fact, these claims are so entrenched that, having pointed out how they conflict with our ordinary picture of the imagination, I fear that some readers might be more inclined to give up this picture than to give up the claims themselves. Insofar as there is any such inclination, it may help to be reminded of the sorts of imaginative exercises that are commonplace in everyday life. As young children, we imagine ourselves to be all sorts of creatures big and small. Our arms become a crocodile's jaw, an elephant's trunk, a bird's wings. We imagine flying over the playground, or burrowing underground, or slithering through the grass. We imagine rocketing to the moon, meeting space aliens on Mars, or travelling through wormholes to distant galaxies. And our imaginings don't stop as we become older. Whether we're imagining winning the Superbowl or an Oscar, becoming President of the United States or being nominated for the Supreme Court, the joy of impending parenthood or the unpleasantness of impending chemotherapy, we continue to stretch our imaginative capacities to all sorts of events and scenarios that lie far beyond anything we've actually experienced.

Granted, we do sometimes say things that seem to suggest limitations on our imaginations beyond mere possibility. In everyday life it's not at all uncommon to hear someone start a sentence with, "I can't imagine..." Sometimes these claims are made when we're confronted with news of tragedy. When we read about a violent crime or a natural disaster, we might profess ourselves unable to imagine the trauma of the survivors, the courage of the victims, or the grief of their families. "I can't imagine how frightened they must have been," we might say when we hear about the children who hid under their desks during a school shooting. When thinking about the parents whose children were killed, we might profess ourselves unable to imagine what they're going through. But what do these claims of unimaginability really mean? Often, it seems that they are simply an expression of our own shock and horror – a way of saying that things are unbearably awful – rather than actual claims about the imagination. Insofar as they are claims about imagination, they will likely be claims not about what we can't do but about what we don't want to do. A refusal to imaginatively engage with these situations might help us to keep some of the horror at bay.

There's one other possibility about how to interpret such claims, however, and it's one that will bring us back to Mary. When certain events and experiences are very foreign from our own, or when they strike us as inexplicable in some key way, we might be inclined to deny that we can imagine them because we don't think we can imagine them *correctly*. There are two ways to understand this suggestion. On the first way, ordinary claims of the form "I can't imagine such-and-such" are shorthand for claims of the form "I'm doing a bad job imagining such-and-such." In this case, they're not really about what we can or cannot imagine but about what we can or cannot imagine well. On the second way of understanding this suggestion, however, such reports are not mere shorthand for claims about correct imaginings but rather follow from them. When we find ourselves unable to imagine something well, when we lack

confidence that our imagining gets it right, that would mean that we're not really imagining it at all.

This analysis would explain why we might think Mary, who has never had any color experiences whatsoever, lacks the ability to imagine red while in her black and white room. No matter how gifted an imaginer Mary is, it seems pretty implausible – at least upon first thought, though we'll return to this question below – that prior to her release she's able to imagine the experience of seeing red correctly. So if imagining an experience requires imagining an experience correctly, then pre-release Mary would lack the ability to imagine the experience of seeing red after all.

What can be said in favor of this analysis, i.e., in favor of the claim that a failure to imagine *well* amounts to a failure to imagine. Here we might compare a related ability, that of memory. Suppose I try to remember how I felt when I got a particular piece of unexpected news. If I now claim to remember being relieved but at the time I was not relieved but disappointed, then my claim is false. Misremembering – or at least radical misremembering – is not really remembering at all; it is at best *seeming* to remember. So perhaps imagining works like remembering, where misimagining – or at least radical misimagining – is not really imagining at all.

Unfortunately, this line of defense proves quickly to be a dead end. On our ordinary picture of imagining, imagining is not like remembering. Though an act is disqualified as remembering in virtue of embodying mistakes (or at least radical mistakes), an act is not disqualified as imagining in virtue of embodying mistakes. In some instances of imagining, of course, the notion of mistake has no purchase. What would it mean for a child to be mistaken when imagining a scary monster or a fairy godmother? But even when we can relevantly talk about mistakes in imaginings – in cases, for example, where we're trying to get things right – the mistakes do not cause the kind of problems for imagining as they do for remembering. A newborn baby might have considerably more hair than his mother imagined. The collector's edition Millennium Falcon replica bought on eBay might be considerably smaller than the purchaser imagined. The newly painted living room might be considerably darker than the homeowner imagined. Yet all of these mistaken imaginings are still imaginings.

So if mistakes do not keep an imagining from being an imagining, then the fact that Mary's imagining of the experience of seeing red will almost surely contain mistakes does not give us grounds to deny that Mary has the ability to imagine this experience. But perhaps there is a related move in the vicinity that's more promising. Though the mistakes do not keep her act from being an imagining, perhaps they keep it from being an imagining *of the experience of seeing red*. Though Mary is trying to imagine the experience of seeing red, and though she takes herself to be imagining seeing red, she's really only imagined some shade of gray.

Here too, however, the line of thought is problematic and, again, the problem becomes obvious when we reflect on ordinary imaginings and the way we think about them. The fact that her infant has more hair than the new mother imagined does not mean that during the nine months of her pregnancy she'd been failing to imagine him. "You're not at all like I imagined," she might coo at him – but not: "Wow, I was imagining some different baby altogether." In general, the fact that an imagining mischaracterizes its target does not mean that it misses its target altogether.<sup>5</sup>

In fact, as philosophers working on imagination have long recognized, the connection between an imagining and its target is an especially tight one. Consider the following passage from Jean-Paul Sartre:

[T]he imaged cube is given immediately for what it is. When I say 'the object I perceive' is a cube, I make a hypothesis that the later course of my perceptions may oblige me to abandon. When I say 'the object of which I have an image at this moment is a cube', I make here a judgement of obviousness: it is absolutely certain that the object of my image is a cube. (Sartre 1940/2010, 9)

As Sartre here suggests, and in line with a general consensus in philosophical discussion of imagination, imaginers enjoy a certain epistemic privilege with respect to their imaginings: they cannot be mistaken about what they are imagining. One way to explain this privilege is in terms of the intentions with which we undertake an imaginative project; as Colin McGinn notes, "the identity of my imagined object is fixed by my imaginative intentions, to which I have special access." In an imagining, the imaginative object is "*given*, not inferred. I know that my image is of my mother because I *intended* it to be; I don't have to consult the appearance of the person in the image and then infer that I must have formed an image of my mother." (McGinn 2004, 5) Indeed, this is one of the many aspects of imagining that differentiates it from perceiving. Though sheer force of will cannot enable me to perceive something, it can enable me to imagine something.

As Nagel notes in the passage quoted earlier, the resources of our own minds are limited, and so we don't have much to go on when we set ourselves the project of imagining what it's like to be a bat. And he is also right that there is an important sense in which our imaginative efforts in this regard will be a failure, namely, that they will fail to teach us what we want to know – they will not teach us what it's like to be a bat. But this instructive failure does not in itself mean that we lack the ability to imagine what it's like to be a bat. And likewise for Mary. Pre-release, Mary's imaginings will not teach her what it is like to see red. But that does

---

<sup>5</sup> For more on the difference between imaginings that miss their targets and imaginings that mischaracterize their targets, see Kind 2016b.

not mean that she lacks the ability to imagine the experience of seeing red while she is in her black-and-white room.

The upshot for the ability hypothesis should be clear: in its traditional form, at least, the hypothesis is false. Mary does not gain the ability to imagine the experience of red when she leaves the room, for this is an ability that she already has inside the room. But here it would be natural to think that the foregoing discussion has revealed a solution to the problem in the course of revealing the problem itself. Having alerted us to the importance of distinguishing the ability of imagining from the ability of imagining correctly, the discussion suggests that we re-formulate the ability hypothesis in terms of the latter ability rather than the former. On this suggestion, then, what Mary gains when she leaves the room is not the ability to imagine the experience of seeing red but rather the ability to imagine this experience correctly.

In the remainder of this paper, I turn to an evaluation of this suggestion. My discussion proceeds in two stages. First, in Section II, I'll consider some particular attempts to revise the ability hypothesis to take account of the distinction between imagining and imagining correctly. As we'll see, however, such revision proves to be more difficult than one might think. Thus, in Section III, I'll turn more directly to the notion of imagining correctly. What does it mean to imagine correctly, and what are Mary's powers in this regard? Ultimately, our discussion will suggest that the ability hypothesis fares no better by employing the notion of correct imagining than it does by employing the notion of imagining.

## **II. Recasting the Ability Hypothesis**

As we've seen, the standard version of the ability hypothesis in play in philosophical discussion proposes that Mary gains three abilities upon her release: the ability to imagine, to recognize, and to remember the experience of seeing red. This is how the view is typically presented in encyclopedia articles and other overviews (see, e.g., Nida-Rümelin 2009; Stoljar and Nagasawa 2004), and it's also how Jackson himself has presented matters after abandoning his defense of the knowledge argument and announcing his allegiance to the ability hypothesis instead (Jackson 2004, 439).<sup>6</sup> As we've also seen, casting the ability hypothesis in these terms is problematic, since there's good reason to think that Mary already has the ability to imagine the experience of seeing red while she's inside the room. Though this problem is not widely recognized, there have been a few instances – I'll here mention three in particular – in which philosophers writing on the knowledge argument have explicitly attempted to avoid this problem by invoking a distinction between the ability to imagine and the ability to imagine

---

<sup>6</sup> Though Jackson has now adopted the ability hypothesis, his main reasons for repudiating the knowledge argument stem from considerations about representationalism.

correctly.<sup>7</sup> Though all three of these philosophers agree that the ability hypothesis will only be plausible if it is cast in terms of the ability to imagine correctly, they differ in how to spell out the notion of correctness.

For D.H. Mellor, who takes up the issue in his “Nothing Like Experience” (1993), imagining something correctly entails being able to recognize it.<sup>8</sup> Like the proponents of the ability hypothesis, Mellor is concerned to give an analysis of our knowledge of what experience is like, but unlike such proponents, he is not concerned to defend physicalism. As Mellor argues, an imagining cannot constitute knowledge of what an experience is like if that experience is imagined incorrectly: “For we can imagine experiences wrongly, and then we do not know what they are like. Someone who imagines sugar to taste like salt, for example, does not know what sugar tastes like.” Thus, he argues, “To know what experiences of a certain kind are like I must, when I imagine them, imagine them correctly, i.e., in a way that makes me recognise them when I have them. (Mellor 1993, 4-5)

Paul Noordhof (2003) agrees with Mellor that incorrect imagining cannot constitute knowledge of what an experience is like, but he offers a slightly different analysis of what it is to imagine an experience correctly. For Noordhof, what matters is not that I be able to *recognize* the experience but that I be able to *identify* it (Noordhof 2003, 23). Recognition, Noordhof worries, seems to imply previous cognition: “I can only recognize things if I have come across them, a picture of them, or a description of them, some time before.” In his view, this threatens to make the ability hypothesis trivial: Rather than offering us some insight into why experience of a certain sort is necessary for knowledge of what such an experience is like, such an analysis simply presupposes what it sets out to explain. He thus prefers to cash out correct imagining in terms of identification, which carries no implication of previous cognition.<sup>9</sup>

In a more recent discussion, Bence Nanay raises a problem that would seem to apply to both of these formulations. At least some people who become blind late in life claim that they can still imagine the experience of seeing red, and presumably they can do so correctly (for discussion, see Sacks 2003). But since such an individual no longer has recognitional capabilities with respect to this experience – “they are most certainly incapable of recognizing red, given that they are blind” (Nanay 2009, 704) – it would be a mistake to cash out correct imagining in terms of recognition. A similar point seems to apply to identification. The accounts of correct imagination offered by Mellor and Noordhof are thus problematic, and

---

<sup>7</sup> It's perhaps not surprising that two of the philosophers to notice this problem – Paul Noordhof and Bence Nanay – have published extensively on imagination.

<sup>8</sup> Though Mellor mentions the Mary case, he is primarily concerned with Nagel's discussion of what it's like to be a bat.

<sup>9</sup> Insofar as Mellor takes recognition to imply only lack of surprise and not previous cognition, Noordhof's account might not be a dramatic departure from Mellor's – as Noordhof himself admits. See Noordhof 2003 fn5.



correspondingly, so too are their versions of the ability hypothesis in which these accounts are embedded.

In an attempt to rectify this problem, Nanay offers an alternate understanding of what it means to have the ability to imagine an experience E correctly: one has this ability when one can imagine the experience “in such a way that would enable one to distinguish imagining experience E from imagining or having any other experience.” (Nanay 2009, 705). Since blind imaginers remain capable of distinguishing their imaginings from one another, Nanay’s account allows that such imaginers can imagine color experiences correctly and thus protects itself from the objection that he posed to the kind of account offered by Mellor and Noordhof.

But in doing so, Nanay opens himself up to a different objection. To see this, it will help to look at how his analysis of correct imagining gets incorporated into his new version of the ability hypothesis. Like the original ability hypothesis, Nanay’s revised version aims to provide an explanation of knowing what an experience is like:

Knowing what it is like to experience E is having the ability to distinguish imagining or having experience E from imagining or having any other experience.

Nanay’s account is disjunctive in nature, and it’s precisely this feature that protects him against what might otherwise seem to be objections or counterexamples. On his view, each of the abilities mentioned is individually sufficient for knowing what an experience is like but they are not individually necessary; rather, what’s necessary for an individual to know what experience E is like is that she have at least one of the two abilities. Thus, the fact that an individual may lack one of these abilities while still having knowledge of what an experience is like will not count against Nanay’s version of the ability hypothesis.

My worry about the account, however, stems from a slightly different direction. Surprisingly, perhaps, my worry about the account does not concern its adequacy. For our purposes here, I’m happy to accept for the sake of argument that Nanay has proposed a version of the ability hypothesis that offers an adequate account of knowledge of what an experience is like. To my mind, however, it would be a mistake to treat this version of the ability hypothesis as one which assigns any real importance to imaginative ability. As I read the account above, not only does it shift away from the ability to imagine but it also shifts away from the ability to imagine correctly. What’s central to Nanay’s ability hypothesis are not imaginative abilities but rather *discriminative abilities*.

Granted, Nanay takes himself to be invoking these discriminative abilities in an effort to explain what correct imagining is. But while an ability to discriminate might be essential for one’s knowledge of what an experience is like, it’s hard to see why it should be essential for the correctness of an imagining. What makes an imagining correct or not has to do with how well it

matches its target. And just as a drawing might correctly capture its target without enabling someone who sees it from distinguishing that target from all other possible targets, so too an imagining might correctly capture its target without enabling an imaginer to distinguish that target from all other possible targets.

It may help to consider things this way. As stated, Nanay's version of the ability hypothesis doesn't require anything about Mary's imaginative abilities to change when she exits the room. She might be producing exactly the same imaginings once she exits the room that she was producing while she was inside the room. What's changed once she leaves the room is that she is now able to distinguish this imagining from all other imaginings and experiences; that's what her knowledge of what it's like consists in.

At this point, it might seem that things have somehow gone astray. One of the key intuitions motivating the ability hypothesis seemed to have to do with imaginative ability. It seemed plausible that we could explain Mary's newfound knowledge of what the experience of seeing red is like at least partially in terms of an imaginative gain. But we've now ended up with a version of the ability hypothesis that – perhaps inadvertently – requires us to jettison that intuition. (Indeed, though I've been assuming for the sake of argument that such a version of the ability hypothesis might well be an adequate account of knowledge of what an experience is like, once we realize how far we've shifted away from the original ability hypothesis, this assumption now starts to seem questionable.) So where do we go from here?

### **III. What Can Mary Imagine Correctly?**

As we saw in Section I, the ability hypothesis cannot be specified in terms of the ability to imagine the experience of seeing red, because on any plausible understanding of that ability, Mary already has it inside the room. In uncovering this problem, however, our discussion led to what looked like a promising alternative: Instead of relying on the ability to imagine, the ability hypothesis might instead rely on the ability to imagine correctly. Indeed, as we saw in Section II, this alternative has been pursued by several proponents of the ability hypothesis. Now that we have explored their proposals, however, it has seemed that the alternative has failed to live up to its promise. In particular, it proved difficult to flesh out the notion of correct imagining in such a way as to save the ability hypothesis. As a result, it appears that we've ended up with a version of the ability hypothesis that, rather than developing the notion of correct imagining, instead seems to abandon it altogether.

These difficulties notwithstanding, one might still have the sense that there is something importantly right about explicating the ability hypothesis in terms of correct imagining. Perhaps the problem with the previous proposals was not the notion itself but rather the specific ways

of fleshing it out. If this is the case, then we might be better served simply by relying on an intuitive understanding of it. Might such a version of the ability hypothesis work? Can such an intuitive understanding do the philosophical work that's needed? It's these questions that will motivate the discussion of this section. Rather than addressing them directly, however, I propose to come at them somewhat sideways. Just as we've previously explored the question of what pre-release Mary can imagine, here I want to explore the question of what pre-release Mary can imagine correctly.

To start, it will be helpful to recall a basic point about imagination that we encountered in Section I: Human imaginers typically imagine all sorts of things that lie beyond anything we've actually experienced, in some cases *far* beyond anything we've actually experienced. Let's call these *distant* imaginings. Of course, whether an imagining counts as distant or not will vary from imaginer to imaginer. For those of us with normal color vision who haven't been locked in a black-and-white room all our lives, imagining the experience of seeing red would not be a distant imagining at all. It would be more distant for someone with red-green color blindness, and yet even more distant for Mary.

As this suggests, not only is the matter of distance relative to the experience one has, but it is also a matter of degree. Distant imaginings lie on a continuum – what I'll call the *distance continuum* – from the not so distant to the very, very distant. Consider the fact that trained musicians can look at an unfamiliar music score and imagine correctly the experience of hearing the piece played, or that trained chefs can look at an unfamiliar recipe and imagine correctly the experience of tasting the cooked dish. Though these acts might naturally be characterized as distant imaginings – after all, the musician has never before heard this composition and the chef has never before tasted this dish – there's also a sense in which they're not *very* distant. As Lewis notes in discussing sightreading, “new music isn't altogether new – the big new experience is a rearrangement of lots of little old experiences,” and presumably the same holds true for new dishes (501 in Lycan). In his view, our capacity for correct imagining gives out when an experience is “new enough.” As he admits, the question of when an experience is new enough – or, to put it in my terms, when an experience is sufficiently distant – is a very hard one to answer.

That said, almost everyone seems to agree that whatever the answer, that however we draw the line between experiences that are not too distant to be imagined correctly and experiences that are, the experience of seeing red is for Mary on the far side of the line. Note here that that I said *almost* everyone. There are two notable exceptions. In discussing the knowledge argument, Daniel Dennett has consistently denied that Mary learns anything at all when she leaves the room (see, e.g., Dennett 1991, 398ff). Given her knowledge of a completed color science, there is no “aha” moment when she sees a ripe tomato for the first

time. If we tried to trick her by showing her a blue tomato, she would not be fooled. Though Dennett's original responses to the knowledge argument do not directly address the issue of Mary's imaginative capacities, his staunch insistence that Mary already knows what the experience of seeing red is like while she's inside the room suggests that he would also insist she has the ability to imagine the experience of seeing red correctly while she is still inside the room. In a recent paper, he has been more explicit on this point:

We are told that Mary in her cell can't imagine what it's like to experience red, try as she might. But suppose she doesn't accept this limitation and does try her best, cogitating for hours on end, and one day she tells us she just got lucky and succeeded. "Hey," she says, "I was just daydreaming, and I stumbled across what it's like to see red, and, of course, once I noticed what I was doing I tested my imagination against everything I knew, and I confirmed that I had, indeed, imagined what it's like to see red!" (Dennett 2007, 23)

As Dennett notes, if we subsequently tested her by showing her various color samples and she passed the test, why wouldn't we conclude that she could correctly imagine the experience of red while she was inside the room?

Paul Churchland has long pushed a similar line. On his view, pre-release Mary has the ability not only to imagine having the experience of red but also the ability to imagine it correctly. Like Dennett, Churchland stresses just how much Mary knows. Given this knowledge, Mary may begin to reconceptualize her inner life:

So she does not identify her visual sensations crudely as "a sensation-of-black", "a sensation-of-grey", or "a sensation-of-white"; rather she identifies them more revealingly as various spiking frequencies in the *n*th layer of the occipital cortex (or whatever). If Mary has the relevant neuroscientific concepts for the sensational states at issue (viz., sensations-of-*red*), but has never yet been *in* those states, she may well be able to imagine being in the relevant cortical state, and imagine it with substantial success, even in advance of receiving external stimuli that would actually produce it. (Churchland 1985, 25-6)

For our purposes here, there are two ways we might take these remarks. On the one hand, Churchland might be granting that the experience of red is, for Mary, a very distant one but trying to show that even this very distant experience can still be correctly imagined. On the other hand, Churchland might be denying that the experience of seeing red is as distant for Mary as we might have thought. To my mind, Churchland is most plausibly read as offering this second kind of proposal. It's not that we were wrong about where to draw the line between

experiences that are not too distant to be imagined correctly and experiences that are, it's that we were wrong about where color experiences lie for Mary on the distance continuum.

Recall again the sightreading example from above, i.e., that trained musicians can sightread scores they've never heard before. Extending this kind of example, Churchland notes that musicians with sufficient training can identify the individual notes of a chord they're hearing for the first time, and conversely, can auditorily imagine an unfamiliar chord from the specification of the notes. Such imaginative feats are possible in virtue of the fact that chords are structured sets of elements, i.e., in virtue of the fact that even new and unfamiliar musical experiences of chords are not that distant. If color sensations are likewise structured, then new and unfamiliar color experiences would be considerably less distant than we'd initially thought. Though color experiences seem to us to be undifferentiated wholes, the same holds true – at least initially – for chord experiences. So, asks Churchland, “Why should it be unthinkable that sensations of color possess a comparable internal structure, unnoticed so far, but awaiting our determined and informed introspection?” (1985, 26-7), i.e., the sort of informed introspective abilities that one might develop were one to master completed neuroscience.

It might help to cast the basic issue here in terms of *scaffolding*. One reason that color experience seems to be so distant for Mary is that there seems to be no way to get to them from the kinds of experiences she has. If color experience is structured, however – and in particular, if it has some kind of phenomenal structure – then that structure could provide a way for Mary to scaffold out from the experiences she has to those that she hasn't. Moreover, this scaffolding would provide Mary not only a way to imagine the unfamiliar experiences but a way to imagine them correctly. More generally, scaffolding also provides us with a useful way to think about the distance continuum.

Churchland's point in raising these considerations is slightly different from ours. Rather than putting them forth to criticize the ability hypothesis, he means to be criticizing the knowledge argument more directly. As a result, he doesn't intend the story he's spun about Mary to show that she actually has the ability to correctly imagine the experience of red while inside the room; rather, the considerations he offers are designed to show simply that *it's not unthinkable* that she has this ability, that her having this skill is not “beyond all possibility” (Churchland 1985, 26). But regardless of whether such considerations are enough to defuse the knowledge argument – a question that it would take us too far afield to settle – it's clear that such considerations are not quite enough for our purposes here. Our interest is in determining what Mary can and can't imagine correctly, in how to draw the line between imaginings that are not too distant to be imagined correctly and those that are. Though it would probably have been too much to expect that we'd be able to be provided with decisive answers to these questions, we might reasonably have hoped for a slightly stronger case one way or the other.

So can we do any better? Here I think we are helped by reflecting upon other cases of relatively distant experiences where correct imagining seems to be within our grasp. To take just one example, consider Mark Haddon's *The Curious Incident of the Dog in the Nighttime*. The story is set in motion when Christopher, the 15-year narrator of the novel, decides to investigate the murder of his neighbor's dog. But Christopher is no ordinary teenager. He's mathematically gifted and has highly developed powers of visual thinking but is socially awkward – physically lashing out at others when he becomes uncomfortable. He hates the colors brown and yellow, refuses to tell jokes or use metaphors, and hides in small spaces when he's frightened. Though Haddon never explicitly uses the terms "autism" or "Asperger's Syndrome" in the course of the book, and deliberately so, Christopher's behavioral difficulties suggest that he lies somewhere on the autistic spectrum.<sup>10</sup>

Haddon, who does not have autism, has been widely praised for his ability to get inside the mind of someone with this condition. The book has been described as a "triumph of empathy," as brimming with "with imagination, empathy, and vision," and as "flawlessly imagined and deeply affecting."<sup>11</sup> The experience of having autism differs from person to person, but Haddon seems to have successfully imagined what it's like to be one such person. (And, moreover, he seems to have enabled his readers to imagine it as well.) Though this imagining is less distant than imagining what it's like to be a bat, say, for people who do not have autism it is still a case of relatively distant imagining.

A different example takes us even closer to the bat case. Though Nagel claims we cannot know what it is like to be a bat, animal scientist Temple Grandin claims that she can adopt a "cow's eye view" of a situation and thereby, in at least a limited sense, know what it is like to be a cow. Like Christopher, the fictional narrator of Haddon's novel, Grandin has autism. Also like Christopher, Grandin describes herself as a visual thinker, and she credits her visual thinking with enabling her "to build entire systems" in her imagination. Her visual thinking is also what enables her to take up a cow's perspective:

When I put myself in a cow's place, I really have to be that cow and not a person in cow costume. I use my visual thinking skills to simulate what an animal would see and hear in a given situation. I place myself inside its body and imagine what it experiences. It is the ultimate virtual reality system. (2006, 168)

---

<sup>10</sup> The term "Asperger's" was used on the book's cover, and both terms appear in various promotional materials for the book that have been put out by the publisher. Haddon discusses his own reluctance to use these terms and his views about how best to describe Christopher at <http://www.markhaddon.com/aspergers-and-autism>.

<sup>11</sup> The first quotation is from a review in *The New Yorker*, the second is an endorsement by author Myla Goldberg that's printed on the back cover of the book, and the third quotation is from a review in *Time Out New York*. All quotations can be found in the editorial reviews included on the Amazon page for the book, [http://www.amazon.com/Curious-Incident-Dog-Night-Time/dp/1400032717?ie=UTF8&ref\\_=asap\\_bc](http://www.amazon.com/Curious-Incident-Dog-Night-Time/dp/1400032717?ie=UTF8&ref_=asap_bc).

Throughout her career, Grandin has revolutionized the handling of livestock with her innovative equipment designs. Facilities using the equipment that she's designed report that the animals are considerably more comfortable, cooperative, and calm than they were previously. Her tremendous success suggests not only that Grandin managed to imagine what it is like to be a cow but that she has managed to do so with at least some degree of correctness.

Reflection on examples such as these helps to shed further light on Mary's powers of imagination. To my mind, these sorts of examples continue the work begun by Churchland's considerations of eroding a knee-jerk skepticism that Mary could imagine color experiences correctly. Granted, such considerations do not go so far as to show definitively that Mary does have the ability to imagine correctly the experience of seeing red inside the room. After all, the examples that we've considered seem to involve experiences that are plausibly less distant for the imaginer than color experiences are for Mary. But even though there still seems reason to doubt that pre-release Mary can correctly imagine color experiences, there is an important moral to be drawn from the examples just considered. In particular, they seem to pose something of a dilemma for a proponent of the ability hypothesis who wants to employ the notion of correct imagining.

If Haddon can correctly imagine what it's like to be autistic, and Grandin can correctly imagine what it's like to be a cow, then that means that many experiences that are fairly far along the distance spectrum can be correctly imagined. In utilizing the notion of correct imagining, then, the proponent of the ability hypothesis has to walk a very careful tightrope. Too restrictive an interpretation of the notion of correct imagining – that is, one that makes correct imagining very hard to achieve – runs the risk of ruling out these cases. But too permissive an interpretation of the notion of correct imagining – that is, one that makes correct imagining very easy to achieve – runs the risk of ruling in Mary. Of course, one might try to achieve the appropriate balance by connecting correct imagining to knowledge of what it's like, e.g., an imagining is correct if and only if it coheres with the imaginer's knowledge of what an experience is like. But this kind of analysis deprives the notion of correct imagining of any usefulness in a version of the ability hypothesis. One can't usefully invoke the notion of correct imagining in one's analysis of what knowledge of an experience when one has relied on knowledge of what an experience is like to distinguish between correct and incorrect imagining. Thus, it starts to look like there is very little room to flesh out the notion of correct imagining in a non-question begging way such that it allows for imaginings like Haddon's and Grandin's to be correct without also allowing imaginings like Mary's to be correct.

Faced with this problem, one might be tempted to deny that the imaginings by Haddon and Grandin are really correct after all. But such a denial is not likely to help a proponent of the ability hypothesis who wants to employ correct imagining. This notion, remember, is meant to

go along with knowledge of what an experience is like. Such knowledge is meant to consist, at least in part, in this imaginative ability. But once we adopt a very restrictive conception of correct imagining, a conception that makes correct imagining out of reach not just for pre-release Mary with respect to color experience and us with respect to bats but for all sorts of other experiences as well, this imaginative ability starts to seem disconnected with – and even more restrictive than – knowledge of what an experience is like. Ultimately, it looks increasingly unlikely that the notion of correct imagining will be able to bear the philosophical weight that proponents of the ability hypothesis want to place on it.

We might put this point in terms of a dilemma for the proponent of the ability hypothesis – or at least for such a proponent who wants to invoke imaginative abilities. If the ability hypothesis is explicated in terms of a *restrictive* notion of correct imagining, then though it will turn out to be true that pre-release Mary can't correctly imagine red, we will also have to deny that Haddon can correctly imagine what it's like to be autistic or that Grandin can correctly imagine what it's like to be a cow. And this seems like the wrong result. Moreover, since it seems plausible that Haddon (and his readers) do have at least some knowledge or what being autistic is like, and likewise for Grandin with respect to the cow's experience, we'd also end up with cases where it looks like someone can know what an experience is like despite being unable to imagine it correctly. On the other hand, if the proponent of the ability hypothesis tries to avoid this problem by employing a weaker notion of correct imagining, then it looks like pre-release Mary's imaginings of color experiences will end up counting as cases of correct imagining. In short, either the ability hypothesis gets the intended result about Mary but relies upon an implausibly strict account of correct imagining, or it relies on a more plausible analysis of correct imagining but doesn't get the intended result about Mary.<sup>12</sup>

#### **IV. Concluding Remarks**

The argument of this paper poses a problem for the ability hypothesis. But it's important to be clear about the force of the considerations I've raised. In particular, nothing that I've said here shows that no version of the ability hypothesis might be usefully employed to defend physicalism from the knowledge argument. What I have shown, however, is that any such version of the ability hypothesis will have to be developed without reliance on imaginative abilities.

In a sense, however, this paper hasn't really been about the ability hypothesis at all – or at least, not just about it. Rather, the true motivation of this paper has been to rescue imagination from some misunderstandings that have arisen in the context of the knowledge

---

<sup>12</sup> Thanks to Sam Coleman for suggesting the dilemma formulation.



argument. Insofar as the discussion of this argument – and, in particular, discussion of the ability hypothesis – has promoted the impression that Mary inside the room cannot exercise her powers of imagination, that she can't imagine color experience, it has promoted (even if inadvertently so) a misleading picture of imagination, a picture that paints imagination as considerably more limited than it in fact is. Moreover, insofar as the discussion has conflated the ability to imagine with the ability to imagine correctly, this misleading picture of imagination and its limits has been exacerbated.

There's an even deeper respect in which discussion of the knowledge argument has had pernicious consequences for our understanding of imagination. The issue that I have in mind was hinted at in the discussion of scaffolding above. In brief, the assumption that underlies the ability hypothesis is that knowledge of what an experience is like consists in an imaginative ability. When Mary comes to see red for the first time, she thereby comes to know what the experience of seeing red is like, and what this means is that she can now imagine things that she couldn't before. Her having the imaginative ability is simply part and parcel of her having knowledge of what the experience is like; that's what such knowledge *is*. But I'm inclined to think that this way of thinking of the relationship between our imaginative abilities and our knowledge gets things importantly wrong.<sup>13</sup> On this conception of imagination, we can't learn from our imagination. Imagining – or even correct imagining – doesn't provide us with new knowledge, or new understanding. How could it? After all, it's what that understanding consists in. And in fact, we have such understanding even if we've never exercised our imagination.

In this way, discussion of the knowledge argument threatens to obscure the epistemic relevance of imagining.<sup>14</sup> It threatens to obscure the sense in which our imaginings can teach us things that we didn't already know. Given the importance of imagination in everyday life, and given the widespread reliance on imagination in various philosophical contexts, this misrepresentation of the epistemic utility of imagination is one that is well worth avoiding.<sup>15</sup>

## References

Alter, Torin. 1998. "A Limited Defense of the Knowledge Argument." *Philosophical Studies* 90: 35-56.

Alter, Torin. 2001. "Know-How, Ability, and the Ability Hypothesis." *Theoria* 67: 229-239.

---

<sup>13</sup> For related worries about the ability hypothesis, see Coleman 2009.

<sup>14</sup> I defend the epistemic relevance of the imagination in Kind 2016c and Kind forthcoming.

<sup>15</sup> I am grateful to Torin Alter, Sam Coleman, and Frank Menetrez for helpful comments on previous drafts of this paper.

- Churchland, Paul. 1985. "Reduction, Qualia, and the Direct Introspection of Brain States." *Journal of Philosophy* 82: 8-28.
- Coleman, Sam. 2009. "Why the Ability Hypothesis Is Best Forgotten." *Journal of Consciousness Studies* 16 : 74-97.
- Conee, Earl. 1994. "Phenomenal Knowledge." *Australasian Journal of Philosophy* 72: 136-150.
- Dennett, Daniel. 2007. "What RoboMary Knows." In Torin Alter and Sven Walter, eds., *Phenomenal Concepts and Phenomenal Knowledge*, 15-31. Oxford: Oxford University Press.
- Dennett, Daniel. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Gendler, Tamar Szabó. 2000. "The Puzzle of Imaginative Resistance." *Journal of Philosophy* 97 (2):55-81.
- Grandin, Temple. 2006. *Thinking in Pictures: My Life With Autism*, second edition. New York: Vintage Books.
- Hume, David. 1739/1985. *A Treatise of Human Nature*. Edited by P.H. Nidditch. Oxford: Oxford University Press.
- Jackson, Frank. 1982. "Epiphenomenal Qualia." 1982. *Philosophical Quarterly* 34: 147-152.
- Jackson, Frank. 1986. "What Mary Didn't Know." *Journal of Philosophy* 83: 291-295.
- Jackson, Frank. 2004. "Mind and Illusion." In Peter Ludlow, Yujin Nagasawa, and Daniel Stoljar, eds., *There's Something About Mary*, 421-442. Cambridge, Mass.: The MIT Press.
- Kind, Amy. 2016a. "Introduction: Exploring Imagination." In Amy Kind, ed., *The Routledge Handbook of Philosophy of Imagination*, 1-11. London: Routledge.
- Kind, Amy. 2016b. "The Snowman's Imagination." *American Philosophical Quarterly* 53: 341-348.
- Kind, Amy. 2016c. "Imagining Under Constraints." In Amy Kind and Peter Kung, eds., *Knowledge Through Imagination*, 145-159. Oxford University Press.
- Kind, Amy. Forthcoming. "How Imagination Gives Rise to Knowledge," in Fabian Dorsch and Fiona Macpherson, eds., *Perceptual Memory and Perceptual Imagination*. Oxford University Press.
- Kung, Peter. 2016. "You Really Do Imagine It: Against Error Theories of Imagination." *Noûs* 50: 90-120.

- Lewis, David. 1983. "Postscript to 'Mad Pain and Martian Pain'." In his *Philosophical Papers*, Vol. 1., 130-132. New York: Oxford University Press.
- Lewis, David. 1988. "What Experience Teaches." *Proceedings of the Russellian Society* 13: 29-57.
- McGinn, Colin. 2004. *Mindsight: Image, Dream, Meaning*. Cambridge, MA: Harvard University Press.
- Mellor, D.H. 1992/1993. "Nothing Like Experience." *Proceedings of the Aristotelian Society* 93:1-16. j nnn
- Nagel, Thomas. 1974. "What Is it Like To Be a Bat?" *Philosophical Review* 83: 435-450.
- Nanay, Bence. 2009. "Imagining, Recognizing and Discriminating: Reconsidering the Ability Hypothesis." *Philosophy and Phenomenological Research*. 89: 699-717.
- Nemirow, Laurence. 1980. Review of *Mortal Questions*, by Thomas Nagel. *The Philosophical Review* 89: 473-477.
- Nemirow, Laurence. 1990. "Physicalism and the Cognitive Role of Acquaintance." In William Lycan, ed., *Mind and Cognition: A Reader*, 490-499. Oxford: Blackwell.
- Nida-Rümelin, Martine. 2009. "Qualia: The Knowledge Argument." In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*.  
<http://plato.stanford.edu/archives/sum2015/entries/qualia-knowledge/>
- Noordhof, Paul. 2003. "Something Like Ability." *Australasian Journal of Philosophy* 81: 21-40.
- Sacks, Oliver. 2003. "The Mind's Eye: What the Blind See." *The New Yorker*. July 28, 48-59.
- Sartre, Jean-Paul. 1940/2010. *The Imaginary*. Translated by Jonathan Webber. London: Routledge.
- Tidman, Paul. 1994. "Conceivability as a Test for Possibility." *American Philosophical Quarterly* 31: 297-309.