

2. Mary and Max and Jack and Ned

John Perry

INTRODUCTION

There is more in Ned Block's rich chapter than I can discuss with the allotments of time, space, and, especially, wit and acumen, that various higher authorities have allotted me. I confine myself to replying to some of his criticisms of my treatment of the knowledge argument in *Knowledge, Possibility and Consciousness* (Perry 2001), mainly by restating my view in ways that connect with at least some of his reservations.

Pace Block, the knowledge argument is about knowledge. In Frank Jackson's classic statement it is a simple and gripping three-step argument. Mary has *new* knowledge when she steps out of the black and white room and sees a red fire hydrant. But while in the black and white room she knew *all* the physical facts relevant to color vision. Conclusion: her new knowledge is of a non-physical fact. It is Mary's new knowledge that is the crucial step. Some physicalists deny that she has new knowledge. I do not deny this. Instead, I offer an account of Mary's new knowledge that is consistent with the identity of qualia and physical brain states. There is no more direct way to confront the knowledge argument.

MARY

Mary thinks something like this: "The type of *this*; color experience is what it is like (for me now) to see the color of *that* fire hydrant; that fire hydrant is red; I am normal and conditions are normal, so this type of experience is what it is like (for people with normal vision in normal conditions) to see red; that is, the type of *this*; color experience is Quale_{RED}."¹ It is the relation between types of color experiences and

¹ Ordinarily we feel free to report knowledge that a subject would express with indexicals in indirect discourse. I will, however, use quotation marks around a sentence that Mary

colors that is crucial, at least in the original form of the argument, for this is what Mary, it seems, could have already known in the black and white room, if qualia were physical states of the brain.²

Block wants us to concentrate on what Mary knows when she is actually having the experience; in Perry (2001) I gave an account that also covered the knowledge Mary retains after having had the experience, but I'm happy to set that aside.

Mary could have known "The type of *this*; color experience is what it is like (for me, in these conditions) to see the color of *that* fire hydrant"; even if her belief that the color of *that* fire hydrant is red was false, perhaps because some trickster had painted all the local fire hydrants green. She believes that there is nothing special about her color vision at the present moment, that there is nothing special about the light and other relevant conditions, that her color vision is normal, and that the fire hydrant she is seeing, like virtually all fire hydrants, is red. It is only the basic new knowledge that needs to be explained, that would be intact if all these beliefs were false, that is philosophically problematic. This knowledge, though philosophically problematic, is relatively trivial, but not completely so. There might not be a fire hydrant; she might not be seeing anything at all, while people play brain games with her in a completely dark room.

There are then four things involved in the truth of Mary's thought: her color experience, the type of color experience it is, the fire hydrant, and its color.

Consider Mary's phenomenology, as she studies the fire hydrant. I will use the term "experience" so that Mary's visual experience is a complex experience, and she can distinguish between the experiences that are parts of the complex. The color experience she has when she has an experience of the fire hydrant is different than the experience she has of the yellow dandelion next to it, and different than the experience she has of the fire engine. The experiences of the colors of the fire engine

could naturally use to express the knowledge, sometimes inserting a gratuitous "that" because it sounds better, as any indirect discourse formulation raises a number of questions about attitude reports that aren't essential to the points I make and on which almost no two philosophers agree. Although this way of formulating things might suggest it, I do not believe we need indexicals and demonstratives to think the thoughts we naturally express with them.

² Or, for that matter, if they were objective states, in some sense of that word, of any kind. See Perry (2001).

Mary and Max and Jack and Ned | 81

and the fire hydrant are of the same type, and different in type from that of the dandelion.

Mary knows that colors are properties that material objects have that are detectable by normally sighted people in favorable light, and that the color is seen at the surface of the object, and can be either uniform or varied. She knows that the visual experience of the color of an object involves the part of the visual field enclosed by the visually perceived boundaries of the object. She has had similar experiences before, but only of black and white and, I suppose, various shades of grey, the fillings for parts of her visual field delineated by the boundaries of objects while she was in the black and white room. She takes it that she is seeing a color, and having, for the first time, the ordinary experience one has when one does so. She knows that the object does not cease to be colored when she closes her eyes and her color experience ceases. So she knows that there are two types of things involved: the fire hydrant, with its color, and her color experience, which is of a certain type.

We can use “mode of presentation” both for the way that things are perceived and the way things are thought about. In the former sense, we would usually have in mind the particular way the object was presented to the subject. So Mary’s mode of presentation of the fire hydrant has to do with the type of impressions, in Hume’s sense, that it gives rise to, which will in turn depend on her position, the light, and so forth. In this sense, there are no modes of presentation of one’s own experiences. One doesn’t perceive them; they are not presented to one as the cause of experiences.

However, to *think* about her particular experience, Mary does need a mode of presentation of it. How does Mary *think of* her present experience as she has it? She thinks of it as playing a certain role in her life, as the present color experience she is having, occupying a certain part of her visual field, and due to the object that is determining what goes on in that part of her visual field. To think of something as playing a role in one’s life in this sense, one does not need to have the concepts to articulate the role; it suffices to be attuned to the facts. Naive Mary would exhibit attunement to these facts in a variety of ways: by closing her eyes if she doesn’t like it; by getting closer to the fire hydrant if she does like it; by focusing her attention on it if something about it interests her; by thinking the sort of thoughts I am getting at with the locution “this_i experience is so and so”, and so on. Sophisticated Mary

will not only be attuned to such facts, but be able to think and talk explicitly about them.³

To form an idea or concept of a kind or a type of thing, it usually suffices to have an exemplar and a similarity relation. Mary has both for the type of color experience. She has the experience to which she is attending. She contrasts that experience with other experiences that she has had in the past and is having now: the experience she remembers having when she saw objects in the black and white room, the experience she has when she diverts her attention to the dandelion next to the fire hydrant (not similar), the experience she has when she diverts her attention to the fire engine (similar), and so forth. She knows that by playing with her brain, or lights, or some combination, her captors could give her similar experiences when the material objects seen were not of the same color, or when there were not material objects seen at all. Even if she is super-cautious, she can think: “this_i type of color experience, whether it be Quale_{RED} or Quale_{GREEN} or whatever, is the type of experience I have, at least right now, when I see that color, whether it be red or green or some other color”. In so thinking, she would be employing what seems to me to be a good candidate for what Block calls a “phenomenal concept”.

How do the truth-conditions of Mary’s doxastic states change, and why, when she acquires this new knowledge? In particular, the most basic new acquisition, that she would express as:

The type of *this_i* color experience is what it is like for me to see the color of *that* fire hydrant.

What is required for Mary’s new belief to be true? It depends on what we take as given. Given only that Mary’s thought is appropriately expressed by the quoted English words, “*This_i* type of color experience is what it is like (for me in these conditions) to see the color of *that* fire hydrant,” what is required is that there be some fire hydrant, and some type of color experience, so that Mary is seeing and attending to the fire

³ Block describes my view in terms of “being attuned to concepts”. But this isn’t a phrase I use in this situation, nor do I quite understand what Block has in mind. If Mary refers to the fire hydrant as part of a speech act, this might require her to be attuned to facts about the concepts of other people. To use demonstratives effectively in speech one needs to be attuned to facts about what other people see and don’t see; this is the sort of situation in which I would talk about being attuned to concepts, i.e. being sensitive to who has what concepts without having the concepts to articulate that to which one is sensitive.

Mary and Max and Jack and Ned | 83

hydrant, and is having and attending to an experience of that type, and the color experience is caused in the appropriate way by the color of the object and the conditions she is in. So the conditions of truth involve all four things, the experience, the type, the fire hydrant, and its color.

Given, in addition, that the type of her color experience is Quale_{RED} and the color of the fire hydrant is red, and she is a normal person, her thought will be true iff

Quale_{RED} is the type of experience normal people have when they see red in normal conditions;

and this is something she already knew in the black and white room. So the truth conditions of her thought *given* those facts does not get at the new knowledge.

So, if we think Mary has new knowledge, but that she could have known in the black and white room which types of color experience go with which colors, we must find the new knowledge, the conditions that the truth of her new belief imposes on the world, in abstraction from those facts.

The natural answer to this is the one we all learned from Frege. Although the objects (the type of experience, the color) are old, the modes of presentation are new. She has thought of red many times, but never as the color of which she is having a normal experience. She has thought of Quale_{RED} many times, but never as the type of an experience she is having and to which she is attending.

Frege's idea that one object can have numerous properties that individuate it, each of which, or at least many of which, can serve as modes of presentation,⁴ which he introduces in the first long paragraph of *Über Sinn und Bedeutung*, needs to be kept distinct from his theory of *Sinne*, as it is developed in the rest of that essay. According to this theory, when one thinks of an object, via a mode of presentation, as meeting a certain condition, that mode of presentation is a constituent of the *Gedanke*, the proposition that corresponds to one's thought that the object meets the condition. The proposition that is the object of one's thought, in turn, is the *Sinne* of the complement sentences of true reports of the thought. Jumping from the plausibility of Frege's insight about modes of presentation, to the validity of his theory of *Gedanken* and *Sinne*, is a little like jumping from the distinction between up and down, to Newton's theory of absolute space.

⁴ Block says something here I'm not sure what to make of....

Indexicals and demonstratives make this pretty clear, for the need for distinguishing modes of presentation is vivid, but the rest of the theory of *Sinne* has to give somewhere. I see Ned Block and believe, on the basis of what I observe, "that man is suave and debonair". My perceptual mode of presentation of Block is something like: *the man I am looking at and attending to*. I am attuned to facts about occupants of this role; I use demonstratives reliably for such objects; I am sensitive to the difference between objects I am looking at and attending to and those I am not; I know how to pick up information about such objects.

But my perceptual mode of presentation does not seem to be part of *what I think*. I avert my gaze, in order not to be overly charmed; I continue to think the same thing, but via a new mode of presentation, my memory of the man I saw.⁵ You can truly report my belief by pointing at Ned Block and saying, "Perry believes he is suave and debonair." Your mode of presentation of Block is as the man *you* are seeing, attending to, and calling attention to. I believe the same thing, while I am looking at Block and remembering him; but my modes of presentation differ; your report is correct, although your mode of presentation is not mine.

One of the identities of Frege's theory needs to be gainsaid, that between modes of presentation and constituents of *Gedanken*, or that between *Gedanken* and what is thought, or that between what is thought and the *Gedanken* that correspond to the complement sentences of true reports of the thoughts. My approach is (roughly) to hold to the second and the third, and give up the first. Other responses to the data provided by indexical and demonstratives are certainly possible; Stalnaker and Lewis can be thought of, ignoring subtleties and differences, as holding the first and second and giving up the third.⁶

If we accept all three identities, then if Mary has new knowledge, there has to a new proposition *P*, so that we can truly report "Mary knows that *P*" and not just a new way of knowing a proposition already known. While not all advocates of the knowledge arguments are Fregeans, something like the Fregean identities are always lurking.

Give up the identities and it does not follow from the fact that Mary has a new belief, a new opportunity to be wrong or right about things, new conditions on the truth of her mental states, that there is a new fact

⁵ See Perry (1980, 1997) for an exploration of such issues.

⁶ See Stalnaker 1981 and Lewis (1979).

Mary and Max and Jack and Ned | 85

known, about a new property of experiences, that wasn't any of the properties she knew about in the black and white room. Her new belief puts the same conditions on colors and types of experiences as one of her old ones, but it puts new conditions on other things, the things that are parts of the modes of presentation and not of the subject matter. For the new belief to be true Mary has to be having an experience of Quale_{RED}, but this experience isn't part of the subject matter of her new belief or her old belief, but has to do with the modes of presentation of her new belief.

Perhaps Mary is so intelligent that she predicted that she would see a red fire hydrant and have Quale_{RED} when she steps out of the room. She can predict this as an existential generalization: "there is a unique fire hydrant and a unique time when I am to be released, and there will be an experience caused by the fire hydrant at that time, and the hydrant will be red at that time and the experience an instance of Quale_{RED}..." She could existentially instantiate and assign names: Call the fire hydrant *A* and the experience *E* and the time *T*. She then could formulate her prediction, using the tenseless 'be': "The type of *E* be what it's like to see the color *A* be at *T*; the type of *E* be Quale_{RED}; the color of *A* be red." That thought, uttered at any time, will have the same subject matter truth-conditions as her thought, at the time of the experience, "The type of *this_i* color experience is what it's like to see the color *that hydrant is now*; the type of *this_i* color experience is Quale_{RED}; the color of *that hydrant* is red. But the modes of presentation will be different—in my theory reflexive truth-conditions of these thoughts will be different. However confident Mary was of her prediction, the equations "*this_i* color experience = *E*" and "*that fire hydrant* = *A*" and "*T = now*" still contain new information, because the modes of presentation are different. Super-intelligent Mary provides a more complex knowledge argument, but no new issues of principle.

NED

Block thinks my account leaves out phenomenal concepts. Whether Mary acquires a phenomenal concept depends on what we mean by this phrase. At one point Blocks says "A phenomenal concept of the experience of red is what Mary lacked in the black and white room and what she gained when she went outside it." By having the experience of

86 | John Perry

seeing red, she gained a new way of thinking of Quale_{RED}, as the type of an experience she was having and attending to. This way of thinking of the qualia of red did not require that she realize that it was the qualia of red, the one she called “Quale_{RED}” even while in the black and white room. She could consider the possibilities that if there were a trickster, or her vision was not as normal as she assumed, that it was Quale_{GREEN} or Quale_{YELLOW} and so forth.

Blocks also says,

With all this emphasis on phenomenal concepts, you might wonder what they are supposed to be. A phenomenal concept is individuated with respect to fundamental uses that involve the *actual occurrence* of phenomenal properties. In these fundamental uses, an actually occurring experience is used to think about that very experience. No one could have a phenomenal concept if they could not in some way relate the concept to such fundamental uses in which the subject actually has an instance of the phenomenal quality.

Again, I see no problem, although the word “individuation” almost always sends shivers up my philosophical spine when it is used with respect to concepts. Mary is having an experience, and using that experience to think about the type to which it belongs. She uses the concept to bring various other experiences in her total visual experience under it, the ones she deems to be similar, and contrast them with others she is having and remembers from her days in the black and white room.

Or consider this:

Consider a specific phenomenal property, Q, e.g. the property of feeling like the pain I am having right now. (If pain just is a type of feel, then Q is just pain.) The physicalist says, let us suppose, that Q = cortico-thalamic oscillation . . . This is an *a posteriori* claim. Thus the identity depends on the expressions on either side of the ‘=’ expressing distinct concepts, that is, having distinct modes of presentation, for if the concepts and modes of presentation were the same, it is said, the identity would be *a priori*.

No problem again. As we have seen, Mary can have the phenomenal concept, without knowing which quale it is of and so, should qualia be physical states, without knowing which physical state it is of.

But then problems develop:

‘Q’ in my terminology is very different from ‘Q_R’ in Perry’s terminology since ‘Q_R’ is a term that Mary understands in the black and white room. ‘Q’ by contrast is meant (by me even if not by Perry and Smart) as the verbal expression of a *phenomenal* concept. A phenomenal concept of the experience

Mary and Max and Jack and Ned | 87

of red is what Mary lacked in the black and white room and what she gained when she went outside it.

There is nothing mysterious about “ Q_R ”; it’s just short for “the quale caused in normal people by seeing red objects in normal light”. There is no reason Mary shouldn’t understand that term in the black and white room.

But what is Block’s “ Q ”? What is this term that Mary couldn’t understand until she had the experience? Since it is the verbal expression of a phenomenal concept, it looks like it should just be “the quale I am having now, as I attend to the color of the fire hydrant” or “the property of being the sort of sensation I am having now, as I attend to the color of the fire hydrant”. What reason is there to suppose that Mary couldn’t understand such terms in the black and white room? Indeed, there may have been black or white or grey fire hydrants in her room, or visible from her room, or on the black and white videos she was allowed to see in the room. She would have known “the property of being the sort of sensation I am having now, as I attend to the color of the fire hydrant, is not Quale_{RED}, since I am confined to a black and white room”.

These words would have expressed a *different* concept than they do after she is allowed outside the room—just as “being the size of *that* man” could express different concepts at different times as one attended to different men. Maybe there is some superior way of expressing Mary’s new phenomenal concept, using terms that she couldn’t understand in the black and white room. Could be... but what would these terms be? I assume Block’s “ Q ” is supposed to be shorthand for the verbal expression of a phenomenal concept, but we are never given the longhand version of it, so how are we to be sure that Mary couldn’t understand this term in the black and white room?

Block continues:

Why do I insist that ‘ Q ’ express a phenomenal concept? Because the mind–body identity claim under consideration must be one in which the phenomenal property is referred to *under a phenomenal concept of it* for the Property Dualism Argument—in any of its forms—*even to get off the ground*. (The Knowledge Argument also depends on the use of a phenomenal concept in my sense.)... If the original paradigm of mind–body identity were “the property whose onset of instantiation here was at 5 p.m. = cortico-thalamic oscillation”, the property in virtue of which the left-hand term presents the referent would not be a special candidate for non-physicality. It would be the property of being instantiated here starting at 5 p.m. The Property Dualism Argument

depends on an identity in which a *phenomenal concept* is involved on the mental side. To allow a non-phenomenal concept is to discuss an argument that has only a superficial resemblance to the Property Dualism Argument.

I don't see much here to disagree with; I think Mary acquires a phenomenal concept; it is quite different from the sorts that Block mentions as not interesting; it is different because it is exactly the concept that Mary would have, the one she would naturally express with "This experience" or other words to that effect.

Block also says this (n. 10):

Chalmers (2003) argues that phenomenal concepts cannot be demonstrative concepts. The main argument could be put as follows: for any demonstrative concept, say ' $this_i$ ', $this_i$ has phenomenal property P would be news. But if the demonstrative concept was genuinely a phenomenal concept, there would be some claims of that form that are not news. I agree with the "not news" rule of thumb, though I would not go so far as to agree that it shows no demonstrative concept can be phenomenal. However, whether or not it shows that there can't be a concept that is both demonstrative and phenomenal, the demonstrative concepts that Perry is talking about are not phenomenal concepts in the sense required to motivate the Knowledge Argument and the Property Dualism Argument, the sense required to ground the metaphenomenal move.

This leaves me perplexed. As we saw, demonstrative phrases can be used to express different concepts in different situations. As Mary turns her attention from the fire hydrant to the lawn next to it, she might use the phrase " $this_i$ color experience" for different concepts; the two concepts might be part of the same thought: " $this_i$ color experience is much more soothing than [turning her head back to look at the fire hydrant] $this_i$ color experience". But how exactly is it news to Mary, as she looks at the fire hydrant, that " $this_i$ color experience has phenomenal property P", if phenomenal property P is exactly the property she has never experienced until now, and is now experiencing, and attending to, and referring to with " $this_i$ color experience"?

CONCLUSION

As I said, Block's paper is rich and interesting—and long. I don't claim to have digested all that he has to say, and have not tried here to discuss all of it. I hope to return to the issue of Black's argument and the variety of

Mary and Max and Jack and Ned | 89

modes of presentation at a later time, after having enlisted the help of seminar students in coming to grips with more of it.

Stanford University

REFERENCES

- Frege, Gottlob (1892) *Über Sinn und Bedeutung* (repr. Göttingen: Vandenhoeck & Ruprecht, 1994).
- Lewis, David (1979) "Attitudes 'De Dicto' and 'De Se'", *Philosophical Review* 88: 513–43.
- Perry, John (1980) "A Problem about Continued Belief", in Jerome Dokic (ed.), *Pacific Philosophical Quarterly* 61/4: 317–22.
- (1997) "Rip Van Winkle and Other Characters", in Jerome Dokic (ed.), *European Review of Philosophy*, ii. *Cognitive Dynamics* (Stanford, Calif.: CSLI), 13–39.
- (2001) *Knowledge, Possibility and Consciousness* (Cambridge, Mass.: MIT).
- Stalnaker, Robert (1981) "Indexical Belief", *Synthese* 49: 129–51.

