SO JUST WHAT IS IT ABOUT MARY? ON KNOWING "WHAT IT'S LIKE"

David Beisecker

I. There's Something about Mary

Frank Jackson's knowledge argument continues to be one of the most compelling and curious illustrations of the so-called "hard problem of consciousness."[1] Jackson's initial aim was to draw out an intuition that there is some sort of epistemic gap between phenomenal and non-phenomenal facts. To do so, he invites us to consider the celebrated case of Mary, a neuroscientist who is supposed to know everything there is to know about the mechanics of the human visual system, but for some fantastic reason (typically imprisonment in a wholly black-and-white environment, perhaps the Kansas depicted in *The Wizard of Oz*), she has never had a red sense impression Most are inclined to agree that despite her vast knowledge of neuromechanics, Mary nevertheless lacks "knowledge of what it's like" to see red**.** Were she to escape Kansas and spy a ripe tomato (which she somehow already knows to be red), she might well say something like, "Ah, so *this* is what it's like to see red." And so the thought experiment suggests that phenomenal knowledge cannot be reduced to, or derived from, theoretical knowledge of physical, physiological, or even functional and representational facts. Originally, Jackson went on to elevate this epistemic gap into a metaphysical one. That is, he took the thought experiment to support the thesis that phenomenal facts are ontologically distinct from the mundanely physical, physiological, or functional.

Materialists are predictably spooked by the idea that there is a phenomenal realm that lies above and beyond (and perhaps not grounded in) the physical. To be sure, this thought experiment is fanciful and woefully underdescribed, and in that respect, it resembles many other famous thought experiments in philosophy. Nevertheless, it seems hard to resist the intuition that Mary learns *something* when she escapes her black and white environment There is just something about Mary which cries out for explanation. The challenge for materialistically inclined philosophers of mind, then, is to explain (or explain away) her post-release enhanced epistemic standing without invoking mysterious, non-material "phenomenal" facts**.**

Perhaps the most popular response to this challenge is the "ability hypothesis" made famous by David Lewis.[2] Lewis suggested that the knowledge Mary gains is not factual knowledge at all (or knowledge *that*), but rather some sort of ability (or knowledge *how*). Originally, Lewis proposed that upon having her first red visual impressions, Mary gains new imaginative capacities—e.g., the ability to conjure up a red impression in memory. More sophisticated versions of this strategy hold that Mary gains recognitional capacities or something like the ability to access physical facts in a new "quasi-indexical" fashion**.** Without going into great detail, the trouble with these proposals is that it is hard to pinpoint exactly what the ability or abilities in question are, for it

seems that one can always pry them apart from the knowledge Mary gains upon her first red sense impression. For instance, one could reasonably suppose that Mary learns what it is like to see red, even if she could not later come to envisage it in imagination.

Now I find it significant that several presentations of hard problems of consciousness are presented in *epistemic* terms—for example, that Mary is said to lack *knowledge* of what it is like to see red, or that we cannot *know* what it is like to be a bat. I think we can make the most progress understanding Mary's curious epistemic situation by concentrating on the *function* of such knowledge attributions. Representational theories of phenomenal consciousness focus, as it were, upon the *belief component* of knowing what it's like, say, to see red; they seek to unpack claims about the phenomenal character of experience in terms of the representational contents of perceptual states. Thus we see their advocates making up types of so-called "non-conceptual intentional content to correspond to our intuitions pertaining to phenomenal distinctions.[3] However, we all know that knowledge attributions involve considerations beyond simply specifying the content of a subject's belief. Specifically, knowledge has a *justification component* too, and this, I argue, is the dimension where Mary (as well as bats and zombies) falls short. In this paper, I present a twist on the ability hypothesis, according to which Mary gains not just an ability, but rather a *justificatory status* with respect to her abilities to deploy observation concepts in experience.

II. Sellarsian Sense Impressions

Let us begin with a mundane observation. Competent speakers of natural languages must learn how to apply observation vocabulary in experience. We all need to be trained to make even the most basic observation reports. Our ability to classify things *as red*, or even *as looking red* is not innate. While we have biologically innate predilections for certain classification schemes, other speakers of our language must teach us how to make observation reports that accord with the specific classificatory dispositions of our fellows. Subjects thus face the task of coordinating or calibrating states of themselves with the application of observation concepts in experience. Simply put, speakers must learn to report the presence of a certain property, say, "red," whenever they are struck in certain fashions—that is, whenever they are in certain internal discriminatory states. And they must further learn when to restrain their acquired dispositions to report such a presence when circumstances are such that the subject's being in a particular state is *not* a reliable indicator of something's actually exhibiting that property. In those circumstances, subjects learn to say that something only "looks" or "seems" to exhibit that property. They learn to report merely that they are stimulated in a way that, under normal circumstances, *would* reliably indicate that property's presence.

Although these internal discriminatory states are presumably physiological states (and states of the nervous system in particular), usually subjects are not able to identify them in those terms. Even our most accomplished neuroscientists have trouble identifying the neural correlates of the simplest

sensations**.** The point here is that ordinary observational vocabulary—including "looks"-talk—is conceptually prior to a developed neuroscience**.** In speaking about the task facing speakers as they learn to apply observation concepts in experience, we should remain theoretically neutral regarding the underlying physiological substrate.

Following Sellars, we can introduce the philosophical notion of a *sense impression* as a theoretically neutral way to refer to such discriminatory states of subjects**.**.[4] Here we understand a sense impression to be the imprint that is characteristically left upon us by the presence of some property under normal circumstances, which can then be used by that subject to elicit observation judgments of that property's presence**.** And so the task described above is that of subjects learning how to coordinate sense impressions with the application of appropriate observation concepts.[5]

Elsewhere, I have argued that many of the curious things philosophers have said about phenomenal consciousness can be understood as making perfectly straightforward assertions about Sellarsian sense impressions**.**[6] In particular, the expression "what it's like" picks out the particular manner in which a subject realizes its sense impressions**.** Such a proposal makes evident and intuitive sense of inter- and intra-personal comparisons of the qualitative dimension of experience**.** Due to differences in our perceptual constitutions and available discriminatory states, "what it's like" for us to see red might be completely different from "what it's like" for a bat or bug-eyed alien to see red, so much so that we could not understand what it would be like to be either**.** And while what it is like for me to see red is presumably pretty much what it is like for most everyone else, it might differ slightly for non-standard folk (e.g., synaesthetes, or those outfitted with those spectrum-inverting lenses of philosophical legend). Furthermore, since my perceptual apparatus may change over time, even though I might not realize it, what it is like for me to see red now might well *not* be what it is like for me to see red in the future or the past**.** Indeed, it is possible (albeit highly improbable) for what it is like for me to see red eventually to shift all the way across the spectrum and become what it is like for me to see green**.** One can see that the functional unspecificity of sense impressions allows one to suspect that what it is like to have a particular kind of experience could have been other that what it actually is.[7] Consequently, another especially appealing aspect of this proposal is that it allows certain internalist intuitions or prejudices to take hold, while at the same time enabling us to remain externalists about intentional or representational content**.** On this account, internally indistinguishable subjects (those "molecule-by-molecule" duplicates of philosophical fantasy) will have experiences with similar phenomenal characters, even though external considerations dictate that the representational contents of their experiences are radically different.

III. Back to the Knowledge Argument

For the purposes of this paper, however, the most significant thing is that this story about sense impressions can help to explain what is going on in the

case of Mary and the knowledge argument more generally. Sellars himself famously thought that sense impressions play little role justifying empirical knowledge. He also had little truck with the skeptical worries generated by absent and inverted qualia scenarios. Nevertheless, it is clear Mary has not come to face the task that the language of sense-impressions has been introduced to describe—namely that of coordinating her own internal states with the application of particular observation concepts in experience The knowledge argument gains its force because it is unclear how Mary's assumed vast knowledge of physiological facts ever could help to overcome this task. Doing brain science by itself will not tell Mary when to report the internal occurrence of a red sense-impression.

Accordingly, I propose that attributions of "knowledge of what it's like" to have a certain kind of experience should be understood as claims that a subject has a justified ability to apply corresponding observation concepts in experience. While most speakers would qualify as capable enforcers of the norms governing color reports,[8] it would be irresponsible to extend this authority to just anyone. A color-blind person would be an incompetent teacher of color terms, even if he knows a great deal about the human visual system, as well as all the inferential connections between colors and other empirical concepts. It is also reasonable to withhold this authority from those, such as young children, who haven't been sufficiently indoctrinated into our reporting practices. Unlike the color blind, Mary (we suppose) has the *potential* to make accurate color discriminations. She also knows all the inferential connections color terms bear to the other terms in our language. However, until she *has actually demonstrated* that she can apply color concepts in experience as reliably as competent speakers, we can reasonably deny that she truly *knows* what it is like to have perceptual experiences of color. She clearly lacks a justificatory status, which manifests itself in our reasonable reluctance to grant her authority enforcing the norms governing our observational vocabulary. And this would be so, even if she happens to possess an uncanny *ability* to make accurate color discriminations. As a result, "knowing what it's like" requires more than just having an ability to apply concepts appropriately; one must be in a position to *justify* this discriminative capacity as well.

Notice crucially that I have not claimed that Mary is unable to entertain any specific *beliefs.* Even when she is trapped inside "Kansas," she might *suspect* that something looks red to her yet fail to *know* this, for she fails to have the appropriate "experience" to justify this suspicion. Thus we can hang onto Jackson's conclusion that Kansas-bound Mary lacks propositional knowledge that we would normally express with observation reports. Rather than missing the ability to form certain beliefs, she lacks the history or "experience" required for her to entertain those beliefs *responsibly.* And it is her assumed epistemic responsibility, not simply her lack of experience, which really prevents her from ever entertaining beliefs that things look red to her.

Consider how this proposal applies to those ever-popular subjects of philosophical fantasy: our physical and functional duplicates spontaneously generated out of swampmuck. Such abominations might make all sorts of claims

about how things look to them, and they might try to convince us that they have the requisite experience and know-how to enforce the norms governing our observation vocabulary. But the justifications they give for entitlement to this authority would fail, simply on the grounds that they would be *false*. So even though by some remarkable coincidence, some such beings happen to have the discriminatory capacity and classificatory dispositions to be competent reporters, it still would be inappropriate for us to so treat them. For if it truly were a cosmic coincidence that they have this gift, then we would be in no position to believe this to be the case.

Insofar as their perceptual apparatus differs from our own, we would also be justifiably reluctant to grant perceptually exotic creatures—bats, bots, or bug-eyed aliens, for instance—knowledge of what it is like to use our color terms. Lacking the perceptual capacities to deploy our observation concepts in experience, they might not ever attain the status of full-fledged (norm-enforcing) members of our linguistic community. In particular, their different physiology might well prevent such beasties from being able to tell when things are likely only to "look red" to a human observer. That is, their different perceptual equipment might prevent them from anticipating our justifiable perceptual errors. Likewise, we would be unable to master a bat's observation concepts. Hence we can respect the intuition that we are unable to know what it is like to be a bat, without having to claim we can not so much as entertain the same beliefs. Some perceptually exotic creatures might even make the same color discriminations that we do (in their own terms, of course). The conditions of proper application for some of their observation concepts would *mirror* those of our own color concepts. Still, if we lack sufficient contact with these creatures to believe this extensional equivalency with sufficient justification, we can reasonably deny them the authority to enforce the rules governing the use of *our* color concepts. So while they would know what it is like *for them* to see red, they might not know what it is like *for us* to see red.

This last bit shows how we can deal with an objection which might be raised to approaches like the one endorsed here, which tie knowing what it is like to have certain experiences with linguistic mastery or grasp of particular perceptual concepts. The objection has us suppose that Mary, while still in Kansas, has gained access to several unlabeled paint chips, at least one of which happens to be red. With that chip in view, Mary comes to have her first red visual sense impression, and so one might be tempted to say that she comes to know what it is like to see red. Yet she fails to realize that it is a *red* sense impression that she has come to experience So while she comes to know what it is like to see red, she nevertheless does not yet have a justified ability to apply the concept red in experience. It might seem, then, that our proposed analysis fails.

But this objection ignores the perspectival nature of knowledge attributions generally. While I grant that Mary has failed to demonstrate mastery over our concept of red, presumably she can still classify future visual experiences as being of roughly the same type as she has when she views the red paint chip. Thus we might still claim that she has acquired a justified ability to apply an

observation concept which she could demonstratively identify as "the shade of *that* chip," and which is more or less extensionally equivalent to our concept of red. And so, in a *de re* sense (or from our perspective), we might say *of the property red* that Mary has learned what it is like to see *it.* But in a *de dicto* sense (or from the perspective of *her* concepts), we can reasonably deny that she knows what it is like to see red. Not until Mary comes to realize that her experience is one that *we* would classify as a seeing of red, would she characterize herself as knowing what it is like to see red. Once we register that attributions of knowledge of what it is like admit to the same *de dicto/de re* distinctions as attributions of knowledge more generally, we can see that the objection fails to provide a true counterexample to the proposed analysis. Indeed, I take this consistency with other types of knowledge attributions to be a great virtue of the present account.

IV. Conclusion

As mentioned before, this paper is part of a much larger project, which seeks to understand phenomenal consciousness by making sense of the puzzling things philosophers are inclined to *say* about "what it's like" to have certain experiences.[9] Rather than focusing upon the nature of that which is attributed, the strategy is to explain what we are *doing* when *we attribute* to subjects knowledge of what it is like to have certain observational experiences. Now some might object that such an approach is too deflationary to explain phenomenal consciousness. While it might explicate our *judgments* about phenomenal consciousness, it fails to address the puzzling features of consciousness itself.[10] However, notice that on the account just given, attributions of knowledge of what it is like are "objective" or attribution-transcendent in that it can be *proper* to attribute to one's knowledge of what it is like to see red without anyone actually doing so, and that everyone's attributing such knowledge to another (a "swampzombie," perhaps) doesn't make it the case that it would be *proper* to do so. So while knowledge of what it is like makes sense only in the context of attributing such knowledge to others, it does not follow that facts involving phenomenal consciousness are "merely attributed."

Although one cannot discern whether subjects "know what it's like" to have an experience just by examining the causal transactions inside their heads as they have those experiences, that merely shows that such a narrow focus abstracts away from the *epistemically* significant, historical facts required for them to have such knowledge. The proper moral is not that phenomenal consciousness must remain objectively ineffable, for these further social and historical features are by no means inaccessible from a third-person perspective. Moreover, we can finally see why subjects would find it important to possess knowledge of what it is like. For justifiably applying observation concepts in experience entitles subjects to do things forbidden to the inexperienced. Although "swampzombies" might be inclined to behave as I do, others would be

disposed to treat them differently. Hence having knowledge of what it's like really can "matter" or "make a difference" to conscious subjects.[11]

NOTES

[1] Frank Jackson, "Epiphenomenal Qualia," *Philosophical Quarterly* (1982) 127-136.

[2] See "What Experience Teaches," *Mind and Cognition: A Reader*, ed. William Lycan, (Oxford: Blackwell, 1990).

[3] See, for instance, Michael Tye's PANIC account of qualia in *Consciousness, Color, and Content* (Boston: MIT UP, 2000).

[4] See the final episode of his "Myth of Jones," Part XVI of "Empiricism and the Philosophy of Mind," recently reprinted in Devries and Triplett, *Knowledge, Mind, and the Given* (Indianapolis, IN: Hackett, 2001).

[5] To borrow a term from Dretske, one might say that subjects must face the task of "recruiting" appropriate internal states of theirs to play the role of indicators of particular observable features of the world.

[6] See my "Phenomenal Consciousness, Sense Impressions, and the Logic of 'What It's Like,'" *Consciousness and Emotion: Agency, Conscious Choice, and Selective Perception*, eds. Ralph D. Ellis and Natika Newton. (Amsterdam: John Benjamins, 2005) 137-153.

[7] These, of course, are the very intuitions which allow so-called "modal arguments" for the explanatory gap to get off the ground.

[8] Of course, perceptual concepts can be much more exotic, and less accessible to everyone, than colors. Think of those describing the subtle differences in the taste of wine, beer, or coffee.

[9] To be sure, I cannot pretend to have offered a complete account of consciousness, for I have not addressed the issue of what would make a mental state conscious, as opposed to unconscious.

[10] See in particular, David Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford UP, 1996) 186-9.

[11] I am aware that some critics will contend that such extrinsic, normative differences are, from a narrowly scientific perspective, explanatorily otiose. I think that such criticisms rest upon implausibly scientistic (perhaps physicalistic) assumptions that systematically exclude the rational types of explanations, in which attributions of phenomenal consciousness (and other intentionally-freighted concepts) most naturally find their home. See my "Functionalism and Folk Psychology: How the Mental Earns its Keep" *Southwest Philosophical Studies* 26 (2004).