

What Mary Might Have Known

An Exposition of the Knowledge Argument
Against Physicalism

Michael Lara

Advisor: Dr. Wayne Wright

California State University, Long Beach

Spring 2019

Introduction

What is it like to have a particular phenomenal experience? It's like *this*. There is an ineffability when attempting to answer this question. There must be a set of facts about the matter, and anyone who has had some particular experience surely knows them. Or do they? The challenge of expressing the facts about subjective phenomenal experience in comparison to facts such as the boiling point of water or the hardness of a diamond produces the intuition that the facts about phenomenal consciousness might be irreducible to the physical facts of the world. Thus, the *explanatory gap* arises. I hope to persuade the reader that the facts *about* phenomenal experience and the facts *about* boiling points and hardness are probably more similar than they seem. Even if we cannot close the *explanatory gap*, clarifying it by appealing to certain Physicalist arguments can be extraordinarily helpful in pushing back against the intuitions generated by the Knowledge Argument.

This work aims to show that the burden is on the claimant that Physicalism is false, and not on the Physicalist to show that Physicalism is true. In order to place the burden on the non-Physicalist, I will proceed by outlining Frank Jackson's views, before he becomes a Physicalist, as regards non-physical 'qualia,' and show why these beliefs are formed purely by intuition. Then, I review Michael Tye's *perspectival Physicalist* account of phenomenology to show how the conclusion of the Knowledge Argument against Physicalism can be rendered potentially unsound by demystifying the *explanatory gap*. Once I have shown how relocating phenomenology from in the mind to the external world helps us to surmount certain claims of the non-Physicalist, I assess Chalmers's claims of the necessity of non-Physical 'qualia' to explaining consciousness. After a short look at Daniel Dennett's response to Jackson's thought experiment about the physically omniscient yet experientially restricted color scientist Mary, I outline a Physicalist Jackson's Representationalist approach wherein he claims that the conclusion of his own Knowledge Argument is false.

At the outset of this project, it was my goal to further my own understanding of 'qualia,' specifically as the term is in play in the Knowledge Argument. I had begun firmly in the non-Physicalist camp based on my own intuitions regarding phenomenal consciousness. Yet, at every turn, for every non-Physicalist account of the possibility of non-physical facts about 'qualia' into which I inquired, I was met with a Physicalist account that provided clarification, a higher level of explanatory power, and had a lesser reliance on intuition. Where I once held a belief that introspection could provide insight regarding experience itself, it seems that I have come to sympathize with Michael Tye in that, "...when I turn my attention inward to the experience itself, I always seem to end up scrutinizing external features..." of the world, and not the 'features of experience itself' (Tye, *Ten Problems of Consciousness*: 167). This inability to introspect on the features of experience itself is often referred to as *phenomenal transparency*. *Phenomenal transparency* marks the notion that even if there are some explicit features of experience that differ from the objects which they represent, they are diaphanous. We cannot turn our attention on the experience, but always refer back to the external referent experience. Gilbert Harman provides an example of *phenomenal transparency*: "Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict you will find that the only features there to turn your attention to will be features of the presented tree, including relational features of the tree from 'here'" (Harman, 1990: 39).

The conclusion of the Knowledge Argument comes with an ontological commitment to non-physical entities. If, in order to explain some aspects of the world, non-physical facts were required, what kind of access might we possibly have to them? How would we *know*, beyond intuition, that non-physical facts were necessary to understand some phenomena? Any entity of the sort would force Physicalism's falsity. David Chalmers gives us a useful definition of 'Materialism,' which, in this case, is synonymous with Physicalism. In order to gain some traction, I will adopt Michael Tye's definition of 'physical': 'anything to which the laws of a completed set of physics—which includes microphysics—apply.' In

accepting Chalmers' definition of Materialism and inserting Tye's definition of 'physical,' we see the thesis of 'physicalism' take form as follows:

“...everything in the world is [something to which the laws of a completed set of physics—which includes microphysics—apply], or that there is nothing over and above the [things to which the laws of a completed set of physics—which includes microphysics—apply], or that the...facts [about that to which the laws of a completed set of physics—which includes microphysics—apply] in a certain sense exhaust all the facts about the world. ...[M]aterialism is true if all the positive facts about the world are globally logically supervenient on the...facts [of a completed set of physics]...” (Chalmers, 41; Tye, 39).

For Chalmers, Physicalism is the notion that once the physical facts are in place in the world, *all the facts* are in place in the world. Now that a definition of Physicalism is adopted, I will focus on the central question posed by the Knowledge Argument: What is the nature of the knowledge acquired when someone actually has a phenomenal experience for themselves? Each author's presentation and response to Jackson's thought experiment about Mary provides insight into the character of their theories of consciousness as regards the possibility—or in the case of Chalmers, the necessity—of facts which are physically irreducible. It has elicited responses from the greatest thinkers in its genre, and it is impressive that such a short work has garnered such widespread attention. It seems that in some way or another, every philosopher who has grappled with the mysteries of phenomenal consciousness has taken it upon themselves to address what it is that Mary's hypothetical predicament does, or fails to do, in moving the conversation of consciousness forwards.

Ch 1 – The Knowledge Argument and Epiphenomenalism

Jackson's Knowledge Argument against PHYSICALISM claims that being in possession of the totality of the physical facts about experience, one could not deduce *all* the facts of experience. They will fail to deduce the knowledge that is acquired in actually having a phenomenal experience oneself. Watching someone have a color experience, even at the most microneural level, is not the same as seeing red oneself. Jackson's Knowledge Argument can be presented formally:

- (P1) Mary knows all of the physical facts about color experience.
(P2) Upon leaving the room, Mary learns some new fact(s) about color experience.
 (C3) Physicalism is false.

Jackson begins his essay *What Mary Didn't Know (1986)* by painting a picture of a color-deprived scientist, Mary, who lives her whole life in a completely black-and-white existence:

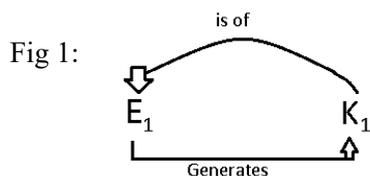
“Mary is confined to a black-and-white room, is educated through black-and-white books and through lectures relayed on black-and-white television. In this way she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of ‘physical’ which includes everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles” (Jackson, 291).

There are some obvious and immediate objections to this hypothetical that would ‘smuggle color into the room.’ Jackson notes a few, such as the color she would experience should she rub her eyes. His response to this is to suppose she is completely colorblind. Any amendment can be made to these kinds of objections in order to allow for the prescribed conditions of the thought experiment to hold. An important caveat regarding restrictions imposed on Mary while she is alienated from color experience is that those restrictions be lifted upon her emergence from the black-and-white room into the world where she can have her first phenomenal color experience.

Jackson argued that upon leaving her room and having her restrictions lifted, Mary will absolutely learn some genuinely new fact about phenomenal color experience, and thus the world, the first time she sees a red rose for herself. Namely, “...she will learn what it is like to see something red...” (Jackson: 291). *What it is like*, according to Jackson in 1986, is not a physical fact about Mary or her environment; the object of phenomenal knowledge, or ‘qualia,’ is non-physical. If she knows all of the physical facts about the world—this includes all of the physical facts about color, the processes undergone when an eye is hit by the light reflected from a surface, every functional property of the brain and the processes undergone when it receives the electrical signals from the optic nerve, etc.—yet she still learns some genuinely new fact, then the physical facts fail to describe the world simpliciter. “Hence, Physicalism is false. This is the knowledge argument against Physicalism in one of its manifestations” (Jackson: 291).

Jackson's criticism of Physicalism is based on the assumption of what has come to be known as the *learning claim*, or the claim that Mary absolutely learns *something* when she leaves the room. She knows *all* the physical facts; if she learns one genuinely new fact, Physicalism fails. Despite the disagreements regarding what it is that Mary does or does not learn when leaving her room, everyone agrees as follows (See Fig 1): There is some experience ‘E₁’ which Mary has never had before leaving the room, and, upon leaving the room and having experience ‘E₁,’ there is some knowledge, ‘K₁,’ which she

learns. The question is: are the facts that make up ‘K₁’ as pertaining to ‘E₁’ physical or non-physical, and what does the Knowledge Argument tell us about them?



Jackson explicitly states that, “...the knowledge Mary lacked [before seeing her first red rose] which is of particular point for the knowledge argument against Physicalism is *knowledge about the experiences of others*, not about her own” (Jackson: 292). She learns *what it has been like for others* to see red all along. Although she had complete physical knowledge about the world, she lacked any facts that are learnable exclusively by means of a necessarily subjective phenomenal color experience. Jackson suggests that, “...what she did not know until her release is not a physical fact about [others’] experiences. But [what she learns] *is* a fact about [their experiences]. That is the trouble for Physicalism” (Jackson, 293).

I argue that Jackson’s claims about the knowledge that Mary learns upon her release are based solely on his own intuitions because of the premises he set out for us. He is claiming that if Mary knows all the physical facts about the experience of others, but not some particular fact because it is only learnable or acquirable by means of having said subjective experience for oneself, then that fact must be a non-physical fact. Thus, ‘qualia’ must be non-physical in nature. Why, though, ought we accept this conclusion? It is a striking claim to argue that a fact is non-physical simply because of the perspective through which it must be obtained, and undermines its own stipulation: Mary knows *all* the facts about the physical nature of the world, including those which are necessarily subjective, if they are physical. So, the Knowledge Argument grants Mary more knowledge than anyone can explicitly account for. The epistemic status which she is granted by Jackson does not countenance necessarily private experiential facts as remaining private, even if they are necessarily private under normal phenomenal circumstances. Thus, if the facts regarding *what it is like* are necessarily private physical facts, then Mary knows them, regardless of her inability to access them through means of having the phenomenal experience herself. This is not supposed to be a knock-down refutation of the Knowledge Argument as a whole, but one inadequacy that must be attended to by Physicalist and non-physicalist alike.

As we will see, although some Physicalists accept the possibility of necessarily private facts about phenomenology, the overarching claim that it is just a brute fact of nature that some knowledge is necessarily private runs against the reductive mode of thought generally hoisted onto Physicalism by both supporters and critics. For this reason, the concept of necessarily private, physically irreducible knowledge is difficult to account for as a meaningful addition to the Physicalist thesis. This is another reason that the intuitions produced by the Knowledge Argument are so strong. In order for the claim that some knowledge is necessarily private to be held by a Physicalist, the claimant must present a reason *why* it is a brute fact of nature that this is the case, or some mechanism which forces the irreducibility of these facts.

This possibility places a hefty burden on Jackson—or any non-Physicalist who relies on the intuitions of the Knowledge Argument—to show why ‘what she did not know until her release is *not a physical fact* about [others’] experiences.’ To meet the demands of this burden, they must show how and why the facts regarding *what it is like* to have an experience are inherently non-physical. Though it is the burden of the Physicalist to explain why the *explanatory gap* arises in the case of phenomenology and nowhere else, the intuitions of the *explanatory gap* alone cannot be the driving force behind the non-Physicalist’s view.

The *explanatory gap* regarding the facts of phenomenal consciousness seems to stand somewhat alone in our picture of the world; why is this phenomena different in kind from the boiling point of water or the hardness of a diamond? This *explanatory gap* suggests that while knowledge of phenomena such as orbits can be fully reducible to and explained by means of physical facts, ‘redness,’ or *what it is like* to see red, cannot be reduced or explained in the same way. If we accept that physical facts are ones which are knowable by, reducible to, or deducible by means of the content accounted for by Tye’s definition—Anything to which the laws of a completed set of physics...apply—the *explanatory gap* eases the acceptance of the non-Physicalist intuitions produced by the Knowledge Argument.

It is the explanatory power of reducible physical facts which the non-Physicalist is targeting in relying on the intuitions of the *explanatory gap*. If the reductive Physicalist cannot conclusively close the gap, then they have a problem. In order to untangle this problem, we must separate the kind of knowledge about phenomenal experience which *is* reducible akin to knowledge about orbits, and the kinds which may not be. Once these kinds are classified, we can demystify the *explanatory gap* by telling a reasonable story about how certain facts of phenomenology are analogous to facts about orbits, and how certain facts are not.

If Physicalism holds that the physical facts entail the totality of metaphysically possible fundamental facts about our world, then it must fail on any account in which a genuinely new fact is learned by Mary. There is a sense in which Mary could think she learned some new fact, but actually learn an old fact in a new way. The Physicalist must allow for the possibility that Mary, having all of the physical facts, could fail to deduce the indexical nature of ‘*what it’s like to see red.*’ Jackson, in arguing against Physicalism, explains that it would be false to assume that knowing $(a = F)$ and that $(a = b)$ means that the knowledge of $(b = F)$ automatically follows. If $(a = 2)$, then $(F = 2)$. If $(a = b)$, then $(b = 2)$.

Under what circumstances might one fail to realize that $(b = F)$, or that they both equal ‘2’? One could simply overlook the fact that there was a reducible correlation, or be unable to deduce the fact that $(b = F)$ based on the kind of information that is represented by ‘b’ and ‘F.’ For example, as Terence Horgan suggests, ‘b’ could represent the fact that Superman can fly, where ‘F’ can represent the fact that Clark Kent can fly. Though they share the property ‘being able to fly,’ and it is *true* that Superman *is* Clark Kent, there is not enough information to deduce that Superman is Clark Kent with 100% accuracy. If the fact that Superman *is* Clark Kent is a stand-alone fact, then the non-Physicalist is arguing that a conjunction of two true propositions which share an identity accounts for a genuinely new fact. Still, if this conjunction was entailed by completely *physical* identities, Mary could fail to know that they were numerically the same, as well as qualitatively, because she does not know the totality of *negative* facts; she does not know that Clark Kent and Superman are the *only* people that can fly.

For an earlier, ‘qualia-freak’ Jackson, EPIPHENOMENALISM is the most acceptable alternative to Physicalism in that it (a) still accounts for non-Physical ‘qualia’ as required by the Knowledge Argument and (b) accounts for the physical world being causally closed in a meaningful way without applying some form of substance dualism; Jackson holds to a kind of property dualism in this case, which is a much stronger version of the dualist argument. A mental event which may hold causal power, for instance, would be a belief. If one was to believe that they could jump a 5’ gap over a canyon, they might be more willing to attempt the jump than if they did not believe they could make it. Experience, or *what it is like* to have an experience, would not maintain this same kind of causal power in every instance. If certain instances of these mental states lack causal power, then the Epiphenomenalist argument is forced to accept certain conclusions about decisions made regarding behavior. If a non-Physicalist is prepared to accept that *what it is like*, as an instantiated property in the world, cannot be a player in the formation of beliefs, they must also be prepared to argue *why*. If intuition is the anchor of their dispositions towards ‘qualia,’ then what about those people who intuit that *what it is like* must have some causal power in

belief formation? If ‘qualia’ can actively contribute to the formation of beliefs, and beliefs are the reason for action, then ‘qualia’ are, or can be, a part of the causal chain.

Being aware of the possible consequences of raw feelings lacking causal power is important: do we, as conscious beings, avoid pain-inducing sensations *because* of the possibility of pain, or is pain a byproduct of a situation which we avoid for other causal, psychological, or physiological reasons? Is it possible that our phenomenal experience lacks all causal power in our decision-making abilities? Even as a ‘qualia-freak,’ Jackson argued that qualia lacks all causal power in the physical world, and was nothing more than a by-product of the functionality of our physical systems based on how they happened to evolve into the complex systems they are today.

In *Epiphenomenal Qualia* (1982), Jackson presents three major arguments against Epiphenomenalism, and shows why they are mistaken in their claims about the causal power of qualia:

(i) “It is supposed to be obvious that the hurtfulness of pain is partly responsible for the subject seeking to avoid pain...” (Jackson, *Epiphenomenal Qualia*: 133).

(ii) “According to natural selection the traits that evolve over time are those conducive to physical survival. We may assume that qualia evolved over time—we have them, the earliest forms of life do not—and so we should expect qualia to be conducive to survival” (Jackson, *Epiphenomenal Qualia*: 133-134).

(iii) “We know about other minds by knowing about other behavior...,” and thus, if they behave like me, they must feel like me (Jackson, *Epiphenomenal Qualia*: 134).

In (i), Jackson refutes the idea that qualia, or *what it is like*, is necessarily responsible for actions such as declarations of pain, avoidance of certain situations, and so on. It is only necessarily so that both qualia and these declarations and avoidances arise from ‘happenings in the brain,’ and that we ought not infer that the qualia themselves are responsible for this behavior. If there is a lack of causal power in qualia, though, then what benefit can they be for us as conscious beings? Why would qualia be traits that are passed down through our evolutionary history?

As for (ii), Jackson argues that qualia are evolutionarily similar to thick coats on a polar bear. On one hand, they are necessary to keep the bear warm in Arctic temperatures; yet, on the other hand, the thicker the coat, the heavier it will be. Thus, a by-product of a thick coat for warmth, which is necessary for survival, is a heavy coat, which slows the bear down. This slowing-down is non-conductive to the survival of the bear, should any faster predator find it in its sight. So, qualia are the non-conductive by products of physical happenings in the brain, and not contributions to the overall wellbeing of the system (or person) itself.

In (iii), if there is no direct causal link between qualia and behavior, as suggested in Jackson’s refutation of (i), then qualia are not a means by which we can know that others have minds as we do.

If there is a complete lack of causal power, though, then why posit qualia as existing in the same way as physical entities are posited? As Jackson notes early on in the paper, it is the intuition that something special is going on which provides a reason to accept qualia as existent. Counterfactually, a lack of causal evidence—which is otherwise taken as evidence in general— suggests this lack of causal power. As I have been arguing all along, it is these kinds of intuitions which should be disregarded as ultimately truth-bearing.

A final concern regarding Epiphenomenalism is this: If the conclusion of the Knowledge Argument is true, and Epiphenomenalism is true, then even after Mary leaves the room and has her first color experience, she would fail to grasp the *facts* of it. There is an assumed relationship between knower and knowledge: to know 'x' requires some causal connection to 'x.' Therefore she, nor anyone else, could *know* what some inherently non-causal, non-physical entity *was like*. So, either (P2) of the Knowledge Argument is false and she does not learn any new facts about the experience of others as they are unknowable, the conclusion of the Knowledge Argument is unsound, or Epiphenomenalism is false.

Per this inconsistency, the non-physical facts of 'qualia' are *unknowable* based on their incapability of participation in a causal relationship between a knower and what is known, are argued to exist based on intuition, and have no observable or inferable causal efficacy. Yet, they require complex rules which govern them that are unnecessary to understand the workings of the world *except themselves*. Thus, Occam's Razor suggests that non-physical qualia a la Epiphenomenalist Jackson ought to be eliminated from our ontologies until further evidence for them arises. In moving away from intuition, this is the kind of reasoning that led Jackson—and, subsequently, me—to give up his self-ascribed title of 'qualia freak,' and pick up his 'Physicalist' identification card later in his career.

Chapter 2 – Relocating Phenomenology

In *Ten Problems of Consciousness (1995)*, Michael Tye refers to the seeming incompleteness of Physicalism as “...the problem of perspectival subjectivity” (Tye, 14). For Tye, who is himself presenting a Physicalist account of consciousness, phenomenal subjectivity is an issue that any Physicalist will need to address to have their argument hold. Perspectival subjectivity is that which accounts for the necessity of a first-person experiential perspective in order to acquire knowledge of a subjective experience. There is no set of physical facts that can construe *what it is like* to see a color, without that person adopting a certain experiential perspective. As I have shown, if those facts are indeed physical, then Mary will have them without needing to adopt that particular perspective; according to the Knowledge Argument, she knows *all* of the physical facts. Tye’s presentation of Mary is as follows:

“Consider a brilliant scientist of the future, Mary, who has lived in a black-and-white room since birth and who acquires information about the world via banks of computers and black-and-white television screens depicting the outside world. Suppose that Mary has at her disposal in the room all the objective, physical information there is about what goes on when humans see roses, trees, sunsets, rainbows, and other phenomena. She knows everything there is to know about the surfaces of the objects, the ways in which they reflect light, the changes on the retina and in the optic nerve, the firing patterns in the visual cortex, and so on. Still there is something she does not know. She does not know what it is like to see red or green or the other colors. This is shown by the fact that when she finally steps outside her room and looks at a rose, say, she will certainly learn something. Only then will she appreciate what it is like to see red. So, Physicalism is incomplete” (Tye, 14).

In order to show that the Knowledge Argument does not refute Physicalism, Tye shows that Mary can possibly learn something *about the experience of others* while being in possession of all of the physical facts. He demystifies the *explanatory gap* by first relocating the object of phenomenal experience to the outside world, and second, introducing *perspectival physicalism* as a mechanism whereby Mary can know *all* the physical facts about the nature of the world, yet still fail to know certain facts *perspectivally*. What she can fail to know is *what it’s like* to be in a physical phenomenal relationship with some content in the external world.

For Tye, phenomenal states are neither analogous to neurophysiological brain states nor psychofunctional ones. His analysis of phenomenal states and their properties leads him to outlining his PANIC Theory in order to show the explanatory power of his view; PANIC stands for Poised Abstract Nonconceptual Intentional Content. The PANIC theory states that, “...representations that differ in their PANICs differ in their phenomenal character, and representations that are alike with respect to their PANICs are alike in their phenomenal character” (Tye, 138).

The *Poised* refers to the sensory inputs that are at-the-ready to make an impact on the beliefs of the subject. Beliefs, then, once formed, are no longer appropriately poised to be considered phenomenal. By describing *phenomenal content* as being properly poised to be experienced, Tye is relocating the phenomenal from the internal world, or the mind, into the external world. For Mary, there is no content available to her which is properly poised to provide her with the phenomenal concept ‘red,’ and thus, she cannot be in possession of that concept. Yet, once she comes into a phenomenal relationship with some content which is poised to teach her the phenomenal concept ‘red,’ she will learn it against her will.

Abstract refers to the fact that the relevant external content need not be involved in providing the experience. For example, an amputee can feel an itch in a phantom limb. So, PANIC can account for non-existent or non-instantiated external content as having the same phenomenal affect, as long as it is

properly intentional and represented in the appropriate way. Abstract external content, however, requires an appropriately similar *phenomenal concept* of which to refer. For someone who has lost an arm, a phantom limb would be a representation of the phenomenal concept of ‘arm’ which they had previously learned.

Nonconceptual refers to the fact that the subject doing the experiencing does not need to possess concepts of the phenomena to have a phenomenal experience of them. Mary, for instance, need not have a concept of redness to see red for the first time; phenomenal content is imposed on the subject against their will. Once phenomenal content has been imposed on a subject through the senses, however, the subject learns a *phenomenal concept* of that content. This aspect of the phenomenal stands in opposition to beliefs, which require concepts for the subject to form them.

Intentionality refers to the fact that experience is about, in reference to, or *of* the external world. Included in these are *phenomenal objects*, such as a pain in leg_x at time. This is a critical distinction, as placing the objects of phenomenal experience in the physical world lends nicely to Tye’s Physicalist outlook, while still accounting for phenomenal experience and *what it’s like* to have an experience from a subjective point of view. For Mary, there is no appropriate phenomenal content for her to refer, and thus, she cannot learn the phenomenal concept ‘red’ until she adopts the appropriate experiential perspective. A caveat to Tye’s argument, though, is that he requires color realism; he requires color to be *out there* in the world with pains. If color itself cannot be considered *external phenomenal content*, then *perspectival Physicalism* faces a whole new set of problems as regards the knowledge of Mary.

Content refers to the fact that the referent of phenomenal experience is not the experience itself, but something external to it. When one finds themselves in the proper physical relationship to some *Phenomenal Content*, it is imposed on them through their senses, and they gain a representational token, or learn a *Phenomenal Concept* about that content. Due to the fact that phenomenal

According to Tye, “...knowing what it is like to undergo an experience requires one to have a certain experiential point of view or perspective” (Tye, 12). Per the thought experiment and Tye’s explanation of phenomenal knowledge, Mary’s complete knowledge of the physical facts about the world must also contain the knowledge that a subject acquires from this certain experiential point of view or perspective. Of course, the Physicalist and the non-Physicalist usually agree that the processes by which these subjective experiences arise are physical. Certain brain states and functions are caused by sensory stimuli (in this case, light waves stimulating the cones in the retina of the eye and converting that stimulus into signals to be interpreted by the brain), which allow for phenomenal experience to occur. Why, though, is there some necessarily subjective *experience* that accompanies these functions? For Tye, as noted, phenomenology is not analogous to functional brain states or neurophysiology. Even if the neurophysiological states are explained, there is something left out of the picture: *what it’s like* to undergo some phenomenal state ‘P.’

Tye refers to this problem as the *explanatory gap* between the neurophysiological and the phenomenal. He invites any inquirer to, “Peer as hard as [they] want at the neurons. Probe deeper and deeper into the structure of the brain. According to the PANIC Theory, [they] will not find any phenomenology. It simply is not there” (Tye, 162). In a similar notion, Hillary Putnam expresses the fact that one can, “... ‘Cut the pie any way you like, [...m]eaning just ain’t in the head” (Putnam, 227). This is a basic reiteration of a familiar concept: subjective experience is not available to a third-party inspection, as Tye says, in the relevant objective way. In other words, phenomenology as it is being referred to by the non-Physicalist is not entailed in the physical facts of the world because it is not being referred to as being *in the world* at all.

In Tye's description of the *explanatory gap* and the problems which revolve around it, he makes it clear that a major part of the issue is the assumed location of phenomenology. Where *phenomenal concepts* are learned *about* the phenomenal content in the external world, *phenomenal content* is strictly external to the subject—even in the case of imaginary or fake content, such as phantom limbs. Pains are *in arms*, whereas a *phenomenal concept* is a token representation of that pain which accounts for its being *in the arm*. Phenomenal content is properly poised to be represented via the spectrum which is prescribed by the PANIC Theory.

According to the PANIC theory, Mary can learn a new *phenomenal concept* by being put in a proper physical perspectival relationship with some *phenomenal content*, such as a red rose, for the first time without rendering Physicalism false. This is due to the prospect of the multiple realizability of PANICs. Due to the nature of PANIC states, there can be different lower-level facts which account for the same higher-level realization, and vice versa. Tye explains realizability as follows:

“The general relationship that obtains between higher-level and lower-level types is ones of *realization*. Higher-level property times have *multiple* lower level realizations. ... any object that has the higher-level property, or is an instance of the higher-level type, does so *in virtue of simultaneously* having one of the lower-level properties or types that realizes it. A gas has a certain temperature at time *t*, for example, in virtue of its having mean molecular energy at time *t'*” (Tye, 41).

If the facts regarding phenomenal experience are strictly physical, yet are only acquirable by means of subjective experience, perhaps Mary learns something *about the experience of others* after all; yet, those facts she learns *about the experience of others* must also be physical facts about the world. Tye posits that, “...phenomenal states cannot be counted as physical states in any sense...unless there is an appropriate mechanism for perspectival subjectivity” (Tye, 161).

Tye refers to this view as *perspectival Physicalism*. We can see that it lends to the possibility for Physicalism to accommodate necessarily private physical facts. This is essential to placing the burden on the claimant that Physicalism is false; if Physicalism can account for brute-force necessarily private facts about experience as *physical*, then Mary can learn something *about the experience of others*, and it is not a direct threat to Physicalism.

The argument for *perspectival Physicalism* is as follows:

- (P1) Fully understanding the essential nature of any phenomenal state ‘P’ requires knowing what it is like to experience ‘P.’
- (P2) Knowing what it is like to experience any phenomenal state ‘P’ requires adopting a certain experiential perspective.
- (C3) Fully understanding the essential nature of any phenomenal state ‘P’ requires adopting the appropriate experiential perspective.

C3 above “...expresses the thesis of *perspectival subjectivity*, and thus, argues for *perspectival Physicalism*” (Tye, 166).

The *explanatory gap* is always a fundamental hurdle for the Physicalist; for Tye, however, the *explanatory gap* can be closed in a more traditional way as long as phenomenology is relocated *out there* in the world, instead of *in here* in the brain. The problem is, if the world is comprised of and only of the physical, and physical facts are usually reducible to and deducible by other objective physical facts, where is a necessarily subjective phenomenology's place in it?

In Chapter 1, I suggested that there are two possibilities regarding physical facts and their ownership: either (i) all physical facts are necessarily reducible to or deducible by other physical facts or (ii) it is just a brute fact of nature that some knowledge is necessarily private. In the case of (ii), I argued that a Physicalist who adopts this view must provide a mechanism by which these facts are both privately known and a reason for which they remain necessarily private. Tye provides what he considers to be a possible mechanism for both conditions, and his account of phenomenology as being comprised of necessarily private physical facts is an excellent example of what a mechanistic account physical phenomenology might look like. It is supposed to be the end-all refutation of Physicalism that Mary knows all of the physical facts, yet still learns something *about the experience of others* when she leaves her black-and-white room. If it can possibly be the case that some physical fact about the world maintains necessary privacy and non-deducibility by the other physical facts about the world, yet still pertain to some phenomena that meets the criteria for being *physical*, then the Knowledge Argument's *learning claim* does not seem to threaten Physicalism after all.

It is unnecessary for the Physicalist commit to closing the *explanatory gap* in the way in which non-Physicalists have been trying to force. The Physicalist is not committed by the intuitions of the Knowledge Argument to the full a priori deductibility of all knowledge about the world from the totality of the fundamental facts alone. It could be a brute fact of nature that some physical facts maintain perspectival subjectivity as a necessary mechanism for learning them. However, whether he realized it or not, it does seem that Tye has described an account of phenomenal content that *is* accounted for by reducible microphysical facts, as long as the person doing the reduction has already had the experience *and* knows all of the fundamental physical facts about the world.

If *phenomenal content* is external to the subject, and all of the facts regarding the occurrence of that phenomenal state or the object of it which is being represented by *phenomenal concepts* are physical facts, then there is no difference between explaining the physical facts about experience and the physical facts about, the boiling point of water, the hardness of a diamond, or the temperature of a gas. There is some neurophysiology to be done in explaining the acquisition of *phenomenal concepts*, but if the non-Physicalist is going to allow for other concepts about the world to fail to refute Physicalism, then there is no reason why the concepts learned about some physical, external content would be any different. They would be learned as *phenomenal content* according to the subjective perspective necessary. Tye's approach has many virtues when it comes to surmounting the *explanatory gap*, or at the very least, characterizing it such that it is demystified in a helpful way.

Phenomenology, for Tye, is also causally efficacious. On his view, it is obvious that tokens of phenomenal representation causally affect our behavior. The causal efficacy in this case works to eliminate the possibility of certain instances of Epiphenomenalism. For any non-Physicalist who takes Epiphenomenalism to be the most appealing alternative to Physicalism, it seems that an account of certain phenomenal states being key causes for belief formation which guides our behavior is a refutation of those theories that rely on non-causal, non-physical 'qualia.'

Tye's flavor of Representationalism can respond to the key point of the Knowledge Argument's intuition generator—that she know *all* the physical facts, yet learns something *about the experience of others*—by relying on a completely physical account of both the world and the mechanisms of experience. Mary can be in possession of *all* the physical facts about the nature of the world, still have something to learn, and Physicalism does not necessarily fail. Once again, we can see how the burden is shifted onto the non-Physicalist to show, beyond intuition, exactly *why* the facts about experience are completely different in kind from other facts about the world such as the hardness of diamonds or the temperature in gasses.

Chapter 3 – Fundamental Consciousness

In *The Conscious Mind* (1996), Chalmers expresses Mary's epistemic situation as such:

“Imagine that we are living in an age of a completed neuroscience, where we know everything there is to know about the physical processes within our brain responsible for the generation of our behavior. Mary has been brought up in a black-and-white room and has never seen any colors except for black, white, and shades of gray. She is nevertheless one of the world's leading neuroscientists, specializing in the neurophysiology of color vision. She knows everything there is to know about the neural processes involved in visual information processing, about the physics of optical processes, and about the physical makeup of objects in the environment. But she does not know what it is like to see red. No amount of reasoning from the physical facts alone will give her this knowledge” (Chalmers, 103).

This particular presentation of Mary in her black-and-white room is coaxing our intuitions in a very specific direction. Chalmers wants to paint a picture where the kind of facts which his expression of ‘physical facts’ entails are neurophysiological functions, the processing of visual information, the structures and biology of human eyes, the physics of optics, and the ‘physical makeup of objects in the environment.’ Chalmers specifically states that ‘no amount of reasoning from the physical facts alone will give her...knowledge [of *what it's like* to have some experience]. In order to argue for *why* she cannot do this deduction, he relies on *supervenience*.

Chalmers takes the doctrine of MATERIALISM (or Physicalism) to be, “...the doctrine that the physical facts about the world exhaust all the facts, in that every positive fact is entailed by the physical facts” (Chalmers, 124). For Chalmers, the physical facts do not and cannot exhaust the facts about the world, as the facts of phenomenal consciousness are entirely different in kind. While the facts about consciousness may *naturally* supervene on the physical facts of the world, they are not akin to the physical facts, nor can knowledge of the physical alone provide the explanatory power necessary to close the *explanatory gap*. Chalmers takes Materialism to be a doctrine which claims that the facts of phenomenal consciousness *globally logically* supervene on the physical facts. Chalmers explains supervenience as follows:

“Logical and natural supervenience have quite different ramifications for ontology: that is, for the matter of what there is in the world. If B-properties are logically supervenient on A-properties, then there is a sense in which once the A-facts are given, the B-facts are a free lunch. Once God (hypothetically) made sure that all the physical facts in our world held, the biological facts came along for free. The B-facts merely redescribe what is described by the A-facts. They may be *different* facts (a fact about elephants is not a microphysical fact), but they are not *further* facts.

With mere natural supervenience, the ontology is not so straightforward. Contingent lawful connections connect distinct features of the world. In general, if B-properties are merely naturally supervenient on A-properties in our world, then there *could* have been a world in which our A-facts held without the B-facts. As we saw before, once God fixed all the A-facts, in order to fix the B-facts he had more work to do. The B-facts are something over and above the A-facts, and their satisfaction implies that there is something new in the world” (Chalmers, 41).

Some clarity about *supervenience* can be found in imagining an alien species of giants from planet Basketball. It is a world inhabited by giants whose entire lives are spent repetitively tossing orange spheres into nets with holes in the bottom. They do not track the amount of times that the spheres successfully enter the nets, nor are they engaged in any kind of competition. A world where ‘basketball’

exists, yet no score is kept, is quite simple to imagine. Were the concept to ever arise, even though it would be a *different* fact, it would not be a *further* fact; it would just be a token representation regarding the microphysical facts and occurrent events. Conversely, it seems difficult to imagine a world where a ‘score’ of a basketball game exists without the game itself. What distal property would that *different* fact be tracking? On my view, even ubiquitously acceptable accounts of non-physical facts, such as the score of a basketball game, ought to be considered physical facts due to their being explainable in physical terms and their direct covariation and reliance on microphysical facts about the world which they represent.

Chalmers uses the distinction between *natural* and *logical* supervenience to make a distinct point about phenomenal consciousness: On his view, the B-facts—the facts about consciousness and qualia—are something over and above the A-facts. Like basketball, the game—or the brain and the physical facts about its functionality—may be required for consciousness to arise, but unlike the score of the game, Chalmers can imagine a physically duplicate world that exists without consciousness, and a physically different world where the facts of consciousness remain the same. In this way, the facts about consciousness, for Chalmers, do not *globally logically* supervene on the microphysical facts about our world, even if they *naturally* supervene on it. Chalmers is not denying that consciousness is *in the world*, but is arguing a kind of *property dualism*, similar to an early Jackson’s Epiphenomenalism. Except, of course, for Chalmers, ‘qualia’ have causal power.

What he is hoping to accomplish by denying *global logical supervenience* of the facts about consciousness on the microphysical is to show that the facts about qualia are irreducible to the physical facts. For Chalmers, the *explanatory gap* is indefinitely wide. Without logical supervenience, reducibility becomes impossible. According to Chalmers, “For almost every natural phenomenon above the level of microscopic physics, there seems in principle to exist a *reductive explanation*; that is, an explanation wholly in terms of simpler entities. In these cases, when we give an appropriate account of lower-level processes, an explanation of the higher-level phenomenon falls out” (Chalmers, 42).

In order for a phenomenon to be reductively explainable in terms of lower-level properties, it must be *logically supervenient* on those lower-level properties. If consciousness were to be reductively explainable in terms of physical facts, then consciousness would have to be the ‘free lunch’ that falls out of the A-facts, or the microphysical facts. In short, if some precise combination of physical facts come to be the case, then consciousness comes along with it, for free, every time. Every time one puts wires connected to a battery on one end and a bulb on the other, if the battery is compatible with the bulb, the bulb will turn on.

Establishing this kind of supervenience as a necessity for Physicalism to hold can be a hard intuition draw towards the non-Physicalist attitude. In theory, if a human being was to be replaced with functional duplicate parts, piece by piece, all memories and necessary ‘software’ included, the Physicalist would have to accept that that person would remain *that person*, consciousness and all, if consciousness *logically* supervenes on the physical—and thus, the psychofunctional and neurophysiological—facts about the world and the brain and human body. Why should this be so worrisome, aside from the intuition that *something* must change about that person? If memories are transferred and all information held inside of the brain is also transferred, why would we think that some qualitatively different *person* would arise? Same hard drive backup, different motherboard.

He suggests that it may be helpful to consider phenomenal facts as basic, fundamental feature to the world. This is similar to when, “...Maxwell introduced new fundamental electromagnetic laws. Only this way could the phenomena be explained. In the same way, to explain consciousness, the features and laws of physical theory are not enough” (Chalmers, 127).

The problem with this analogy is that the fundamental laws of electromagnetism are *physical* laws, and express one of the four fundamental forces in physics: gravitation, electromagnetism, weak, and strong interactions. These are said to be fundamental because they cannot be further reduced to more basic interactions. It seems that if Chalmers is to be taken at his word, then consciousness and the facts about it cannot be further reduced. Yet, they must still be *explicable* at their own fundamental level. Thus, the *explanatory gap* remains unacceptable, especially in the case of Mary, who is in possession of all of the physical facts about color experience. We might ask Chalmers the question, “Does Mary lack knowledge of any electromagnetic properties which govern some aspects of consciousness, or if they do not, does she *know* that they do not?” An answer to a question such as this would allow for a more genuine inquiry into what he means by his claim that the facts regarding phenomenal experience require a more fundamental law that is similar to the laws of electromagnetism.

It would be interesting, given the chance, to ask Chalmers another question: What of her *legitimate observations*? What of those experiences which, by all accounts, Chalmers’ included, she would absolutely know *what it is like* to subjectively undergo some phenomenal state? Would Mary’s introspection regarding her *legitimate observations*, when coupled with those observations which she knew to be illegitimate in some sense—perhaps through comparison of her own neurophysiology when looking at a rose on a black and white screen versus the neurophysiology of someone looking at the same rose outside of her color-restricted room—give her (or us) any insight into the kinds of factual deductions that she could perform?

As noted, Chalmers is very explicit in his claim that no amount of reasoning will lead Mary from a state of having complete knowledge of the physical facts regarding color and how it affects the brain to knowing *what it is like* to see red. As such, it follows for Chalmers that Physicalism is false. His view is much more substantial when it comes to qualia and its properties; so much so that his view argues for the necessity of some fundamental, irreducible facts of conscious experience. In the same way that physics is an account of the microphysical facts, a theory of consciousness may very well be looking for a detail of the micro-experiential facts. Even if conscious experience cannot arise without the physical, for Chalmers the facts about it are set apart from the facts about the physical nature of the world. Physics will never have the explanatory power in regards to consciousness and experience that it does when explaining the hardness of diamonds, the liquidity of water, or the temperature of gasses. The issue is that the explanations of these phenomena are a posteriori derivations from a priori intentional claims, and as such, perhaps they are more similar in kind than intuition can show.

In *Phenomenal Concepts and the Knowledge Argument (2004)*, Chalmers explains his response to the Knowledge Argument in very simple terms:

“...[He] assumes *phenomenal realism*: roughly, the view that Mary acquires new factual knowledge (not a priori deducible from physical knowledge) when she sees red for the first time. This excludes views on which Mary merely gains a new ability, or on which she gains no knowledge at all. It is compatible with views on which Mary gains knowledge of an old fact in a new way. The important aspect of this view is that it allows for an *epistemic gap* between physical truths and phenomenal truths, in the sense that phenomenal truths are not entailed a priori by physical truths” (*Chalmers, Phenomenal Concepts and the Knowledge Argument*).

As we have seen, the *explanatory gap* is the axis around which the non-physical argument rotates both because of the intuitions regarding the facts of experience which arise when the *explanatory gap* is considered, as well as the inability of the Physicalist at this point in our understanding to close it using explicitly physical facts. Chalmers’ use of *a priori* in this excerpt is of particular import; Mary, by this account, can be in possession of all of the physical facts that are possibly acquirable from an objective perspective, yet cannot deduce the facts about experience_a without subjectively having experience_a

herself. He also goes on to assert that this view can accommodate any Physicalist theory which relies on perspectival subjectivity, such as Tye's PANIC theory, with no issues. Where he disagrees with Tye is the metaphysical location of phenomenology; for Chalmers, they are 'in' the mind and run parallel to functional aspects of the brain, and are *logically supervenient* on fundamental micro-conscious facts.

Again, we see how the facts about *what it is like* to have an experience can be potentially cashed out under the umbrella term 'physical,' but with stipulations regarding why it is that Mary cannot know them. This is why we must either stipulate a new version of the thought experiment wherein Mary only knows those physical facts which are possibly acquirable by a first-person subjective mode, or whether she is in possession of *all* of the physical facts, period. Still, the Physicalist would do well to hold off on the suggestion that, when all is said and done, the facts of subjective experience will be completely irreducible to the fundamental physical facts until shown otherwise. For Materialism, the facts about consciousness, and everything else, are 'free lunch' when all of the micro-physical facts are fixed and the required physical systems come into the proper organization.

They may have their own fundamental laws *within* the physical world, similar to gravity and electromagnetism, but the mystical *explanatory gap* does not arise between those two, and so why should the facts of phenomenal experience be expected to maintain the inexplicability that they currently have ad infinitum? What is missing is a mechanism that must be put in place in order to do the reduction or comparison between facts which are normally only acquirable by means of first-person subjective experience and third-person objective observation, and perhaps a later Jackson's Representationalism can show one way in which this can be accomplished. We can see how it is important to meet the demands of a thinker such as Chalmers when contemplating what it is that Mary does or doesn't learn when she leaves the room for the first time and sees her first rose.

By analyzing Chalmers' claims of impossibility regarding the deduction of the subjective physical facts about experience by the objective physical facts about the world that we can appreciate the necessity for a mechanism of reducibility by which Mary would be able to connect the objective physical facts about the world to the facts of experience *a posteriori*. After all, as a later Jackson argues, these *a posteriori* explanations are the same kinds of explanations we find acceptable in all other physical cases. As such, the burden is on the non-Physicalist to show why the facts of experience are completely *a posteriori* irreducible to the microphysical facts about the world. In Chalmers's terms, the facts about phenomenal consciousness can fail to *globally logically supervene* on the microphysical facts and be non-deducible by them, but ultimately be reducible to them. The Knowledge Argument alone, nor the intuitions summoned by it, seem to refute that possibility at substantial length.

Chapter 4 – The Blue Banana

Dennett's approach to the Knowledge Argument in *Consciousness Explained* (1991) is not only an argument in defense of Physicalism, but also something of a reprimanding of any philosopher who even momentarily entertains the possibility of Physicalism failing on its account. For Dennett, anyone who fails to accept that Mary could discriminate between different colors when she leaves her black-and-white room for the very first time simply isn't following the directions of the thought experiment. They have fallen victim to what he calls Philosopher's Syndrome: "Mistaking a failure of the imagination for an insight into necessity" (Dennett, 401). The thought experiment is presented in such a way that intuition and a lack of imagination regarding what Mary *really* knows—the depth of her knowledge if she does, in fact, possess a complete set of physical facts—prevents one from doing so. First, he presents another version of the thought experiment, through the words of Jackson. Then, he makes a slight addition:

"Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black-and-white room via a black-and-white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like red, blue, and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue.' ... What will happen when Mary is released from her black-and-white room or is given a color television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false. ... [p.128]" (Dennett, quoting Jackson, 398).

Dennett's presentation of the thought experiment, however, is in stark contrast to how it is presented by others who would respond to it due to what he calls a 'surprising—but legitimate—[addition]':

"And so, one day, Mary's captors decided it was time for her to see colors. As a trick, they prepared a bright blue banana to present as her first color experience ever. Mary took one look at it and said, 'Hey! You tried to trick me! Bananas are yellow, but this one is blue!' Her captors were dumfounded. How did she do it? 'Simple,' she replied. 'You have to remember that I know everything—absolutely everything—that could ever be known about the physical causes and effects of color vision. So of course before you brought the banana in, I had already written down, in exquisite detail, exactly what physical impression a yellow object or a blue object (or a green object, etc.) would make on my nervous system. So I already knew exactly what thoughts I would have (because, after all, the 'mere disposition' to think about this or that is not one of your famous qualia, is it?). I was not in the slightest surprised by my experience of blue (what surprised me was that you would try such a second-rate trick on me). I realize it is hard for you to imagine that I could know so much about my reactive dispositions that the way blue affected me came as no surprise. Of course it's hard for you to imagine. It's hard for anyone to imagine the consequences of someone knowing absolutely everything physical about anything!'" (Dennett, 399-400).

Besides the sarcastic tone with which it is presented, Dennett is portraying a very different Mary in this example. This is a very Dennettian Mary, one that takes on some of his own personality traits. Dennettian Mary, however, is really no different from the Mary presented in any other version of this thought experiment. How can that be? She was confined to the same room, under the same circumstances, and had the same knowledge as every other time she emerged into the colorful world outside of her

restricted existence. What is so different about Dennett's Mary that allows her to have this knowledge of color before having actually experienced color from the external world? For Dennett, the failure of imagination is not a failure of Mary's but a failure of ours, the philosophers. We simply cannot imagine what the totality of physical facts of the world would allow the knower of them to *know*.

If the facts about color are completely physical and mind independent—which is still being debated amongst the philosophers concerned with the metaphysics of perception—and the facts about the processes by which experience occurs are completely physical, it is the burden, then, of the non-Physicalist to show why, and more importantly how the knowledge obtained in *having an experience*, or *what it is like* to see a color, can possibly be anything but contained inside of the collection of the totality of physical facts. The hard challenge for a reductive Physicalist is to show how the facts about color experience can be reduced in a transmittable way such that the *explanatory gap* is effectively closed in the traditional way. Or, as I have shown, that it cannot be because of the perspective through which it must necessarily be learned.

Still, there is a lack of a useful mechanism in Dennett's response which can be used to *prove* the reducibility of the facts of experience to fundamental physical facts. There is just a suggestion of an utterance by Mary wherein she tries to explain that there *is* such a mechanism in place. It seems plausible that Mary would make such an utterance, especially if such a mechanism did exist. Where Dennett fails to provide the mechanism itself, it is still a very helpful addition in shifting the intuition of the truth of the Knowledge Argument's conclusion. I do not think that it was Dennett's intention to provide a solid defense of Physicalism with this refutation. It is more than likely a slightly rhetorical response to the Knowledge Argument and a rejection of it as a meaningful thought experiment altogether. Either way, he succeeds in adding a new dimension to the pool of imaginative aspects regarding Mary's knowledge before and after she leaves the room. Fortunately, as I have shown, it is the burden of the non-Physicalist to prove why the current *explanatory gap* threatens the thesis of Physicalism ad aeternum. To this point, I cannot see how phenomenology strictly meets some criteria that refutes the thesis of Physicalism, even with the intuitions of the Knowledge Argument in play.

Ch. 5 – What Relevant Information Are We Waiting On?

After accepting Physicalism as a true doctrine, Jackson authored *Mind and Illusion* (2003) in order to provide (i) a clear of the commitments of Physicalism, (ii) upon what grounds he finds the Knowledge Argument to fail, and (iii) a new account of what he asserts that Mary will or will not learn when she leaves her room and has her restrictions lifted. Jackson posits an account of Physicalism that somewhat dissents from a view such as Tye's; he claims that it is not experience that is lacking some feature (concepts) as opposed to belief, but that it is the other way around. Phenomenal experience must have some feature which belief lacks; an *inextricable richness*.

Jackson begins the reproach of his original stance regarding the conclusion of the Knowledge Argument with what he calls a 'brief rehearsal' of his thought experiment:

“A brilliant scientist, Mary, is confined in a black-and-white room without windows. She herself is painted white all over and dressed in black. All her information about the world and its workings comes from black-and-white sources, like books without colored pictures and black-and-white television. She is, despite these artificial restrictions, extraordinarily knowledgeable about the physical nature of our world, including the neurophysiology of human beings and sentient creatures in general, and how their neurophysiology underpins their interactions with their surroundings. Can she in principle deduce from all this physical information, what it is like to see, say, red?” (Jackson, *Mind and Illusion*, 2003: pp.2-3).

In this presentation, Mary is 'extraordinarily knowledgeable about the physical nature of our world,' yet no longer explicitly holds *all* of the physical information. He is also asking a very specific question which is slightly different than the usual inquiry: Can Mary '...in principle deduce from all this physical information, what it is like to see, say, red?' Jackson also reframes the question about Mary in terms of abilities as opposed to knowledge. Jackson is faithful to his original claims, however, and Mary knows *all* of the physical facts about the world.

In arguing for Representationalism, Jackson asserts, "...that although physicalists are committed to the experiential being necessitated by a rich enough physical account of our world—otherwise it would take more than the physical nature of our world to secure its experiential nature, contrary to physicalism—they are not committed to the experiential being a priori derivable from the physical" (Jackson, 4). This is a helpful clarification of the Physicalist's commitments regarding experience: the experiential is necessitated by a physical account of the world, and the Physicalist is not committed to the experiential being a priori derivable from the physical. In short, Mary can know all of the facts about the physical nature of the world and still fail to deduce 'what it is like to see red.' Where Jackson is no longer committed to the possibility of a *a priori deduction* of 'redness' for Mary, he is now committed to a *posteriori reducibility*.

In a reassessment of his original claims, Jackson finds the Knowledge Argument to fail by 'mistaking intentional properties for instantiated properties.' This can be understood through the 'I' in Tye's PANIC theory: experiences are *of* something. Experiences are not instantiated objects themselves, nor is the redness experienced when looking at a red rose *instantiated* in the experience in an objective way. It is a subjective representation of the way the world *might* be. Whether or not one ought to believe that that is the way the world is can be debated, but to speak of the redness of experience as instantiated makes no sense. Experiences cannot themselves *be* red, but only *of* red.

Jackson compares 'redness' in objects to properties such as 'temperature' in order to make a point regarding moving from an *a priori* claim to an *a posteriori* one. This is an attempt to narrow the

explanatory gap and further demystify the problem. In this case, a kind of *color realism* is required; color must be a property of *objects*, and more importantly, mind independent for a comparison such as this to gain traction in the debate. That issue aside, Jackson claims that:

“Our belief that gases have temperature and pressure is grounded in their behavior. Moreover, we know that their behavior is fully explained by the various features recognized and named in the kinetic theory of gases. There is no need to postulate any extra features of gases in order to explain their behavior. This makes it very hard to hold that no matter how much information we have framed in the terms of the kinetic theory and in terms of the functional roles played by the properties picked out by the terms of that theory, and no matter how confident we are that the kinetic theory and its future developments provide a complete picture in the relevant respects of the essential nature of gases, the passage from this information to whether or not gases are hot and have pressure is *a posteriori*. What relevant information are we waiting on? We know that all we will get is more of the same. Skepticism about gases having temperature and pressure threatens if we insist that we cannot go *a priori* from the molecular account of gases and the concomitant functional roles to gases having temperature and pressure” (Jackson, 5).

In short, Jackson suggests that we ought to accept the explanation because of the following argument:

(P1) Temperature in gases is that which does so and so (a priori premise about the concept of temperature).

(P2) That which does so and so is mean molecular kinetic energy.

(C3) Therefore, temperature in gases is mean molecular kinetic energy.

Jackson is making a powerful point when asking, “What relevant information are we waiting on?” This question places the burden firmly on the non-Physicalist in this debate. That is not to say that Jackson’s position is correct. At least, however, he is presenting the non-Physicalist with a coherent question that, if answered, can shed light on the debate as a whole. As previously suggested, *if* she in fact has *all* of the physical information in her possession, and knowing *what it is like* is physical information, then she will absolutely be able to deduce it due to her more-or-less physical omniscience. However, I suspect that this is an unsatisfactory response. Anyone inquiring into what it is that Mary does or does not learn upon leaving the room and having her first legitimate phenomenal color experience is much more interested in her predicament as I describe it below. I propose a shorter, clearer version of Mary’s epistemic position:

There is an introvert, Mary, who is happily secluded from the world in a black-and-white room. She has not—and for whatever reason you like—cannot have a phenomenal color experience while she is in the room of any colors but black, white, and shades of gray. She knows all the facts about the physical nature of the world which can possibly be acquired based on her current subjective perspective. If Mary begrudgingly left her room, had her restrictions lifted, and saw a red rose in bloom for the first time, would she learn some genuinely new fact about the world? If she would, then Physicalism might be false.

This version of Mary, introverted and happy to be in her room, leaving only when she *must*, can provide a different set of intuitions altogether. Surely, though, that cannot change the actual outcome of the experiment. While this has been an interwoven theme throughout this work, an explicit statement of the downside of intentional intuition pumps is helpful. Implicit intuition pumps aside, what does this generation of Mary’s leaving her restricted room tell us about the Physicalist stance? For one, we can see why *perspectival Physicalism* is an attractive position for Physicalists due to its claims regarding the possibility of some necessarily private facts. This shifts the burden onto the non-Physicalist to show why those facts *must be* about some non-physical aspect of the world. If there is something that Mary learns

about the world, she might have failed to learn it in the room due to the subjective perspective she was in; namely, a subjective perspective absent of *colored phenomenal content*.

On the other hand, we also know that necessary privacy of some physical facts—phenomenal or otherwise—is generally outside the scope of the thesis of Physicalism. For Jackson, the necessary irreducibility of facts about phenomenal experience to fundamental physical facts cannot be permitted. They may not be a priori deducible, but once they are a posteriori known, then they can be reduced. We can apply this criterion by revisiting his analogy of the referent of phenomenology to the facts about the temperature in gasses. On Jackson's view, if Mary were to learn some genuinely new fact about the world that was completely irreducible to the total set of fundamental physical facts which she already knows, perhaps *then* Physicalism fails on account of the Knowledge Argument, but not before.

In order to further deny the validity of the Knowledge Argument, Jackson argues for the possibility that Mary could be wrong about her own knowledge. She could learn an intentional fact of something, gain a representational state of *inextricable richness* that was completely reliant on the physical facts about the world, and still—wrongly—believe that she had learned some genuinely new fact when she had not. Thus, Mary cannot be the arbiter of the truth value of Physicalism herself. Jackson responds to the *learning claim* in *Mind and Illusion* as follows:

“If feel is a matter of immediacy, inextricability, and richness of representational content, and the right kind of functional role, the difference is that, after her release, Mary has representational states with all those properties. If she makes the mistake of conflating intentional properties with instantiated properties, she will think that she has learned something new about how things are, but she'll be wrong. Rather, she is in a new kind of representational state, different from those she was in before. And what is it to know what it is like to be in that kind of state? Presumably, it is to be able to recognize, remember, and imagine the state” (Jackson, 26).

The *ability argument*, or the idea that what Mary gains is some ability and not some genuinely new knowledge, is not necessarily tied up with Representationalism. The language used by Jackson maintains ‘the representational state’ as a representational token of the physical world. Utilizing this claim, Jackson argues that Representationalism has done the hard work in explaining phenomenal experience as being strictly physical.

Between Tye's account of *perspectival Physicalism* and Jackson's Representationalism, we can see how the facts of experience are probably more similar to the facts about the temperature of gasses than the Knowledge Argument might originally lead one to intuit. Even if the Physicalist is not required to commit to the possibility of a *a priori deduction* of the facts of experience, he shows how the Knowledge Argument's mistaking *intentionality* for *instantiation* can allow for a misclassification of the facts about phenomenal experience. In unearthing this potential mistake, he shows how even a physically omniscient color scientist could fail to connect the dots regarding the objective physical facts about phenomenal experience. Again, we see the burden of the debate hoisted onto the claimant that Physicalism is false. This time, by asking, “What relevant information are we waiting on?” (Jackson, 5).

Conclusion

It has been the continual goal of this work to persuade the reader that the Knowledge Argument is not a *prima facie* refutation of Physicalism. Representationalism (in the vein of Tye and a later Jackson) can provide accounts of why some physical facts require adopting a ‘certain experiential perspective’ to know them, as well as how Mary could be physically omniscient and still fail to recognize that something she learned was not necessarily a genuinely new fact, but an old fact in a new way. By analyzing their views, we see that they can shift the burden of two main claims against Physicalism as regards phenomenology:

- (i) Physicalism fails if the facts of phenomenal experience are not a priori derivable from a complete possession of the fundamental physical facts about the world.
- (ii) The facts which explain phenomenal experience are irreducible to fundamental physical facts.

It ought to be a primary concern of the Physicalist view that the necessarily first-person subjective acquisition of certain facts be eliminable as a kind of synonym for facts about non-physical aspects of the world. The Physicalist will further preserve the integrity of their thesis if they can prove that this distinction does not do any work in giving special classification to some facts. It is the usual availability of instances of phenomenal knowledge to only an individual which has undergone a specific subjective phenomenal experience as opposed to the availability of those facts to multiple persons in an objective sense which is generating the intuitions of the Knowledge Argument’s conclusion. Moreover, I have shown that it does not necessarily follow from that distinction alone that the facts regarding *what it is like* to have an experience are non-physical; it also does not show that they are physical.

If Physicalism holds that the physical facts account for the totality of metaphysically possible facts about our world, then it must fail on any account in which ubiquitously non-physical facts are learned by Mary. Simultaneously, it must fail on any account wherein a true totality of facts are possessed, yet there is some genuine aspect of the world which is unknown. As a species, even though we lack the knowledge of a completed set of physics which is gifted to Mary in the Knowledge Argument, it seems as though we may be closer to showing that Physicalism holds in the case of phenomenal experience than we might intuit. Ultimately, the non-Physicalist requires only one instance of a knowable fact about a non-physical entity to refute the Physicalist thesis as a whole. One such instance can only occur when the criterion for ‘physical’ has been established according to some acceptable Physicalist thesis, and as previously noted, I cannot see how phenomenology can fail to meet that criterion once established.

What the Knowledge Argument reveals is that we have a long road ahead of us in coming to a conclusive answer on whether or not the subjective facts about experience are a priori deducible by or completely reducible to the objective physical facts about it, but it seems as though it is a real possibility that they might be. What it does not do is refute Physicalism by generating intuitions about what the facts about experience are. If we accept representationalism as plausible, and Mary knows all of the physical facts, she would *know* that the blue banana was blue, even if she was not in possession of the perspectival representation of it before seeing it for the first time. She would know it was blue in the same way that, given the mean molecular kinetic energy of a gas, she would know that gas was properly poised to present itself as warm via a *phenomenal concept*. If we allow one as an acceptable explanation, we ought to allow the other. As far as Mary goes, what other relevant information are we waiting on?

Works Cited

- Block, Ned (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science* 9:261-325.
- Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, David J. (2004). Phenomenal concepts and the knowledge argument. In Peter Ludlow, Yujin Nagasawa & Daniel Stoljar (eds.), *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. MIT Press. pp. 269.
- Chalmers, David J. & Jackson, Frank (2001). Conceptual analysis and reductive explanation. *Philosophical Review* 110 (3):315-61.
- Dennett, Daniel C. (1991). *Consciousness Explained*. Penguin Books.
- Harman, Gilbert (1990). The intrinsic quality of experience. *Philosophical Perspectives* 4:31-52.
- Horgan, Terence E. (1984). Jackson on physical information and qualia. *Philosophical Quarterly* 34 (April):147-83.
- Jackson, Frank (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32 (April):127-136.
- Jackson, Frank (1986). What Mary didn't know. *Journal of Philosophy* 83 (May):291-5.
- Jackson, Frank (2003). Mind and Illusion. *Royal Institute of Philosophy Supplement* 53:251-271.
- Ludlow, Peter ; Nagasawa, Yujin & Stoljar, Daniel (eds.) (2004). *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*. MIT Press.
- Metzinger, Thomas (2003). Phenomenal transparency and cognitive self-reference. *Phenomenology and the Cognitive Sciences* 2 (4):353-393.
- Nagel, Thomas (1974). What is it like to be a bat? *Philosophical Review* 83 (October):435-50.
- Nordby, Knut (1990). Vision in a complete achromat: A personal account. In R. F. Hess, L. T. Sharpe & K. Nordby (eds.), *Night Vision: Basic, Clinical and Applied Aspects*. Cambridge University Press.
- Polger, Thomas W. (2008). H₂O, 'water', and transparent reduction. *Erkenntnis* 69 (1):109-130.
- Putnam, Hillary (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science* 7:131-193.
- Raymont, Paul (1995). Tye's criticism of the knowledge argument. *Dialogue* 34 (4):713-26.
- Siewert, Charles (2004). Is experience transparent? *Philosophical Studies* 117 (1-2):15-41.

- Stich, Stephen P. (1992). What is a theory of mental representation? *Mind* 101 (402):243-61.
- Tye, Michael (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. MIT Press.
- Tye, Michael (2000). Knowing what it is like: The ability hypothesis and the knowledge argument. In Gerhard Preyer (ed.), *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*. Rowman & Littlefield.
- Tye, Michael (1991). *The Imagery Debate*. Cambridge: MIT Press.
- Tye, Michael (2009). *The Metaphysics of Mind*. Cambridge University Press.
- Valaris, Markos (2014). Self-Knowledge and the Phenomenological Transparency of Belief. *Philosophers' Imprint* 14.