

THE KNOWLEDGE ARGUMENT AND THE PROBLEM OF QUALIA

A Thesis

Presented to

The Faculty of the Department

of Philosophy

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

Ahmet Kadir Uslu

May, 2016

THE KNOWLEDGE ARGUMENT AND THE PROBLEM OF QUALIA

Ahmet Kadir Uslu

APPROVED:

Josh Weisberg, Ph.D.
Committee Chair

James W. Garson, Ph.D.

David K. Phillips, Ph.D.

Steven G. Craig, Ph.D.
Interim Dean, College of Liberal Arts and Social Sciences
Department of Economics

THE KNOWLEDGE ARGUMENT AND THE PROBLEM OF QUALIA

An Abstract of a Thesis

Presented to

The Faculty of the Department

of Philosophy

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

Ahmet Kadir Uslu

May, 2016

ABSTRACT

In the history of philosophy, the mind-body problem is a traditional problem that examines the relation between mind and body. It simply seeks an answer for the question: what is the relation between mental properties and physical properties? The modern version of the mind-body problem is called the problem of consciousness. Accordingly, philosophers try to explain how and why we have conscious states or phenomenal experiences rather than nonconscious states. In other words, why is there ‘something it is like’ for a person in conscious experience? Conscious experience seems the hardest problem of consciousness because it is subjective. The subjective aspect of experience is considered as a big threat against physicalism since there is still no sufficient explanation about how it arises from a physical basis. ‘Phenomenal experience’ and ‘qualia’ are the other terms that refer to conscious experience. Proponents of qualia claim that no physical theory of mind can explain the qualitative character of subjective experience because qualia are not reducible to the physical properties of the mind. On the other hand, physicalists argue that mental states are brain states and brain states are physical states. Therefore, there is no reason to conclude that qualia lie beyond the scope of physicalist theory of mind. Frank Jackson’s knowledge argument is one of the main objections against physicalism. The main aim of this text is to argue whether ‘the knowledge argument’ really threatens physicalism. In this context, I first discuss what the term ‘qualia’ means and how it became the essential part of the consciousness debates. Then I examine the knowledge argument in detail. In the second chapter, I argue two main objections to the knowledge argument, respectively Daniel Dennett’s radical objection and the ability hypothesis. However, I show that both of them have some flaws. Lastly, I assert that the old fact-new mode analysis is the best possible answer to the knowledge argument.

Acknowledgements

Many thanks and appreciation to the people in the department of philosophy who have supported me during my academic education in the University of Houston.

I would like to thank my advisor Josh Weisberg and my committee members James W. Garson and David K. Phillips for believing in me. Their encouragements were really important to complete this thesis.

I also want to thank my undergraduate advisor Nebil Reyhani and the other academicians who guided me about my academic career.

I am very grateful to the Ministry of National Education (Republic of Turkey) for their financial support.

Lastly, I want to thank my family and my girlfriend for their endless support.

Table of Contents

Abstract.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
CHAPTER 1.....	1
Introduction.....	1
Qualia and Consciousness.....	3
The Knowledge Argument.....	6
Epiphenomenal Qualia.....	10
CHAPTER 2.....	14
Daniel Dennett’s Radical Objections.....	14
The Ability Hypothesis.....	22
CHAPTER 3.....	31
The Old Fact/New Mode Thesis.....	31
Bibliography.....	46

CHAPTER 1

1.1.Introduction

The mind-body problem is one of the fundamental problems in the history of philosophy. It investigates the relation between mind and body. Accordingly, philosophers have been seeking an answer for ages for the question: what is the relation between mental properties and physical properties? While physical properties are supposed to be objective and observable by anyone, mental properties are considered subjective. For example, anyone can see that I have two hands. This is one of the physical properties of my body. Then, it is equally accessible to every person. On the other hand, I may tell you that I am in pain. Or, I may tell you that I am thinking about what I will eat tonight. Unlike physical properties, these conscious events are private to me and only I have privileged access to them. In other words, only I can feel directly whether I am in pain or not. The others can only surmise from my behaviors that I am in pain. This is what makes the problem intractable.

The modern version of the mind-body problem is the problem of consciousness, or with David Chalmers' (1995) term 'the hard problem of consciousness'. In *Facing Up to the Problem of Consciousness*, Chalmers states:

It is undeniable that some organisms are subjects of experience. But the question of how it is that these systems are subjects of experience is perplexing. Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience

arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does (1995, p. 203).

In a word, the hard problem of consciousness is a problem which tries to explain how and why we have conscious states or phenomenal experiences rather than nonconscious states, or why there is 'something it is like' for a person in conscious experience. So the hardest problem of consciousness is the problem of experience because it is subjective. As Thomas Nagel (1974) has introduced it, there is something it is like to be a conscious organism. This subjective aspect of the experience is seen as a big threat for physicalism because we have no good explanation of how it arises from a physical basis. Sometimes, 'phenomenal experience' and 'qualia' are used instead of conscious experience.

In the first chapter of my thesis, I will first discuss what the term 'qualia' means and how it became the central part of the consciousness debates. Second, I will argue how it relates to our main topic, 'the knowledge argument'. Then I will give more details about Jackson's knowledge argument and his 'Mary case'. In the second chapter, I will argue two important physicalist objections to the knowledge argument, respectively Daniel Dennett's radical objections and the ability hypothesis; and I will suggest that they are not the best objections against the knowledge argument since both of them have some flaws. In the last chapter, I will claim that the old fact-new mode thesis is the best possible physicalist answer to Mary argument. I will support my argument with Torin Alter's conceptual mastery thesis. Lastly, I will conclude that qualia do not include anything non-physical like Jackson claims. Namely, the knowledge argument is not sufficient to threaten physicalism.

1.2. Qualia and Consciousness

As I mentioned, the problem of conscious experience is sometimes called as the problem of qualia. It is considered equivalent to the problem of consciousness by some philosophers. For example, John Searle posits that “[t]he problem of consciousness is identical with the problem of qualia...” (1998, p. 28). The importance of qualia comes largely from the fact that it is thought a fundamental problem for materialist explanations (physicalism, functionalism and so forth...) of the mind-body problem. Proponents of qualia claim that no physical theory of mind can explain the qualitative character of subjective experience because qualia are irreducible and non-physical properties of the mind. On the other hand, the opponents of qualia posit that mental states are brain states and brain states are physical states; and they have found no reason for concluding that qualia lie beyond the scope of physicalist theory of mind.

Then, what do qualia mean? To have a clear understanding of the notion of qualia, we need to look at the origins of the notion. The word ‘qualia’ (singular form: quale) derives from Latin, and it means qualities in Latin. In its modern sense, the term was first used by the philosopher Charles Sanders Peirce. Tim Crane gives us more information about it:

The first philosopher to use terms ‘quale’ and ‘qualia’ in something like its modern sense was C.S. Peirce. When Peirce wrote in 1866¹ that ‘there is a distinctive *quale* to every combination of sensation ... a peculiar *quale* to every day and every

¹ Brian L. Keeley (2009) claims that there is some confusion in Crane’s citation of the 1866 date here. According to him, the quoted passage is from Peirce’s collected papers and his notes for a series of lectures in 1898.

week—a peculiar *quale* to my whole personal consciousness’, he was talking about what experience is like, in a general sense, not restricted to the qualia of experience in the sense in which it is normally meant today (2001, p. 177).

However, Clarence Irving Lewis was the first person who introduced the technical use of the term ‘qualia’ into the philosophical discussion. In his book *Mind and the World Order*, he (1929) offers that conscious experiences present the immediate data (the given) to the mind. ‘The given’ here can be interpreted as the antecedent of Nagel’s phrase ‘what it is like to’ since both of them refer to qualia. In other words, Lewis’ term ‘the given’ can be used for expressions such as ‘looks like’, ‘sounds like’, ‘tastes like’, etc. In Lewis’ words, what it is that is given:

In any presentation, this content is either a specific quale (such as the immediacy of redness or loudness) or something analyzable into a complex of such. The presentation as an event is, of course, unique, but the qualia which make it up are not. They are recognizable from one to another experience (1929, p. 60).

Then, we can define ‘qualia’ as properties of what are given.

Also, Crane (2001) asserts that the origin of the term of qualia has relation to the notion of sense-data. According to him, they play similar roles. While sense-data are objects of experience, qualia are their properties. Then, he continues:

Sense-data were supposed to be the immediate objects of experience. So conceived, our awareness of sense-data are among those facts about experience which are open to philosophical reflection, rather than scientific theorizing. One would expect, then, that awareness of sense-data is something that can be gleaned from thinking about ‘what it is like’ to have an experience, in Nagel’s phrase. Since facts about

what it is like to have experience ought, on the face of it, be obvious to us, or obvious on reflection, then the existence of sense-data ought to be obvious to us. (2001, p. 174).

In this case, the puzzle of whether there are sense-data in our experiences is parallel to the puzzle about qualia. Price's (1932) example of seeing a tomato is a good example to see this parallelism. In *perception* he says:

When I see a tomato there is much that I can doubt. I can doubt whether it is a tomato I am seeing, and not a cleverly painted piece of wax. I can doubt whether there is a material thing there at all... One thing however I cannot doubt: that there exists a red patch of a round and somewhat bulgy shape, standing out from a background of other color-patches, naad having a certain visual depth, and that this whole field of color is presented to my consciousness... that something is red and round then and there I cannot doubt... that it now *exists* and that *I* am conscious of it—by me at least who am conscious of it this cannot possibly be doubted... This peculiar and ultimate manner of being present to consciousness is called *being given*, and that which is thus present is called a *datum* (1932, p. 3).

What Price tries to say here that maybe we can doubt there is a tomato I am seeing as a material object. However, we cannot doubt that I am seeing something red and round. So it is obvious that I am conscious of something. This conscious experience is what we call qualia or phenomenal experience. Then as Crane claims, we can talk about a relation between sense-data and qualia. Also, it is worth noting that Price's tomato example can be the source of inspiration for Jackson's Mary argument, namely the knowledge argument.

1.3. The Knowledge Argument

The knowledge argument is one of the main arguments against physicalism, which claims that conscious experience involves non-physical properties. According to this, there are phenomenal truths about consciousness that we cannot deduce them from the complete physical truth. Frank Jackson (1982) introduced the knowledge argument by his famous example of neurophysiologist Mary.

He (1982, 1986) tells us Mary is a clever scientist who has grown up in a black and white room and been forced to learn everything there is to know about the physical nature of the world through a black and white television monitor. That is to say; she has no color experiences although she learns all the physical facts about seeing in color. Then, her captors release Mary from her black and white room or give her a color television. What will happen after Mary is released and sees a red rose for example? Will she learn anything new or not? According to Jackson, it seems obvious that she will learn what it is like to see red. It means that her previous knowledge was incomplete though she had all the physical information. Therefore, “there is more to have than that, and physicalism is false” (Jackson, 1982, p. 130). This is a brief summary of Jackson’s knowledge argument. He uses this example to establish the main epistemic premise of the knowledge argument, ‘*the non-deducibility claim*’ which states that: “there are phenomenal truths that cannot be a priori deduced from the complete physical truth” (Alter, 2008, p. 4). Here is the original Mary argument in Jackson’s words:

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 wave-length 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” (1982, p. 130).

In his later work (1986), Jackson emphasizes that Mary actually becomes physically and scientifically omniscient. He says:

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 (1986, p. 291).

On this basis, Jackson claims that physicalism is false:

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 (1982, p. 130).

Then, Jackson’s original Mary argument can be outlined as following:

1) Mary knows all the physical facts about color vision before she leaves the room

- 2) Mary learns something about what it is like to see in color (new fact about color experiences) when she leaves the room
- 3) Then, there are non-physical facts concerning color experiences
- 4) Therefore, physicalism is false

By the way, many authors (included me) identify the knowledge argument with Mary case. However, Jackson (1982) uses further example in his paper: the case of a person, Fred, who has better color vision than anyone else on record.²

Jackson's argument seems *prima facie* impressive, but it is also highly controversial. Philosophers have advanced many different responses to the knowledge argument since Jackson introduced Mary case. For example, George Graham and Terence Horgan (2000) state that: many philosophers have claimed that the argument is inadequate as it stands and that Jackson's thought experiment of Mary is "poorly conceived" (Thompson, 1995, p. 264).

On the other hand, some dualistic attempts have arisen to refute physicalist accounts of the qualia and knowledge argument. As Jesse Prinz said: "responding Jackson is like killing a worm with a big rock: it's easy enough to kill the initial worm, but when you lift the rock, you expose a writhing tangle of new worms. That's what makes the knowledge argument so good" (2005, p. 1).

In addition to Jackson, the major authors who defend a dualistic (or epiphenomenalist) account of qualia are Nagel and Chalmers. They simply claim that qualia are non-physical properties of conscious experience. In his famous paper *What is it like to be a bat*, Nagel argued that "consciousness is what makes the mind-body problem

² For more information, see Jackson's (1982) *Epiphenomenal Qualia*, p. 128.

really intractable” (1974, p. 435). By the term ‘consciousness’, Nagel means subjective aspect of conscious experience. He says:

Conscious experience is a widespread phenomenon. It occurs at many levels of animal life ... the fact that an organism has conscious experience *at all* means, basically, that there is something it is like to *be* that organism ... Fundamentally an organism has conscious mental states if and only if there is something that it is like to *be* that organism—something it is like *for* the organism (1974, p. 436).

Next, Nagel concludes that if ‘qualia’ are irreducible to any physicalist explanation, then physicalism is false. Because no physical theory is sufficient to explain the subjective character of consciousness.

David Chalmers, another supporter of the knowledge argument formulates the philosophical zombie argument to show that consciousness cannot be reduced only to physical properties. According to him, “consciousness escapes the net of reductive explanation, because no explanation given wholly in physical terms can ever account for the emergence of conscious experience” (1996, p. 83). Chalmers’ zombie argument is as follows:

This creature is molecule-for molecule identical to me, and indeed identical in all the low-level properties postulated by a completed physics, but he lacks conscious experience entirely ... what is going on in my zombie twin? He is physically identical to me, and we may as well suppose that he is embedded in an identical environment. He will certainly be identical to me functionally: he will be processing the same sort of information, reacting in a similar way to inputs, with his internal

configurations being modified appropriately and with indistinguishable behavior resulting (1996, p. 83).

Chalmers thinks that such a zombie world is conceivable. If zombies are conceivable, they are also possible. If they are possible, then physicalism is false. Because they lack qualia although they are molecule-for-molecule identical to humans. Then, qualia are beyond physical explanations.³

I shortly discussed Nagel's and Chalmers' arguments since they are quite related to our main discussion topic. As I mentioned in the footnote, Chalmers' zombie argument depends on a very different structure, but all these views aim at the same thing with Jackson's knowledge argument: they simply claim that qualia are not reducible to physical properties. After giving this information, let's return to our main topic; the knowledge argument. The chief claim of the Mary case was like this: after her release from the black and white room, Mary learns a *new* fact - *what it is like to see red* though her complete physical knowledge. Therefore, there is some non-physical information about colors that Mary did not know before. Then physicalism is false.

1.3. Epiphenomenal Qualia

The notion of 'epiphenomenal qualia' is another important aspect of Jackson's account. In his Mary argument, Jackson intends to demonstrate that qualia are epiphenomena, and they are causally ineffective in the physical world. He considers qualia as just products of brain activities, and thus they make no difference to the physical world. Therefore, Jackson says that physicalism leaves qualia out of the scientific explanation

³ Although Chalmer's zombie argument aims at showing that qualia are not physical like Mary argument, it is structurally different. Unlike Nagel's and Jackson's arguments, it relies on conceivability and modal claims.

because qualia have no causal power in the physical world. In Jackson's own words, they are "inefficacious" ... and "their possession or absence makes no difference to the physical world" (1982, p. 133). On the other hand, some suggest that qualia can be causally efficacious in the physical world. For example, we take some medicine when we have a headache or stomachache. In this case, pain forces me to take a pill. However, Jackson replies that none of these has any real force. He says: "it is simply a consequence of the fact that certain happenings in the brain cause both" (1982, p. 133). As further evidence for the inefficacy of qualia, Jackson shows Darwin's theory of evolution. According to natural selection, qualia evolved over time for helping us to survive in the physical world. We expect qualia to be conducive to survive because we have them while the earliest organisms do not. But Jackson states that the effect of qualia is hardly seen in the physical world. He gives polar bears as a counterexample: "Polar bears have particularly thick, warm coats. The Theory of Evolution explains this (we suppose) by pointing out that having a thick, warm coat is conducive to survival in the Arctic. But having a thick coat goes along with having a heavy coat, and having a heavy coat is *not* conducive to survival. It slows the animal down" (1982, p. 134). At the same time, Jackson emphasizes that it does not mean that Darwin's theory is wrong. He concludes by saying that:

The point is that all we can extract from Darwin's theory is that we should expect any evolved characteristic to be *either* conducive to survival *or* a by-product of one that is so conducive. The epiphenomenalist holds that qualia fall into the latter category. They are a by-product of certain brain processes that are highly conducive to survival (1982, p. 134).

As a third objection, Jackson discusses the problem of other minds. We know about other minds because of their behaviors. For example, I know that you are sleepy since I sometimes become sleepy, and that's why I know what it is being sleepy. According to this, some would ask that how we come to know about other minds if we do not suppose their behaviors are the outcome of the qualia. Their behaviors provide evidence to believe that they have qualia like mine. So, we believe that their qualia cause their behaviors. However, Jackson opposes by saying that "an epiphenomenalist cannot regard behavior, or indeed anything physical, as an outcome of qualia" (1982, p. 134). Then he gives an example:

Consider my reading in *The Times* that Spurs won. This provides excellent evidence that *The Telegraph* has also reported that Spurs won, despite the fact that (I trust) *The Telegraph* does not get the results from *The Times*. They each send their own reporters to the game. *The Telegraph's* report is in no sense an outcome of *The Times'*, but the latter provides good evidence for the former nevertheless (1982, p. 134).

In a word, Jackson states that qualia are effects that cannot be causes of something physical, but they are caused by physical things. He concludes by saying that:

They (qualia) are an excrescence. They *do* nothing, they *explain* nothing, they serve merely to soothe the intuitions of dualists, and it is left a total mystery how they fit into the world view of science. In short we do not and cannot understand the how and why of them (1982, p. 135).

Jackson endorses the physicalistic view that we are a part of nature. He says, "we are the products of evolution" (1982, p. 135). When we need, we use our understanding

and sensation to survive. However, he adds, epiphenomenal qualia are excluded from this survival.

To sum up, Jackson thinks that his argument leads to epiphenomenalism. For this reason, physicalism is a highly optimistic view to explain qualia. Even if we are a part of nature like physicalists claim, qualia do not fit into the physical explanations. Therefore, before her release, Mary's physical information is not enough to know *what it is like to see red*.

In conclusion, the knowledge argument seems a powerful threat against physicalism. However, it is worth noting that Jackson has modified his views about Mary. He now agrees "with Laurence Nemirow and David Lewis on what happens to Mary on her release" (2004, p. 439). But he rejects the idea that Mary only comes to know an old fact in a new way. He replies against the old fact-new mode thesis by applying to the new fact thesis (see Jackson, 2005, p. 318).

In the next chapter, I will discuss two important rebuttals to Jackson's knowledge argument and his views about qualia. First, I will argue one of the most radical objections, namely Dennett's radical attempt. Then, I will examine *the ability hypothesis* that is very popular response to the knowledge argument.

CHAPTER 2

2.1. Daniel Dennett's Radical Objections

When Jackson introduced the knowledge argument, it evoked many philosophical objections. Still, there is no philosophical consensus about Mary case. While some philosophers do not accept one premise of the knowledge argument, others do not agree with its conclusion. Dennett rejects the second premise of Jackson's argument, namely the claim that Mary learns a new fact about colors when she leaves the room. Now, let's see Dennett's radical objections to this premise.

Unlike many other critics, Dennett (1991) thinks that Jackson's thought experiment is ill-founded, and an intuition pump encourages us to misunderstand its premises. According to Dennett, before Mary leaves the room, she is already capable of knowing the color experience she is about to have. Thus, it will not be surprising for her to see a red apple. She will be able to infer from her nervous system that the apple she is seeing is red. Dennett supports his claim by his blue banana example:

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 (...) would make on my nervous system. So I already knew exactly what thoughts I would have

(...). I was not in the slightest surprised by my experience of blue (...). 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!" (1991, pp. 399-400).

So, what is Dennett claiming here? As Michael Beaton (2005) points out, he is trying to say that Mary can know what it is like to see in color without ever having seen anything colored. Then, Dennett seems to refuse Jackson's main claim which suggests that "it seems just obvious that she will learn something about the world and our visual experience of it" (Jackson, 1982, p. 130). He thinks that this is what is wrong with Mary argument. However, some argue that Dennett's aim here is more modest and he does not use his blue banana example to deny that Mary learns anything upon her release. For example, Luca Malatesti (2008) points out that Dennett just tries to disqualify the intuition, which supports the belief that she might learn something upon her release. There is a good reason for thinking in this way because in *Consciousness Explained* Dennett says: "My point is not that my way of telling the rest of the story proves that Mary *doesn't* learn anything, but that the usual way of imagining the story does not prove that she does. It doesn't prove anything [...]" (1991, p. 400). Then, Dennett does not seem to argue that Mary cannot undergo an epistemic progress by seeing colored objects.

Next, in his paper *What RoboMary Knows*, Dennett (2006) presents further objection for the knowledge argument. He claims that the knowledge argument depends on the experience requirement. According to this, someone could not know what it is like to see in color if she never seen in color. As Torin Alter (2008) points out, Dennett thinks

that this is the basis of the knowledge argument's main epistemic premise: the premise that any physical knowledge is not enough for phenomenal knowledge of color experiences. Alter summarizes the reason that why Dennett believes the knowledge argument depends on the experience requirement as follows:

The no-experience-necessary response. The claim that Mary makes epistemic progress upon release would make perfect sense if having color experiences were required for knowing what it's like to have them. But if the experience requirement fails—if it is possible to know what it's like to see in color without having color experiences—then why couldn't Mary put herself in a state that allows her to figure out what it's like to see in color? If there is no logical bar to obtaining this phenomenal knowledge without seeing colors, then there is no reason why Mary could not obtain that knowledge by exploiting her comprehensive physical knowledge (2008, pp. 5-6).

Then, if it is possible to obtain any phenomenal knowledge without having experience, there will be no logical obstacle for Mary to know what it is like to see red without seeing red. Dennett tries to undermine the knowledge argument by removing the experience requirement. He thinks that it is not necessary to have an experience with X to know what it is like to have an experience with phenomenal character X (Dennett, 2006). As an example, he gives Hume's missing shade of blue. According to this example, one can extrapolate the missing shade of blue by experiencing other shades which are phenomenally similar. Namely, we do not experience the missing shade directly; we experience the surrounding shades which are fairly similar to the missing one. That is to say, Dennett removes the experience requirement as opposed to Jackson.

Now, let's look at Dennett's RoboMary argument. Dennett intended to undermine the belief that Mary gains knowledge when she leaves the room through this argument. According to Dennett's argument, RoboMary is a standard Mark 19 robot, but she has one important difference, she was brought online with black and white video cameras unlike standard Mark 19 robots. So, she does not have color vision. However, everything else in her hardware is arranged for color vision, which is same with in the standard Mark 19. Like human Mary, RoboMary is also able to learn all physical facts through her black and white camera eyes. And she learns everything concerning the color vision of Mark 19s while she was waiting for her new pair of color cameras to change. Eventually, she becomes omnipotent about the color-coding system of all Mark 19s. Through using her comprehensive knowledge, RoboMary puts herself in standard Mark 19s' situation and writes some code which enables her to colorize the input from her black and white cameras in light of data she collected. During her research and development stage, RoboMary makes comparisons between herself and other Mark 19s about how different they react while looking at the same objects; then she records all the information she gathers. Finally, her black and white camera eyes are replaced with a pair of color cameras. And then she opens her eyes and see a ripe tomato for example. But she learns nothing new because she already knew what it would be like for her to see a ripe tomato through the information she gathered from other Mark 19s (Dennett, 2006). Dennett gives a better explanation of this situation in his locked RoboMary case:

She obtains a ripe tomato and plunks it down in front of her black and white cameras, obtaining some middling gray scale values, which lead her into a variety of sequel states... She consults an encyclopedia about the normal color range of

tomatoes, and she knows that these gray-scales in these lightning conditions are consistent with redness, but of course nothing comes to her directly about color, since she has black and white cameras, and moreover, she can't use her book-learning to adjust these values, since her color system is locked. So, as advertised, she can't put herself directly into the *red-tomato-experiencing* state. She looks at the (gray-appearing) tomato and reacts however she does, resulting in, say, thousands of temporary settings of her cognitive machinery. Call that voluminous state of her total response to the locked gray-tomato-viewing state A... Then she compares state A with the state that her model of herself goes into... (namely) state B, the state she would have gone into if her color system hadn't been locked. RoboMary notes all the differences between state A...and state B...and... makes all the necessary adjustments and puts herself into state B. State B is, by definition, not an illicit state of color experience; it is the state that such an illicit state of color experience normally causes (in a being just exactly like her). But now she can know just what it is like for her to see a red tomato, because she has managed to put herself into such a dispositional state... (2006, p. 11).

However, Dennett's argument is highly controversial. For instance, Alter argues that Dennett's argument does not threaten *the non-deducibility claim*, which states that "there are phenomenal truths that cannot be a priori deduced from the complete physical truth" (Alter, 2008, p. 4). According to Alter, if *the non-deducibility claim* were false, Mary would not learn any facts when she leaves the room. However, she learns some facts new when she is released, then *the non-deducibility claim* is true. Based on this view, Alter says that Locked RoboMary does not a priori deduce the phenomenology of seeing red from the

physical truth. She just puts herself in a dispositional state that a standard Mark 19 would have. So, she uses her physical knowledge to conceive the required effect, but the way she uses is not a priori deduction. Then, Alter (2008) posits that the only way to threaten *the non-deducibility claim* with RoboMary argument is a priori deduction. On the other side, Dennett gives a response to this claim, and he says that: “I just do not see that this is what matters. So far as I can see, this objection presupposes an improbable and extravagant distinction between (pure?) deduction and other varieties of knowledgeable self-enlightenment” (2006, p. 12). Next, he says that “I didn’t describe RoboMary as “self-programming” herself; I said she “notes all the differences between state A, the state she was thrown into by her locked color system, and state B; the state she would have been thrown into had her color system not been locked, and-being such a clever, indefatigable and nearly omniscient being-makes all the necessary adjustments and *puts herself into state B*” (2006, p.12). According to Alter, Dennett means that “the distinction between a priori deducibility and other sorts of infer-ability is (a) deeply problematic or (b) not relevant to whether physicalism stands or falls” (Alter, 2008, pp. 8-9). That is to say, Dennett points out that this kind of distinction is rather inaccurate, and it is not related with physicalism. But Alter does not agree with Dennett, and he tries to illustrate his idea with an example. He makes a comparison between two cases to know that the sum of a trapezoid’s angles is 360 degrees:

Case 1: You figure out the sum by constructing a proof from Euclid’s axioms.

Case 2: A future neuroscientist time travels back to the present and describes a brain state characteristic of someone who knows the sum. She also gives you a device that can be used to put you in that state and explains that the device works only if

you contemplate Euclid's axioms for a few seconds. You contemplate the axioms and use the device. It works (2008, p. 9).

According to Alter (2008), in case 1 the geometrical information is a priori deduced from Euclid's axioms. On the other hand, in case 2 you do not a priori deduce the information although you use the same knowledge.

In order to show how we apply the same distinction to phenomenal knowledge, Alter also gives the example of Hume's missing shade of blue. He says that although there are several ways to get the phenomenal knowledge of the missing shade through phenomenal knowledge of the surrounding shades, only one fulfills a priori deduction: "deducing the missing-shade phenomenology by combining phenomenal information about the surrounding shades, without relying on other phenomenal information" (2008, p 10). For Alter, this kind of a priori deduction is feasible. However, he claims that it is not possible for Mary to deduce what it is like to see red by combining the information she obtains before leaving the room. This is also valid for RoboMary case. Her reasoning includes more than a priori deduction even if she understands what it is like to see red. Namely, RoboMary's reasoning is similar to the case 2 than case 1 which are mentioned above (Alter, 2008). He states his argument as follows:

Why does putting herself in state B enable Robomary to know what it's like to see red? B is a dispositional and (let us assume) non-phenomenal state; there is nothing it is like to be in B. Nevertheless, B involves color phenomenology in that it contains the relevant phenomenal information. Therein lies the problem for Dennett's argument. By putting herself in a state that involves color phenomenology, RoboMary cheats. Pre-release Mary should be no less puzzled

about B than she is about seeing red. If she lacks phenomenal information about seeing red, then she lacks the phenomenal information that B contains. If there are open epistemic possibilities about the nature of phenomenal redness that she cannot eliminate, then there are open epistemic possibilities about the content of B that she cannot eliminate. RoboMary comes by her phenomenal knowledge of color experience not by a priori deduction from physical information but rather by putting herself in a non-phenomenal dispositional state that contains the relevant phenomenal information (Alter, 2008, pp. 12-13).

To sum up, Alter posits that RoboMary cheats because if the states she puts herself would not involve any phenomenal information about redness, it would be impossible for her to imagine what it is like to see red. So, in order to threaten the knowledge argument, she needs to deduce without having similar phenomenal information.

On balance, Alter may be right about Dennett's arguments. The examples Dennett gives do not seem to pass *the non-deducibility claim*. In this respect, I agree with Alter that RoboMary cheats because she uses abilities or phenomenal information that are not allowed in the knowledge argument, and what RoboMary does is not a priori deduction. However, what I oppose about Alter's claim is this: *the non-deducibility* does not prove that there is something non-physical in Mary case. She is incapable to deduce what it is like to see red before leaving the room because she did not undergo the relevant experience yet. After seeing red for the first time, she just makes epistemic progress about the fact she already knew. She only learns another aspect of the relevant fact. So, as distinct from Dennett I confess that Mary makes epistemic progress upon her release. But it does not mean that what Mary learns after her release is non-physical. In a word, I disagree with

Alter that physicalism is not sufficient to explain Mary argument. Mary makes epistemic progress since she comes to know an old fact she already knew in a *new* way. She has complete physical knowledge about colors before her release, but still she sees red for the first time. Because a physical fact can be gained by countless different ways. This is what makes Mary case interesting. I will argue this view in depth in the last chapter. First, let's discuss the ability hypothesis.

2.2. The Ability Hypothesis

The main claim of the proponents of the knowledge argument is that Mary learns a new fact when she leaves the room. And some critics of the new fact thesis accept that when Mary is released, she learns what it is like to see color. However, these critics suggest that she does not learn any new facts. Laurence Nemirow (1980, 1990) and David Lewis (1983, 1988) introduced one version of this view, which is called the ability hypothesis. This was the first major response to the knowledge argument. Nemirow and Lewis argue that knowing what an experience is like is knowing-how, not knowing-that.⁴ That is to say, knowing what an experience is like is not factual knowledge. Therefore, the fact that Mary learns what color experience is like when she is released does not threaten physicalism.

According to the ability hypothesis, what Mary does learn is just a bundle of new abilities, remembering, imagining, and recognizing. In the postscript to *Mad Pain and Martian Pain*, Lewis comments: "...knowing what it is like is not the possession of information at all. It isn't the elimination of any hitherto open possibilities. Rather, knowing what it is like is the possession of abilities: abilities to recognize, abilities to

⁴ See Gilbert Ryle's (1946, 1949) distinction between knowing-how and knowing-that. Also for further information, see Jason Stanley and Timothy Williamson's "Knowing How" (2001).

imagine, abilities to predict one's behavior by imaginative experiments" (1983, p. 131). Similarly, in *What Experience Teaches*, he states: "the ability hypothesis says that knowing what an experience is like just is the possession of these abilities to remember, imagine, and recognize. ... It isn't knowing-that. It's knowing-how" (1990, p. 516). Nemirow also presents a fairly similar view. He says: "knowing what an experience is like is the same as knowing how to imagine having the experience" (1990, p. 495). Accordingly, we can say that knowing what it is like to experience A is having the ability to imagine or recognize or remember having experience A.

However, is the ability hypothesis persuasive? Does it show that Mary case does not threaten physicalism? While some claim that the ability hypothesis is true, many philosophers argue that the ability hypothesis is not a strong answer against the knowledge argument. First, let's discuss the ability hypothesis in detail. Afterward, I will argue some important objections to it.

Lewis describes knowing what an experience is like with specific abilities. Remembering is one of these abilities. For example, suppose you do bungee jump for the first time. After having this experience, you learn what it is like to do bungee jump. Hereafter, you can remember the experience. Since you remember it, you can imagine what it is like to do bungee jump. Because, if you are able to remember an experience, you must be able to imagine it. Then, by having the experience of doing bungee jump, you acquire the abilities to remember and imagine the experience. Similarly, Lewis (1988/2004) compares knowing 'what it is like' to being able to use chopsticks.

In addition to the abilities to remember and imagine, the third important ability you gain is the ability to recognize the experience if it comes again. Suppose you tasted

Vegemite before. Lewis says: “If you taste Vegemite on another day, you will probably know that you have met the taste once before. And if, while tasting Vegemite, you know that it is Vegemite you are tasting, then you will be able to put the name to the experience if you have it again” (1988/2004, p. 98). David Hugh Mellor (1993), who is another proponent of the ability hypothesis purports that knowing what experiences of a certain kind are like requires at least the ability of recognizing. For instance, in order to know what sugar tastes like, I must be able to recognize its taste. According to Mellor, it does not mean that this is just recognizing that what I am tasting is sugar. Even if I taste sugar without knowing what it is, I will recognize the taste of it. Meanwhile, although Mellor defends the ability hypothesis, he lacks the physicalist motivation that Lewis and Nemirow have.

To sum up, Lewis and Nemirow, and their followers try to show that learning what it is like is reducible to gaining abilities. However, to save physicalism from the threat of the knowledge argument, they must demonstrate that Mary’s new knowledge does not involve new factual knowledge. In other words, the acquisition of Mary should be fact-free in the relevant sense. As Sam Coleman (2009) remarks, Lewis’ account is nuanced on this issue. Lewis concedes that some “aspects of ability are, purely and simply, a matter of information. If you want to know how to open the combination lock on the bank vault, information is all you need. It remains that there are aspects of ability that do *not* consist simply of possession of information and that we *do* call knowledge. The ability hypothesis holds that knowing what an experience is like is that sort of knowledge” (1988/2004, p. 100).

According to Coleman (2009), the essential point here is that the proponents of the ability hypothesis must show that Mary does not add new factual knowledge to her stock.

So, the important question is whether Mary increases her factual knowledge by gaining the abilities with regard to the experience of redness. If the ability hypothesisists answer this question in the negative, is there any support for this claim?

The ability hypothesisists love to use analogies to support their claims. For example, Nemirow states that knowing what it is like is similar to being able to wiggle one's ear. Lewis compares knowing what it is like to being able to use chopsticks. Lastly, Mellor gives the example of riding bike to explain knowing what it is like. He says:

Knowing what feeling warm is like is not knowing any fact, because it is not knowing *that* any proposition is true: it is just knowing *how* to imagine feeling warm. In this respect it is like knowing how to ride a bicycle. I cannot state the fact I know then either, because there is no such fact to state. I must of course know some facts about bicycles to know how to ride one, but having this ability is obviously neither constituted nor entailed by my knowing those facts. And that is why no one thinks it mysterious that I cannot say what fact my knowing how to ride a bicycle is knowledge of: it is too obvious that there is no such fact (1993, p. 8).

What Mellor intended to say here is that knowing how to ride a bicycle does not depend on knowing any propositional knowledge. By refusing this kind of fact, he implicitly answers the question whether Mary increases her factual knowledge by gaining the ability about the experience of redness in the negative. But is Mellor's example cogent to threaten the knowledge argument? According to some, it is not.

As Coleman (2009) remarks, if a person rides a bike for the first time, we normally think that s/he would fall off. And in parallel with their arguments, Lewis, Nemirow, and

Mellor would be happy to explain this result by claiming that what that person needs is the ability to ride a bike, and the only way to gain that ability is practicing. So, academic study has no effect to acquire that kind of abilities. However, Coleman asserts that this explanation has a stipulative whiff. He says:

The first-time bike rider falls off, of course, because they lose their balance. Balancing is essential to riding, and mastery of it can only come through practice. Saying this is not to say anything that yet counts against the Hypothesisists. They can put the point by saying the ability to keep one's balance is a vital *component* of the ability to ride a bike, if we may talk this way. But consider the kind of knowledge that is involved in learning how to keep your balance on a bike. I claim it depends upon *cognitively monitoring the quality of new phenomenal states* (2009, p. 83).

What Coleman tries to say is that having the ability to ride a bike requires new phenomenal state: how it feels to be balanced on a bicycle. And this is what the ability hypothesisists reject. According to Coleman, keeping your balance when riding is crucial since it allows senses to concentrate to other important matters. Also, you can keep the bike straight up by looking to check whether the horizon is exactly horizontal. However, focusing on the horizon takes too much time, therefore you can crash into a kid or a cat. Accordingly, Coleman gives good bike riders as an example. He says that: "Good bike riders don't use what they see as a guide to whether they are balanced; they use their eyes to see what's coming and to find their way. What keeps them up instead is their sense of balance, how balanced they feel" (2009, p. 83). So the important point here is not only about a new feeling; instead, cognitive monitoring is essential part of Coleman's account. Coleman's objection is valid for Lewis' and Nemirow's examples too. To have the ability

of using chopstick for example, one will learn necessarily how it feels to use chopsticks correctly. Similarly, Coleman claims that Mary's experience also involves irreducible phenomenal knowledge. He states: "In each case where a new ability is acquired, specific new phenomenal states must be learned about" (2009, p. 88).

I agree with Coleman that the ability hypothesis is imperfect and knowing what it is like cannot be reducible to the possession of abilities. However, I do not agree with his conclusion that a new ability requires specific *new* phenomenal knowledge. One can gain a new ability about an old phenomenal fact. That is, a new ability does not necessitate a new phenomenal knowledge.

Another and I think stronger critique to the ability analysis is Alter's objection. According to ability hypothesis, having the ability to imagine an experience is a necessary condition of knowing what that experience is like, and Alter (1998) states that this seems implausible. He gives an example to refute that claim:

Suppose John is a non-defective human being except for one thing: he lacks the ability to imagine seeing colors. While staring at a red tomato, he has the typical human experience of seeing red. But if we take away the tomato, he is unable to conjure up a red visual image. His impoverished imaginative capacity does not seem to prevent him from knowing what seeing red is like *while he stares at the tomato* (1998, p. 37).

If the example Alter gives is conceivable, then the ability to imagine seeing red would not be necessary to know what seeing red is like, and this threatens the ability hypothesis. Similarly, a person who has Alzheimer disease can be given as counterargument to the ability hypothesis. Suppose that you have Alzheimer and you forget

what you saw in ten minutes. Then, for example, I show you a picture of your father, and you do not remember it after ten minutes. But it does not mean that you do not know what seeing your father's picture is like while you are looking at the picture. However, the ability hypothesis claims that the ability to remember an experience is one of the necessary conditions to know what it is like to experience something. If my example is plausible, there is something wrong with the ability hypothesis.

Michael Tye (2000) reveals another difficulty for the ability hypothesis. He takes the color experience to show that human sensory experience is highly rich. We know from some estimates that humans can experience about 10 million different colors. But only a few of these colors have unique names since we do not have enough stored representations in memory for the rest of colors. However, my phenomenal experience of blue 11, for example, is different from my experience of blue 13 although I do not have any stored representations in my memory, or any concepts for these specific hues. Therefore, it is hard for me to go into a paint store and identify a color on a chart which is exactly same hue with my living room walls. Certainly I have the concept of blue, and I can recognize it when I see something blue. But the problem is that I do not possess the concepts for every hues of blue. Of course, this point is not limited to colors. It applies to the other sensory experiences; sounds, tastes etc. In this respect, Tye criticizes Lewis and Nemirow. He says:

When Mary first sees the rose and exclaims, "So, that is what it is like to see red," she certainly acquires certain abilities, as Lewis and Nemirow suppose. She is now able to recognize red things by sight; she can identify the experience of red when it comes again; afterwards, she can remember the experience of red; she can imagine what it is for something to be red. So far no obvious difficulty. But she knows more

than just what it is like to experience red. As she stares at the rose, it is also true of her at that time that she knows what it is like to experience the particular determinate hue of red—call it ‘red17’—she is seeing. Of course, she does not know that hue *as* red17. Her conception of it is indexical. She thinks of it only as *that* shade of red. But she certainly knows what it is like to experience that particular hue *at that time at which she is experiencing it* (2000, p. 12).

In that case, does Mary acquire any new ability? She seems not able to recognize red17 as red17 because she lacks the concept red17 like other human beings. She does not have sufficient mental ability to identify the hue of red17 when she sees it again. For example, if someone presents her both red17 and red18, she cannot be certain which experience matches exactly with her previous experience of the red17. Then, Mary lacks the abilities Lewis mentions even though she certainly knows what it is like to experience the specific hue of red she is experiencing. To show this claim’s accuracy, Tye gives an example:

Suppose we inform Mary that she is seeing red17. She replies, “So, this is what it is like to see red17. I had always wondered. Seventeen, you see, is my favorite number; and red the color of my mother’s favorite dress.” We then say to her, “No, you don’t know what it is like to see red17. For you won’t remember it accurately, when you take your eyes from the rose; you won’t be able to recognize it when it comes again; you won’t be able to imagine the experience of seeing red17.” Should Mary then admit that she doesn’t really know what it is like to see red17 even while she is staring at the rose? She won’t know it later certainly. But it seems intuitively bizarre to deny that she knows it *at the time* (2000, pp. 12-13).

If I consider all the objections I have presented so far, I can conclude that the ability hypothesis expired, and it needs a revision. Or, we need an alternative response to the Mary case to save physicalism from the threat of the knowledge argument. In the next chapter, I will argue an encouraging hypothesis against to Jackson's Mary argument.

CHAPTER 3

3.1. The Old Fact/New Mode Thesis

I have presented two strong objections to the knowledge argument, respectively Dennett's radical objection and the ability hypothesis. However, I have argued that both objections seem to have some weaknesses that cannot be modified. Lastly, I will discuss another answer to the Jackson's Mary argument. It is called the old fact/new mode thesis. First, let's see what the old fact/new mode thesis is.

According to the old fact-new mode analysis, Mary gains knowledge after seeing red. However, she merely gets an old (previously known by her) fact in a new way. The old fact-new mode analysis was introduced by Terence Horgan as an objection to the knowledge argument. In *Jackson on Physical Information and Qualia*, he (1984) argues that Jackson's attack on physicalism is fallacious. First, he agrees with Jackson that Mary obtains new knowledge when she is released. Then, unlike Jackson's assertion, he states that Mary might just come to know an old fact she already knew by some new way of epistemic access. To understand Horgan's objection, let's first recall the formulation of the knowledge argument:

1. While still in a black and white room, Mary knows everything there is to know about the physical nature of the world and colors.
2. However, when she is released and sees the red tomato for the first time, she learns what it is like to see red.
3. Before Mary leaves the room, there was some knowledge that she lacked.

4. Therefore, there is a new knowledge about colors that Mary did not know while she was still in the room. Because this knowledge is not included in the physical information Mary knows, it is a non-physical fact.

Although Horgan accepts the first three premises, he demurs the conclusion of the last premise. According to him, the fact that Mary gains some knowledge after leaving the room, it might be an old fact she already knew. In other words, she might just come to know an old fact by a new way of epistemic access. He presents this argument in the following passages:

Consider Mary at the moment when she finally has her first color-experience – say, the experience of seeing ripe tomatoes. Jackson maintains, and I agree, that Mary obtains new knowledge at this moment, and thus new information: she finds out what it is like to see ripe tomatoes (1984, p. 150).

In the next page, he says:

Phenomenal redness, the referent of ‘this property’, may very well be a physical property. This possibility is not ruled out by the fact that Mary learns something new from her experience (1984, p. 151).

In a word, Horgan (1984) claims that premise four expresses new information since Mary has a new aspect on phenomenal redness. Still, this fact is compatible with physicalism. The perspective is new, but the property itself does not require to be new. He gives an example to explain his argument:

- (1) Superman can fly.
- (2) Clark Kent can fly (1984, p. 150).

According to Horgan (1984), if Lois Lane knows that Superman can fly but does not know that Clark Kent can fly, then (1) and (2) must refer to different information although both of them attribute the same property to the same person. Namely, it is quite possible that some sentences can express ontologically physical information but not explicitly physical information. He says:

Even though Superman is Clark Kent, nevertheless we must distinguish between the information that Superman can fly and the information that Clark Kent can fly. Similarly, even if phenomenal redness is a physical property, nevertheless we must distinguish between (i) the information that the given property, as physicalistically described, is possessed by ripe-tomato experiences, and (ii) the information which Mary expresses by saying “seeing ripe tomatoes has this property” (1984, pp. 151-152).

Similarly, consider the following example. Emma knows that Hulk Hogan is 63 years old. However, she does not know that his real name is Terry Gene Bollea. So she has never heard this name before. Then, she sees news on television which states that famous wrestler, Terry Gene Bollea is 63 years old. After seeing this news, she supposes that she has learned a new fact about a wrestler’s age. What she does not know is both ‘Hulk Hogan is 63 years old’ and ‘Terry Gene Bollea is 63 years old’ refer to the same fact. Accordingly, Emma already knew that fact before she saw the news on television.

So it can be applied to the Mary case. She might have two different concepts of a single physical property. In this case, when she says ‘seeing ripe tomatoes has this property’, she would express new knowledge, but it would not be concerning a new property she did not know before. It would be a new concept of an old fact.

Tye is another proponent of the old fact-new mode thesis. In his paper called *The Subjective Qualities of Experience*, Tye (1986) argues that Mary does not acquire knowledge of a new fact, but rather new introspective knowledge of an old fact she already knew. He says: “Knowing what it is like, according to (my account) is grounded upon factual knowledge which is obtained in the appropriate manner, namely by introspection” (1986, p. 9). He gives an example related to his claim. Suppose that Jones is a super scientist of the twenty-third century, who has acquired comprehensive knowledge of what goes on to us physically when we see colors. However, Jones is congenitally blind. The problem here for physicalism is supposedly that Jones does not know what it is like to undergo the characteristic experience of seeing red. Then one day, Jones accepts to have an operation. Now, he is able to undergo the characteristic experience of seeing red. According to Tye (1986), when he does undergo the experience and he identifies it via introspective awareness, then he will know what it is like to have the experience of seeing red. That is, he will learn something after the operation. But he will not come to know a new fact different from those he knew before. Rather he will come to undergo experience of a sort he has not undergone before. Accordingly, he will come to know a fact of an old sort in a new way.⁵

One major objection to the old fact-new mode analysis is “the new fact thesis” (Chalmers, 2004). According to this, “necessarily, whenever someone gains new knowledge of an old proposition they must also gain knowledge of some new proposition” (Cath, 2009, p. 149). If this is true, Mary must come to know some new fact when she

⁵ In his later works, Tye has modified his views about the knowledge argument, but I will not discuss them here.

comes to know an old fact under a new mode of presentation; and this seems to cause trouble for the old fact-new mode analysis. Suppose I do not know that water is H₂O. According to the new fact thesis, I would necessarily learn *new* and substantive fact about the world if I came to know that water is H₂O, and this is what the old fact-new mode thesis rejects. But this is a highly questionable claim. Because there are examples which shows that the new fact thesis is not obviously true. As Cath (2009) reminds, Lewis' (1979) example of two gods who occupy the same possible world is a good one. According to this, both gods are omniscient as regards every proposition true at their world; however, they still lack knowledge as:

[N]either one knows which of the two he is. They are not exactly alike. One lives atop the tallest mountain and throws down manna; the other lives atop the coldest mountain and throws down thunderbolts. Neither knows whether he lives on the coldest mountain or on the tallest; nor whether he throws manna or thunderbolts (Lewis, 1979, pp. 520-521).

Lewis claims that “if the gods came to know which was which, they would know more than they do. But they wouldn't know more propositions” (1979, p. 521). Then, if one suggests that the gods come to know old propositions under new modes of presentation, there would be no objection to such an account to indicate that it challenges the new fact thesis (Cath, 2009).

In fact, we can increase the examples. Suppose that Sadie is five-year-old, and she does not know to count up to 100. However, she knows four operations with small numbers. In this case, (1) can be false even though (2) is true:

(1) Sadie knows that $100-96=4$.

(2) Sadie knows that $2+2=4$.

According to this, if Sadie comes to know that $100-96=4$, then she will come to be in a new state of factual knowledge, but she will not come to know a new fact. She will only come to know an old fact under a new way of presentation. As a possible objection, some may claim that it is not true to compare mathematical truths with physical truths. But this is not the only example against the new fact thesis. Further examples can be given.

Then, we can easily suggest that the new fact thesis cannot be applied to all examples of the old fact-new mode hypothesis. So, like Cath (2009) asks if there is an exception to the new fact thesis, could not there be more? He says, “At the very least, we can say that if cases of coming to know an old proposition under a new first-personal mode of presentation are exceptions to the new fact thesis then (for all we know) the same could be true for practical modes of presentation” (2009, p. 151).

In addition to these critics, I will point out that the concept-mastery objection can be used to support the old fact-new mode thesis. The concept-mastery objection is an objection that Torin Alter appeals it in order to rescue phenomenal concepts and the knowledge argument from the critique of Derek Ball (2009). Unlike Alter, I will argue that the conceptual mastery thesis does not save the knowledge argument; in fact, it threatens the knowledge argument’s main claim that Mary gains new knowledge when she leaves the black and white room and sees the red for the first time. To understand the problem, let’s recall Jackson’s Mary argument. Mary, a brilliant scientist raised from birth in a black and white room. She has never seen color. But she learns all of the physical facts there is to know about human color vision by black and white television and computer screens. Jackson argues that despite all of her vast knowledge, there is still something that Mary

does not know about colors; that is what it is like to see red, for example. According to him, this is a fact that she would learn only when she is released and experiences to see red for the first time.

In this respect, one natural comment of Mary's case will be that when Mary sees red for the first time, she acquires a new concept of red that can be acquired only by undergoing such experiences. So this enables her to have new thoughts about red, thoughts might be expressed by her like, "That's what it's like to see red! I never imagined it could be so wonderful!" (Ball, 2013, p. 498).

However, this strategy works only if there are concepts that Mary can acquire only through experiencing red. According to Ball's (2009) usage, these are called 'phenomenal concepts'. Alter (2013) admits that there are no phenomenal concepts in this sense. But he claims that by appealing to concept mastery, the phenomenal concept strategy can be saved. Now, let's see Alter's examples to introduce the notion of concept mastery. First, he gives Tyler Burge's (1979) well-known example:

A man believes he has developed arthritis in his thigh. When his doctor explains that arthritis is a disease of the joints and so cannot occur in the thigh, he will presumably concede that his earlier belief was false. Even before she enlightens him about the nature of arthritis, the doctor and her patient can agree that he has arthritis in his ankles. As Ball emphasizes, the possibility of such agreement seems to require that the two share a single concept of arthritis. It is no mystery how that came about. Presumably, both picked up a public-language term that expresses the concept – a term such as 'arthritis' (2013, pp. 484-485).

According to this, before the doctor corrects his misunderstanding, the patient has not mastered the concept of arthritis yet. Alter gives another example to explain the notion of concept mastery. He states that:

Consider Joe and Josie. Joe knows practically nothing about chemistry. For example, although he has heard of atoms and molecules, he could not begin to explain what they are, how the two are related, or how chemical bonding works. By contrast, Josie is an expert chemist. Now take the truth water is H_2O . Josie knows_M⁶ that truth. Joe might know it too, despite his poor grasp of chemistry. After all, he might possess the concept of H_2O by acquiring a term for it. But if he knows that water is H_2O , he knows it only in the sense of knowledge_P. Unlike Josie, he does not know_M that water is H_2O (2013, p. 487).

What Alter intends to say here that phenomenal concepts cannot be mastered only by possessing those concepts; also we need to undergo relevant phenomenal experiences. His claim rests upon the idea that the phenomenal concept strategists can clarify Mary's epistemic progress upon seeing red for the first time by referring the conceptual mastery. In this respect, Alter concedes that Mary does not learn any new fact when she sees red for the first time. However, she acquires conceptual mastery of a concept of red. Then, her epistemic progress is related with gaining mastery of a concept of red. She gains her epistemic progress by this mastery.

⁶ Alter (2013) claims that the concept-mastery objection relies on a distinction between (a) knowledge under concepts that the knower possesses with mastery (or non-deferentially) and (b) knowledge under concepts that the knower possesses with or without mastery (or deferentially or non-deferentially). So he refers to (a) and (b) respectively as *knowledge_M* and *knowledge_P* ('M' for 'mastery' and 'P' for possession).

Alter tries to justify his claim by another analogy. He considers Nagel's (1974) thought experiment concerning a pre-Socratic philosopher being told by someone that matter is energy. And this philosopher does not know anything else about modern physics. However, he trusts his source and believes that she is reliable. In this case, we might suppose that he knows_P that matter is energy. Later, if he learns more physics, he gains full mastery of the relevant concepts. "He thus comes to know_M the same truth that he already knew_P" (Alter, 2013, p. 491). Then Alter concludes that the same thing can be applied to the Mary case. According to this, Mary's post-release gain can also be explained by conceptual mastery. When she sees the red for the first time, she does not come to know new facts. She comes to know_M the old fact that she already knew_P. That is to say, Alter admits that what Mary learns after leaving the room is not a *new* fact. This is the crucial point for my thesis because I will claim that Alter's view is compatible with the old fact-new mode thesis. If so, physical explanation of the knowledge argument seems possible, unlike Alter's basic claim. As Ball (2013) points out, Alter thinks that his view of phenomenal concepts saves the dualistic accounts of Jackson's knowledge argument against physicalism. According to Jackson's non-deducibility claim, Mary cannot a priori deduce the phenomenal truths about what it is like to see red from the complete physical facts. Namely, if ontological reduction of the phenomenal truths is possible, then they should be deducible from the physical facts (Chalmers and Jackson 2001, Jackson 1998). However, Mary cannot a priori deduce what it is like to see red from her physical knowledge. Therefore, physicalism is false. Alter gives an example related to this issue. He says that:

Consider someone who has heard talk of prime numbers and so possesses the concept 'prime number' without mastery... Such a person might not be in a position to deduce that there are infinitely many prime numbers, even if her reasoning capacity were in all other respects ideal. It does not follow that it is not a priori that there are infinitely many prime numbers. Now consider someone who has mastered prime number and assume her reasoning capacity is ideal. If she could not deduce that there are infinitely many prime numbers, this would be strong evidence that it is not a priori that there are infinitely many prime numbers. This brings out the general point that concept mastery is tied to apriority in a way that concept possession is not (2013, pp. 488-89).

According to this, if Mary's physical knowledge were enough for the conceptual mastery of redness, she would necessarily deduce what it is like to see red before she leaves the room. However, she cannot a priori deduce this fact unless she leaves the room and sees the red for the first time. Therefore, Alter concludes that physical facts are not sufficient to explain phenomenal facts.

On the other hand, Ball (2013) argues that this kind of reasoning is wrong. First, he gives a definition of a priori deducibility: "The A truths are nothing ontologically over and above the B truths only if: if an ideal reasoner knows the B truths, then she can come to know the A truths on the basis of a priori reasoning" (2013, p. 502). Then he argues Alter's arthritis example in a different way. Suppose a person who has a misconception about arthritis. He believes that arthritis can only occur in some joints such as ankles, knees, and fingers, "but that it is a matter of something like definition or conceptual truth that arthritis cannot occur in the hip joints" (Ball, 2013, p. 502). Obviously, there is no ontological

difference between the truths about arthritis and the microphysical truths about joints. But such a thinker, Ball says: “might be in no position to come to know the truths about arthritis on the basis of a priori reasoning from the microphysical truths about joints. In particular, she is in no position to know the truths about arthritis in the hip joints in this way” (2013, p. 502). Therefore, Alter needs to refuse *a priori deducibility claim* under this interpretation.

Now, let's see how I correlate the old fact-new mode thesis with Alter's conceptual mastery. First, we need to recall the main claim of the old fact-new mode thesis. Unlike other physicalist accounts, it grants that Mary gains knowledge after leaving the room and seeing red for the first time. However, she merely gets an *old* (she already knew) fact in a new way of epistemic access. On the other hand, Alter also confesses that post-release Mary does not learn any *new* fact; she already knows all of the facts before leaving the room. She only gains mastery of the relevant fact. He says, “she can know_p those truths (relevant phenomenal truths) before leaving the room” (2013, p. 488).

Therefore, I see no fundamental difference between these two arguments. Both of the arguments basically claim that Mary acquires conceptual mastery about what she already knew before. I want to explain this more clearly by giving David Papineau's example:

Suppose a researcher into educational history knows of all the 117 children in Bristol Primary School in 1910—including Archie Leach. She then learns, on reading *Movie Magazine*, that Cary Grant was also at the school in 1910. In a sense, she has learned something new. But this does not mean that there was an extra child in the school, in addition to the 117 she already knew about. In truth, Cary Grant is

one and the same person as Archie Leach. Her new knowledge is only new at the level of concepts. At the level of reference there is nothing new. The objective fact which validates her new knowledge that Cary Grant was at that school is no different from the objective fact that validated her old knowledge that Archie Leach was at the school. (Moreover, if she comes to learn that Cary Grant = Archie Leach, the fact which makes this identify true is similarly none other than the fact she always knew, that Archie Leach = Archie Leach) (2011, p. 177).

If we apply this example to Alter's conceptual mastery thesis, we can say that the researcher gains only conceptual mastery by learning that Cary Grant was also at the school. He does not learn anything new. He just learns a fact that was implicitly included in the old fact. That is to say, he makes epistemic progress about the same fact; with Alter's words, he masters what he knew before. Then, there is no reason not to say the same thing for Mary case. When she learns what it is like to see red, Mary seems to come to know something new. However, it does not mean that what Mary learns is an additional knowledge to her old knowledge. Like Papineau (2011) emphasizes in his example, her new knowledge is only new at the level of concepts. There is nothing new at the level of reference.

We certainly know that pre-release Mary's complete knowledge is physical. If my claim is true, then the post-release Mary's knowledge has to be physical. Because there is no difference between pre-release and post-release Mary's knowledge regarding the level of reference. There is no *ontological* difference. If a fact is physical, then it is always physical. Therefore, there is no reason to claim that what Mary learns after leaving the room is non-physical. Both pre-release and after-release Mary's knowledge refer to the

same fact. Tye (2009) also presents an argument that is similar to mine. He believes that Mary makes no discovery after her release because what she knows before leaving the room are the same as what she knows after she leaves. He states:

The trouble with this suggestion is that it entails that what Mary knows later is just the same as what she knew before, for there is no change in the fine-grained facts she knows. So there is no new propositional knowledge. Here is the parallel: Consider my remarking, to a friend of mine who is the world's leading authority on elm trees, 'That's an elm over there.' I can know that this is the case ... but my grasp of the concept *elm* is deferential. My friend also knows that that is an elm over there, but his grasp of the concept *elm* is non-deferential. What he knows is the same as what I know. There is no difference in the fact we know here ... but the way we grasp what we know is different. If later I become the world's leading authority on elm trees, and I repeat my earlier remark in the same situation, what I know is what I knew earlier. I make no discovery. So this strategy seems to me to offer no real progress (2009, p. 129).

However, there seems a nuance between Tye's example and my account. He does not seem to accept epistemic progress that Alter offers. As distinct from Tye, I admit that upon her release, Mary makes epistemic progress about an old fact she already knew.

Further examples might be given against Alter's thesis. What all these examples try to underline is basically the same thing: Mary does not make a genuine discovery when she leaves the room.

As a possible objection to my thesis, Alter would say that "by 'mastery' I mean 'substantial mastery'" (2013, p. 487). By the term 'substantial', Alter refers to an essential

change. According to this, if I gain a substantial mastery about an old fact I knew before, my new gaining would involve something more than my old gaining. In other words, they would not be the same facts anymore. My mastery would involve a phenomenal progress. But what is the basis of this argument? If Mary already knows all of the facts before she leaves the room, how can she gain substantial mastery of the relevant concept? As Ball (2013) points out, Alter does not give any detailed description of the term of conceptual mastery. Instead, he only gives some examples to explain his notion, and this is not enough to convince physicalists to believe that Mary's gain of conceptual mastery is substantial.

A further objection might be like this: Alter would claim that even if Mary learns an old fact in a new way, still she cannot a priori deduce the relevant fact from her complete physical knowledge. Therefore, it does not undermine the knowledge argument. A priori deducibility seems the crucial part of Alter's thesis. With regard to this, Alter states that:

To say that Mary's epistemic state changes from having knowledge_p to having knowledge_m is in a certain sense to say she learns an old fact in a new way. But this does not undermine the knowledge argument, at least in the non-deducibility/non-necessitation formulation. The gap with respect the knowledge_m suggests that the fact in question is not a priori deducible_m from microphysical facts. And that result is what generates the problem for physicalism that the knowledge argument exploits (2013, p.492).

Unlike Alter, I do not think that *the non-deducibility* formulation poses a real problem for physicalism. Because *the non-deducibility claim* can be applied to all physical facts we undergo in our lives. For example, I know what it is like to eat cheesecake because I ate cheesecake before many times (suppose fifty times). Still, my fifty-first experience of

eating cheesecake will be different than all of my previous experiences. The place will be different, the time will be different etc. Therefore, I cannot a priori deduce what it is like to eat my fifty-first cheesecake. In this case, can we say that I do not know what it is like to eat cheesecake? I do not think that we can say it. By eating my fifty-first cheesecake, I only gain new aspect about an old fact I already knew. But it does not increase my factual knowledge. I just learn what it is like to eat cheesecake in The Cheesecake Factory, for instance. So *the non-deducibility* seems to be a problem because the uniqueness of phenomenal experiences. All phenomenal experiences are unique, even if they are about the same fact.

To sum up, *the non-deducibility* seems irrelevant to whether physicalism stands or falls. That is, Alter's possible objections do not seem to be effective. In conclusion, I assert that Alter's conceptual mastery thesis might be support for the old fact-new mode thesis unlike Alter's main objective when he introduced it.

Bibliography

- Alter, T. (1998). A Limited Defense of the Knowledge Argument. *Philosophical Studies*, 90(1): 35-56.
- Alter, T. (2008). Phenomenal Knowledge without Experience. *The case for qualia*, Edmond L. Wright Ed., 247.
- Alter, T. (2013). Social Externalism and the Knowledge Argument. *Mind*, 122: 481-96.
- Ball, D. (2009). There are No Phenomenal Concepts. *Mind*, 118: 935-62.
- Ball, D. (2013). Consciousness and Conceptual Mastery. *Mind*, 122(486): 497-508.
- Beaton, M. (2005). What RoboDennett Still Doesn't Know. *Journal of Consciousness Studies*, 12(12), 3-25.
- Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, 4: 73-121.
- Cath, Y. (2009). The Ability Hypothesis and the New Knowledge-how. *Noûs*, 43(1), 137-156.
- Chalmers, D. J. (1995). Facing up to the Problem of Consciousness. *Journal of consciousness studies*, 2(3), 200-219.
- Chalmers, D. J. (1996). *The Conscious Mind: In search of a fundamental theory*. New York: Oxford University Press.
- Chalmers, D. J. and Jackson, F. (2001). Conceptual Analysis and Reductive Explanation. *Philosophical Review*, 110: 315-61.
- Chalmers, D. J. (2004). Phenomenal Concepts and the Knowledge Argument. In Peter Ludlow, Yurin Nagasawa and Daniel Stoljar. Eds. *There's Something About Mary*: 269-98. Cambridge, Mass: MIT Press.

- Coleman, S. (2009). Why the Ability Hypothesis is Best Forgotten. *Journal of Consciousness Studies*, 16(2-3), 74-97.
- Crane, T. (2001). The Origins of Qualia. *History of the mind-body problem*. T. Crane and S. Patterson. Eds., Routledge: 169-94.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston MA: Little, Brown & Co.
- Dennett, D. C. (2006). What Robomary Knows. *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Torin Alter and Sven Walter. Eds. Oxford University Press.
- Graham, G., & Horgan, T. (2000). Mary Mary, Quite Contrary. *Philosophical Studies*, 99(1), 59-87.
- Horgan, T. (1984). Jackson on Physical Information and Qualia. *Philosophical Quarterly*. 34: 147-83.
- Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, 32(127), 127-136.
- Jackson, F. (1986). What Mary didn't Know. *The Journal of Philosophy*, 83(5), 291-295.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. New York: Oxford University Press.
- Jackson, F. (2004). Mind and Illusion. In Peter Ludlow, Yujin Nagasawa and Daniel Stoljar. Eds. *There's something about Mary*: 421-42. Cambridge, Mass: MIT Press.
- Jackson, F. (2005). Consciousness. In Frank Jackson and Michael Smith. Eds. *The Oxford Handbook of Contemporary Philosophy*: 310-33. Oxford: Oxford University Press.
- Keeley, B. (2009). The Early History of the Quale and its Relation to the Senses. *Routledge Companion to the Philosophy of Psychology*, Routledge Press: 71-89, New York.

- Lewis, C. I. (1929). *Mind and the World-order: Outline of a theory of knowledge*. Courier Corporation.
- Lewis, D. (1979). Attitudes De Dicto and De Se. *The Philosophical Review* 88: 513-43.
- Lewis, D. (1983). Postscript to 'Mad Pain and Martian Pain' in his *Philosophical Papers* Vol. 1., New York: Oxford University Press: 130-2.
- Lewis, D. (1988/2004). What Experience Teaches in *There's Something about Mary: Essays on Frank Jackson's Knowledge Argument Against Physicalism*. P. Ludlow, Y. Nagasawa and D. Stoljar. Eds. Cambridge MA: MIT Press: 77-103.
- Malatesti, L. (2008). Maryno Znanstveno Znanje. *Prolegomena: časopis za filozofiju*, 7(1), 37-59.
- Mellor, D. H. (1993, January). The Presidential Address: Nothing Like Experience. In *Proceedings of the Aristotelian Society* (Vol. 93, pp. 1-16). Aristotelian Society, Wiley.
- Nagel, T. (1974). What is it like to be a Bat?. *The philosophical review*, 83(4), 435-450.
- Nemirow, L. (1980). Review of Nagel's *Mortal Questions*, *Philosophical Review*. 89: 473-77.
- Nemirow, L. (1990). Physicalism and the Cognitive Role of Acquaintance. In Lycan 1990.
- Papineau, D. (2011). Phenomenal Concepts and the Private Language Argument: *American Philosophical Quarterly*, Vol. 48, No. 2, Wittgenstein and Naturalism: 175-184.
- Price, H. H. (1932). *Perception*, London: Methuen & Co.
- Prinz, J. (2005). *Mental Maintenance: A Response to the Knowledge Argument*.

- Ryle, G. (1946). Knowing How and Knowing That. *Gilbert Ryle: Collected Papers*. Gilbert Ryle Ed. 2: 212-25. New York: Barnes and Noble.
- Ryle, G. (1949). The Concept of Mind. Chicago: University Press.
- Searle, J. R. (1998). How to Study Consciousness Scientifically. *Consciousness and human identity*. J. Cornwell. Ed. New York, Oxford University Press: 21-37.
- Stanley, J., & Williamson, T. (2001). Knowing How. *The Journal of Philosophy*,98(8), 411-444.
- Thompson, E. (1995). Colour Vision, London: Routledge.
- Tye, M. (1986). The Subjective Qualities of Experience. *Mind*, 95: 377, pp. 1-17.
- Tye, M. (2000). Consciousness, Color, and Content. Cambridge MA: MIT Press.
- Tye, M. (2009). Consciousness Revisited: *Materialism without Phenomenal Concepts*. Cambridge, MA: The MIT Press.