PHYSICALISM AND THE PHENOMENAL-PHYSICAL GAP: CAN A POSTERIORI NECESSARY PHYSICALISM ADEQUATELY RESPOND TO THE PROBLEM OF PHENOMENAL SUBJECTHOOD?

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MURAT ARICI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF PHILOSOPHY

April 2011

Approval of the Graduate School of Social Sciences

_____

Prof. Dr. Meliha Altunışık
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

_____

Prof. Dr. Ahmet İnam
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

_____

Assoc. Prof. Erdinç Sayan
Supervisor

Examining Committee Members

Prof. Dr. David Grünberg	(METU, PHIL)	_____
Assoc. Prof. Erdinç Sayan	(METU, PHIL)	_____
Prof. Dr. Ayhan Sol	(METU, PHIL)	_____
Assist. Prof. István Aranyosi	(BILKENT, PHIL)	_____
Assist. Prof. Hilmi Demir	(BILKENT, PHIL)	_____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Murat ARICI

Signature:

ABSTRACT

PHYSICALISM AND THE PHENOMENAL-PHYSICAL GAP: CAN A POSTERIORI NECESSARY PHYSICALISM ADEQUATELY RESPOND TO THE PROBLEM OF PHENOMENAL SUBJECTHOOD?

Arıcı, Murat
Ph.D., Department of Philosophy
Supervisor: Assoc. Prof. Erdinç Sayan

April 2011, 155 pages

Phenomenal consciousness presents a recalcitrant problem for the scientific conception of the world and the physicalist thesis that claims that everything that exists (including whatever is involved in any mental phenomena) is physical and physically explainable. Thus, on this view, every truth is a physical truth. By Putnam-Kripkean considerations and for several other reasons, I defend the claim that any version of such a physicalist thesis must be a necessary thesis, which ultimately means that contingent physicalism is not tenable.

Against this thesis, philosophers have put forward several anti-physicalist arguments including the knowledge argument, the conceivability/modal argument, the explanatory gap argument, and the property dualism argument. All these arguments rest on the assumption of an epistemic/explanatory gap, which I call the "phenomenal-physical gap," between the phenomenal and the physical. I claim that the phenomenal-physical gap (the PP-gap) is unbridgeable, from which it can be concluded that a priori physicalism is not tenable.

The phenomenal concept strategy (PCS), which is a specific strategy within a posteriori necessary physicalism, aims at offering an explanation in physical terms of why we have such an unbridgeable gap by differentiating between phenomenal and physical concepts in a fundamental way. Nevertheless, proponents of PCS—the most promising version of a posteriori necessary physicalism—face a severe problem that I call "the problem of phenomenal subjecthood" in explaining in physical terms why we have the PP-gap.

The phenomenon of "experiencing" consists of three substantially existing elements: the phenomenal subject (the experiencer), the experiential item (what is experienced by the subject), and the phenomenal s-v-o relation (the experiential relation) between the first two. I argue for the substantial existence of phenomenal subjects based on an argument I provide, the reality of some mental phenomena such as phenomenal unity and continuity, and the mental facts concerning phenomenal peculiarity, phenomenal agency, and the sense of phenomenal I-ness, the reality of all of which one cannot deny.

Since PCS accounts are mostly qualia-centered accounts that ignore the reality of phenomenal subjects and the phenomenal s-v-o relation, they cannot account for the PP-gap in physical terms without first offering substantial theories of phenomenal subjecthood. But once they grant the substantiality of phenomenal subjects, they face severe difficulties in establishing their accounts of the nature of phenomenal concept, and thus the PP-gap in physical terms.

Keywords: phenomenal consciousness, physicalism, phenomenal concept strategy, phenomenal subjecthood, phenomenal I-ness.

ÖZ

FİZİKALİZM VE FENOMENAL-FİZİKSEL GEDİĞİ: APOSTERİORİ ZORUNLULUKÇU FİZİKALİZM, FENOMENAL ÖZNELİK SORUNUNA YETERLİ BİR ÇÖZÜM ÜRETEBİLİR Mİ?

Arıcı, Murat
Doktora, Felsefe Bölümü
Danışman: Doç. Dr. Erdinç Sayan

Nisan 2011, 155 sayfa

Fenomenal bilinç, birbiriyle ilgili iki bakış açısına direnen inatçı bir problem olma niteliği taşır: Bunlardan birincisi genel olarak bilimsel anlayıştır ve diğeri de zihinsel süreçler de dâhil olmak üzere evrende var olan her şeyin fiziksel olduğunu ve fiziksel terimlerle açıklanabileceğini—ki böylelikle bütün doğruların fiziksel doğrular olduğunu—iddia eden fizikalist tezdir. Ben bu tezde Putnam-Kripke'ci yorumlamalar ve diğer bazı nedenlerden yola çıkarak bu adı geçen fizikalist tezin her türlü versiyonunun zorunlulukçu olması gerektiğini savunuyorum. Ki bu da olumsal fizikalizmin savunulamaz olması anlamına geliyor.

Bu fizikalist teze karşı filozoflar, bilgi argümanı, düşünebilirlik/kiplik argümanı, açıklayıcılık gediği argümanı ve nitelik ikiciliği argümanlarını da içeren çeşitli karşıt argümanlar ileri sürdüler. Bu argümanların hepsi de benim "fenomenal-fiziksel gediği" diye adlandırdığım, fenomenal ile fiziksel arasındaki epistemik/açıklayıcılık gediğine dayanmaktadır. Bu fenomenal-fiziksel gediğinin (FF-gediği) kapatılamaz olduğunu iddia ediyorum. Bundan da apriori fizikalizmin savunulamaz bir tez olduğu sonucu çıkmaktadır.

Aposteriori zorunlulukçu fizikalizm içinde spesifik bir strateji olan fenomenal kavram stratejisi (FKS), neden böyle kapatılamaz bir gedikle yüzleştiğimizin açıklamasını fiziksel terimlerle vermeyi amaçlıyor. Bunu, fenomenal ve fiziksel kavramlar arasındaki çok temel bir farka işaret ederek yapmak istiyor. Fakat aposteriori fizikalizmin en ümit vadeden versiyonu olan FKS'nin savunucuları neden böyle bir gedik (ya da boşluk) ile yüzleşmek zorunda kaldığımızı fiziksel terimlerle açıklarken benim "fenomenal öznelik problemi" diye adlandırdığım ciddi bir problemle karşı karşıya gelmektedirler.

"Deneyimleme" fenomeni tözel/tözsel olarak var olan üç unsurdan oluşmaktadır: fenomenal özne (deneyimleyen), deneyimlenen öge (özne tarafından deneyimlenen şey) ve bu ilk ikisinin arasındaki fenomenal özne-yüklem-nesne bağı (deneyimleme bağı). Bu tezde fenomenal öznelerin tözel/tözsel olarak var oldukları savımı, ileri sürdüğüm bir argümana, fenomenal birlik ve devamlılık gibi zihinsel fenomenlerin gerçekliğine ve yine gerçekliğini hiç kimsenin inkar etmek/yadsımak istemeyeceği fenomenal kendine özgülük, fenomenal faillik ve fenomenal ben-lik duygusu ile ilgili zihinsel olgulara dayandırıyorum.

FKS açıklamaları fenomenal öznelerin ve fenomenal özne-yüklem-nesne bağının realitesini görmezden gelen ve çoğunlukla kuale merkezli açıklamalar oldukları için, fenomenal öznelikle ilgili ciddi teoriler sunmaksızın FF-gediğini fiziksel terimlerle açıklayamazlar. Öte yandan, bu açıklamalar fenomenal öznelerin tözel/tözsel varlığını kabul ettikleri anda da fenomenal kavramların doğasını ve dolayısı ile FF-gediğini fiziksel terimlerle açıklama girişimlerinde ciddi zorluklarla ve önemli bir ikilemle yüzleşmek zorunda kalırlar.

Anahtar Kelimeler: fenomenal bilinç, fizikalizm, fenomenal kavram stratejisi, fenomenal öznelik, fenomenal ben-lik.

To Those Who Has Supported Me in My Education

TABLE OF CONTENTS

PLAGIARISM iii

ABSTRACT iv

ÖZ vi

DEDICATION viii

TABLE OF CONTENTS ix

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiii

CHAPTER

1. INTRODUCTION: THE PROBLEM OF PHENOMENAL CONSCIOUSNESS 1

    1.1. Why Is Phenomenal Consciousness a Puzzle? 1

    1.2. Materialist/Physicalist Approach to Phenomenal Consciousness 10

2. MINIMAL PHYSICALISM AND THE THREE TYPES OF RELATION: IDENTITY, SUPERVENIENCE, AND REALIZATION 20

    2.1. Monism vs. Dualism 20

    2.2. The Identity Relation 22

        2.2.1. Reductive Physicalism Using the Identity Relation 23

        2.2.2. Type vs. Token Identity 25

        2.2.3. A Priori vs. A Posteriori Physicalism 27

        2.2.4. The Problem of Modal Status: The Necessity vs. Contingency of the Physicalist Thesis 29

2.3. The Supervenience Relation 31

2.3.1. Supervenience Physicalism 32

2.4. The Realization Relation 36

3. ANTI-PHYSICALIST ARGUMENTS AND THE GAP BETWEEN THE PHENOMENAL AND THE PHYSICAL 45

3.1. Arguments against Physicalism 45

3.2. Contemporary Arguments against Physicalism 47

3.3. The Epistemic/Explanatory Gap 55

3.4. Is the Phenomenal-Physical Gap Bridgeable? 59

3.5. The Phenomenal-Physical Gap and the Intuition of Distinctness 62

4. TYPE-B PHYSICALISM, PHENOMENAL CONCEPT STRATEGY AND THE PROBLEM OF PHENOMENAL SUBJECTHOOD 66

4.1. Two Types of Physicalist Response to the Phenomenal-Physical Gap 67

4.2. The Phenomenal Concept Strategy (PCS) 71

4.3. The Unique Features of Phenomenal Experience/Concepts 79

4.4. PCS and the Problem of Phenomenal Subjecthood 82

5. SUBSTANTIAL EXISTENCE OF PHENOMENAL SUBJECTS 90

5.1. Consciousness and Phenomenal Subjecthood 91

5.1.1. Creature vs. State Consciousness 92

5.1.2. Conscious vs. Non-Conscious States 95

5.2. Two Dimensions of Phenomenality: Subjectivity vs. Qualitativity 98

5.2.1. The Connection between Subjectivity and Qualitativity 99

5.2.2. Phenomenal Unity 101

5.2.3. Phenomenal Continuity 103

5.3. The Nature of Phenomenal Subjectivity 106

5.3.1. Phenomenal Peculiarity 107

5.3.2. Phenomenal Agency 108

5.3.3. The Sense of I-ness 110

6. A DILEMMA FOR PHYSICALISTS IN RESPONDING TO THE PROBLEM OF PHENOMENAL SUBJECTHOOD 116

6.1. Two Models for the Metaphysical Structure of "Experiencing" 117

6.2. Type-B Physicalism Revisited 123

6.3. Phenomenal Concept Strategy Revisited 126

6.4. Paths to Go 131

BIBLIOGRAPHY 135

APPENDICES

A. CURRICULUM VITAE 142

B. TURKISH SUMMARY 144

LIST OF FIGURES

FIGURES

Figure 4.1 A type-B physicalist's way of thinking about phenomenal consciousness 71

Figure 6.1 Non-Constitutive model for the structural relation between a phenomenal subject and qualia 120

Figure 6.2 Three versions of the constitutive model for the structural relation between a phenomenal subject and qualia in a phenomenal state 121

LIST OF ABBREVIATIONS

ABBREVIATIONS

(p)-statements:	Physical-to-physical statements
(m)-statements:	Mental-to-mental statements
(mp)-statements:	Mental-to-physical statements
The PP-gap:	The Phenomenal-Physical Gap
(NC):	The Nomological-Physical Constraint
(LC):	The Logical-Phenomenal Constraint
(PC):	The Provability Constraint for Referential Terms
PCS:	The Phenomenal Concept Strategy
(ES):	Desideratum for the Existence of Phenomenal Subjects
(PUT):	The Phenomenal Unity Thesis
PPB:	The Problem of Phenomenal Binding
(PG):	The Thesis of Phenomenology of Change
(PCT):	The Phenomenal Continuity Thesis
(RE):	Desideratum for the Relation of Experiencing
(SS):	Desideratum for the Substantiality of Phenomenal Subjects
(NP):	Desideratum for Non-Publicity of Phenomenal Tokens
(PA):	Desideratum for Phenomenal Agency
(SI):	Desideratum for the Sense of I-ness

CHAPTER 1

INTRODUCTION: THE PROBLEM OF PHENOMENAL CONSCIOUSNESS

1.1. Why Is Phenomenal Consciousness A Puzzle?

For millennia, philosophers have been asking numerous perplexing questions about the nature of the world and of the human being. Some have been solved, at least within the current conceptions of what counts as a solution to philosophical problems, but some have been continuing to confuse minds. It is true that philosophers have made a lot of progress in understanding the material world. This was not achieved by mere philosophical thinking, of course. Modern physicists and chemists have revealed many aspects of the nature of the material world. Modern biologists, too, have discovered quite many physical features of living organisms. What we see around us including our physical bodies as well as the realms of animals and plants, mountains, oceans, celestial objects, etc. are now much less mysterious to us than they were to those who lived a couple of centuries ago. Not only these material objects but also phenomena resulting from the interactions between material objects, such as electricity, radiation, electromagnetism, gravitation, etc. are now less puzzling than before. As a result, philosophers and scientists developed a rightful self-confidence in their belief that physical and biological sciences can in principle explain all the phenomena there are in the world.

There are, however, some recalcitrant phenomena that damage this self-confidence—phenomena that have not been adequately explained for thousands

of years. Phenomenal consciousness[1] (among other mental phenomena) is specifically one of them. It appears to be one of the most mysterious phenomena in the world. How can a mere material substance like a slushy brain posses such an astonishing and mysterious feature? Suppose you know little about human biology and have never seen an image of the human brain or any part of it. And someone shows you a close-up image of neuronal activity in the brain on a screen. There is no way for you to figure out why such brain activities would be accompanied by such an amazing phenomenon like consciousness, and why, cell activities, say, in plants would not. What is more is that there seems to be no way to figure out why a given human brain is accompanied by *this* particular phenomenal consciousness but not *that* one. On the other hand, someone with appropriate knowledge about the behavior of H_2O molecules might predict many of the surface features of water like liquidity and transparency.

What kind of phenomenon is phenomenal consciousness then? It is really hard to give a satisfactory definition of phenomenal consciousness. For one thing, any such definition will require one or other non-agreed-upon theory of consciousness. For another thing, the term 'consciousness' itself[2] is used with many different meanings. But we can begin with some of these different meanings, and then get a grip on the phenomenal consciousness of a creature. The term 'consciousness' can mean, for example, "wakefulness" as in the sentence "The patient regained her consciousness," or mean "awareness" as in the sentences "I am conscious of the situation we are in," and "You should be more conscious environmentally." 'Consciousness' can also mean "sentience" as

[1] When used as an adjective attached to 'consciousness,' the term 'phenomenal' expresses an intrinsic quality of conscious mental states, which will be explained in more detail shortly. The term is being used in different senses in a variety of philosophical contexts, however. The readers should not be confused by other usages of the term throughout this thesis, and only keep in mind its specific use in the given context.

[2] I assume no difference between the adjective form 'conscious' and the noun form 'consciousness' in terms of ontological commitments. The noun form is just an abstraction of the property attributed by the adjective form.

in the sentence “This alien creature is a conscious being; it can sense the stimuli.” And finally, the term can indicate “self-awareness,” meaning “the awareness of one’s own awareness,” as in the sentence “Infants and most animals are not self-conscious.”

In one sense, the variety of usages for ‘consciousness’ does not help us understand phenomenal consciousness in a better way unfortunately. It might complicate the issue. In another sense, however, the various employments of the term ‘consciousness’ present a rich set of tools to comprehend different aspects of consciousness. All in all, the employment of the term in quite different contexts does not mean we cannot speak of the nature of phenomenal consciousness. We can at least describe or characterize crucial aspects of its nature. This may involve using synonyms or metaphors, or giving examples. Nowadays, philosophers use several terms to characterize the whole or some part of the nature of phenomenal consciousness, such as “experiential character,” “subjective feel,” “raw (sensory) feel,” and Thomas Nagel’s famous term “what-it-is-likeness” (Nagel 1974) of having a mind and mental states. These terms are basically aimed to distinguish phenomenally conscious beings from other types of being. I am phenomenally conscious, for example, but the computer I am using now to type these lines is not, because I *experience* several (visual, auditory, tactile, etc.) aspects of the writing activity at this moment; I can be in various emotional states (like joy, sadness, fear, anger, etc.) having *qualitative characters*; I sometimes desire things (wanting a cup of Turkish coffee, longing for a person, hoping to see a movie, etc.) and my desires involve *subjective feels*; and *there is something it is like* to be me in general. Nevertheless, my computer lacks all these features. It is not capable of experiencing anything. It is not phenomenally conscious. In fact, it does not have any type of consciousness at all; neither do any other non-living objects around me like tables, chairs, bookshelves, etc.

The notion of "experience" is usually taken to be definitive of phenomenal consciousness. But here we should carefully note that some phenomena such as blindsight[3] motivate philosophers to interpret the nature of an "experience" in two fundamentally different directions. In one direction, some interpret an experience as essentially involving a *subjective feel*, which is what exactly constitutes the phenomenality of consciousness. So, on this construal phenomenal consciousness can be equated with experience. In another direction, on the other hand, some construe experience as involving merely a *distinctive feel*, a purely qualitative character that nevertheless lacks subjectivity. On this construal, therefore, an experience can be *non-conscious*, and thus cannot be equated with the phenomenality of consciousness. We may call the former construal of experience "phenomenal construal" and the latter one "qualitative construal." Accordingly, if you are inclined towards the phenomenal construal, cases that involve qualitativity but lack subjectivity (or awareness of a subject) do not exhibit an experience for you. If you adopt the qualitative construal, on the other hand, you hold that the same cases present experiences that are non-conscious. It seems that the phenomenal construal of experience is more intuitively appealing. When asked, ordinary people would say "I would not call 'experience' those mental states that occur without a full-blooded awareness." I will use the term 'experience' and 'experiential' in this more intuitive sense based on the phenomenal construal throughout the rest of this thesis.

There is another way to interpret those special cases that involve qualitativity but lack a subject's awareness. Adopting the qualitative construal, one may here

[3] Blindsight patients suffer from a complete blindness in *some area* of their visual field because of some damage on their retina or in a certain part of their brain. On specifically designed experiments, they nevertheless surprisingly guess with a high degree of correctness what is happening visually on the area to which they are blind to, though they insist that they do not see anything at all. This is taken by many to imply that these patients are non-conscious of their visual experiences on those areas, which means they have *non-conscious experience*. But, I believe, there are two different conceptions of "experience" in play here.

think that subjects of those cases have phenomenal consciousness but lack a kind of non-phenomenal consciousness that Ned Block calls “access consciousness” (1995). According to Block, a mental state is access conscious if, in virtue of having that state, a subject can use its content for rational inference, deliberate control of behavior, and verbal report. Accordingly, blindsight patients are not access conscious of the blind areas in their visual fields since they do not satisfy any of the above three conditions, though they are phenomenally conscious (when qualitatively construed) of the same areas since they are able to correctly guess what is happening visually there albeit being unaware of it.

So far we have been speaking of phenomenal consciousness as a feature of creatures; i.e. as a creature’s being phenomenally conscious. It is clear that if a creature is a phenomenally conscious being, then some relevant components involved in the conscious domain of that creature will also be said to be phenomenally conscious. If I am phenomenally conscious at the moment, then the current relevant *mental sates* of mine can be said to be phenomenally conscious as well. The same goes for the properties of those mental states, and for the processes and events occurring in my mind when I am phenomenally conscious. They are phenomenally conscious *mental states, processes, events and properties* too.

Are there *conscious* mental states that do not involve an experiential character or any phenomenology? It seems that there are. Our brain, for example, processes a lot of information when we are asleep or even when we are awake without our phenomenal awareness. They may not have any *subjective feel*. So we are not *phenomenally* conscious of those processes though we may be capable of cognitively employing the information processed. We may control our behaviors based on that information; and we may even give a verbal report of the information being processed without phenomenally experiencing any *subjective*

and *qualitative* features involved. These and some other similar mental processes that David Chalmers calls the "easy problem" of mentality (Chalmers 1996 and 2003) do not pose a serious threat to the scientific conception of the world as phenomenal consciousness does.

Returning to our original problem, phenomenal consciousness described above as having experiential (subjective and qualitative) character resists to the scientific conception of the world in two ways: metaphysical and epistemological. Metaphysically speaking, most contemporary scientists and philosophers believe that whatever is involved in any observable phenomena of the world is included on the list of ontologically fundamental entities of physical and biological sciences—the list that only comprises material/physical entities. In other words, there exists nothing beyond the fundamental entities studied by physics and biology. Is phenomenal consciousness included on that list too? Is it something material/physical (or functional, realized by material/physical entities)? You might think it must be, in accordance with your scientific conception of the world. But most people have intuitions that phenomenal consciousness presents qualities that can hardly be included on the list of ontologically fundamental entities of physics and biology. These qualities are the ones that are related to the subjective and qualitative aspects of phenomenal consciousness.

The situation is not better when you think of the issue epistemologically. There are tools (like observation and experimentation) that are legitimate to use in science to acquire knowledge of nature. If phenomenal consciousness is material/physical, then we must be able to acquire knowledge also of phenomenal consciousness using those tools. But how are we going to do that while we are having trouble in understanding even its most basic qualities and do not know the proper way of understanding its nature? Using scientific tools to comprehend the nature of phenomenal consciousness seems to give no

philosophical insight into it at all. This is why we need philosophical tools over and above the scientific ones to understand it.

To illustrate the metaphysical and epistemological predicament we face in the case of phenomenal consciousness, suppose that I am having a perceptual experience with a subjective and qualitative character: I am looking at and touching the red leather case of my camera on the table at this very moment. Clearly a lot of physical/chemical, neurological and cognitive processes are occurring in my brain during this experience of mine. But these processes supposedly involve only material entities that can be perfectly analyzed in depth by physicists/chemists, neurologists and cognitive scientists. Nevertheless, my experiencing the perception of the red leather case has subjective and qualitative features too. My experience is distinctively subjective since no one else can have *the one and the same* experience: I have the strong feeling that it is only *my* experience, and no one else can be subjected to this experience as *my experience*. Someone else can, perhaps, have an experience qualitatively identical to mine, but that would be a numerically different experience than mine, which is very unique and private to my mental life.

Furthermore, the experience I am having of the perception of red leather case presents certain distinctive qualities—called "qualia" in the literature—such as the redness and softness that I enjoy in the domain of my consciousness. These distinctive qualities do not seem to be features of the thing I perceive—the red leather case of my camera. For when I close my eyes, and visualize the same object I perceived a moment ago, I am still presented the same qualitative features. Even in the absence of the object of my perception, my experience of visualizing the red leather case has the same kind of visual qualities: redness, opaqueness, softness of texture, etc. It is these qualities and subjectivity that scientists have to explain in material/physical terms, but so far could not.

There is one further issue concerning phenomenal consciousness, which is fundamentally relevant to the subjective character of phenomenal consciousness: The problem of the "sense of self" (or as I will call later "sense of I-ness"). We might deny that there is a "self-like entity" within our existence, adopting a nominalist stance or Humean position embracing his bundle theory of mind, but we must admit that there is at least an undeniable "*sense* of self" that we feel when we turn into our inner mental lives. This sense of self might be an illusion corresponding to nothing, ontologically speaking, as the Humean position claims, but we cannot deny the existence of the "sense" itself. And there must be a reason for its existence. We must at least investigate why there is a sense like that. I believe an adequate answer would be quite relevant to, and even be definitive of the nature of subjectivity and qualitative features of mind. Most philosophers of mind usually prefer not to deal with this problem when they examine subjectivity and qualitativity. It is because of the Humean position they adopt and their belief that the sense of self we feel should ultimately be dissolved when the true nature of subjectivity and qualia is completely revealed. Nevertheless, this is mistaken since the sense of self might not be a product of, but may directly or indirectly be a constitutive element of phenomenal consciousness. Indeed I will treat it as evidence for the substantial nature of phenomenal subjecthood[4] when I examine my chief target, the subjective character of phenomenal consciousness. So, it deserves an equal attention as subjectivity and qualia receive.

Having clarified the essential features of phenomenal consciousness this way, there are several reasons for the resistance of phenomenal consciousness to the incorporation to the scientific conception of the world. Two methodological ones, which I call epistemological barriers, are the following. First, phenomenal

[4] A notion that I will introduce in chapter 4, and fully examine in Chapter 5.

consciousness is not a publicly observable phenomenon. Whatever we see around us, including every part of our material bodies, can be examined publicly by more than one observer at the same time. We can check whether what we suppose exists or is happening at a given time is the same as what others observe exists or is happening at the same time. We can at least intelligibly assume that we and others are in the same epistemological position with respect to a material being or an event involving material entities. I can quite rationally assume that the red leather case of my camera on my desk can be equally observed and examined by other people with an equal epistemological status. The same goes for the events occurring around me. I can compare what I am observing to be happening at a certain time, say raining outside of my house, with what others observe to be happening at the same time in the same location, and find out if the features of the event of raining are presented to me and the others in the same way. Phenomenal consciousness, on the other hand, cannot be observed publicly—whether it is considered to be a thing or event. The phenomenon of consciousness literally belongs to a subjective and very private domain, which prevents it to be examined from a third-person perspective. We can open up someone's skull and observe the neurobiological processes happening in the brain and examine the relevant material elements, but we cannot observe and examine (from the third-person or public perspective) phenomenal events (which have subjective and qualitative features) supposedly occurring in the same location.

Second, investigating the nature of phenomenal consciousness is a reflexive or self-investigative action. I cannot observe and examine someone else's phenomenal experiences from the public perspective. But can I not observe and examine my own phenomenal experiences from the public perspective as I myself can surely observe and examine the neurobiological events of my own brain, say with the help of mirrors or cameras and monitors? No. Examining our own

phenomenal consciousness is a self-investigative effort that can be done only from our very own first-person perspective, which also makes the investigation itself epistemologically subjective and closed to third parties—other epistemic agents. It is clear that investigating phenomenally conscious events is fundamentally epistemologically different from all other investigations of natural phenomena in the world.

Consequently, these epistemological barriers do not allow the scientific method to treat the conscious phenomena in the same way it does other natural phenomena. Phenomenal consciousness with its fundamentally different character explained so far damages the self-confidence scientists and philosophers achieved for the last couple of centuries. This even causes some to believe that physical and biological sciences cannot even in principle reveal the real nature of phenomenal consciousness since its purportedly non-physical qualities are beyond the scope of these sciences (Chalmers 1996 and 2003; Strawson 2000; and Stoljar 2001). It is because of this reason that we will investigate whether a naturalist explanation of phenomenal consciousness can be given within a completely materialist framework.

1.2. Materialist/Physicalist Approach to Phenomenal Consciousness

The problem of phenomenal consciousness as presented above is a problem from two standpoints: naturalistic and scientific. Though it has no precise meaning in philosophy, naturalism in contemporary philosophy basically stimulates the ontological principle that nature (as substances, properties, relations, etc.) is all that there is. There is nothing "supernatural" (Kim 2003). And epistemologically speaking, the doctrine simply urges the idea that in investigating reality, philosophy should always be in close contact with the scientific method. Considered as such, we may think of these two standpoints as a single one, and call it the "standpoint of *empirical philosophy*," which emphasizes several

metaphysical and epistemological assumptions common to Anglo-Saxon analytic philosophy and modern science. We have to ask, on the other hand, whether phenomenal consciousness presents itself as a problem from other standpoints. The answer is clearly "yes." But the character of the problem, and thus the central questions raised around the issue, changes depending on the standpoint one adopts. If one's philosophical stance, for example, embraces supernatural entities, one will still want to explain things within the nature of phenomenal consciousness, but integrating phenomenal consciousness into our scientific conception of the world will be neither a central nor a subordinate issue that has to be dealt with.

What makes phenomenal consciousness a puzzle for the naturalist and scientific conception of the world then? Is it because of its metaphysical character or because of our epistemic access (or lack of access) to its nature? It is certainly because of both. In fact, the metaphysical character determines the conditions of our epistemic access to it. And the conditions of our epistemic access to it determine how much we can know about its metaphysical character. But the question of its metaphysical character has become a more central issue among philosophers of mind because of the popularity of naturalism. For many, naturalizing mind (the project of integrating our conception of mind into our scientific conception of the world, i.e. explaining mind in terms that refer only to natural properties) will also provide answers to metaphysical issues regarding phenomenal consciousness. It is the materialistic view which is the background ideology behind the project of naturalizing mind. On the metaphysical level, the materialistic view claims that everything that exists is material; there is nothing above and beyond material entities. On the epistemological level, the view basically adopts the scientific approach and defends the claim that every truth is a scientific (or physical) truth and that to know about nature we need only employ standard empirical methods guided by relevant rational tools.

Let us now briefly look at several possible approaches to the problem of phenomenal consciousness. We can then go into the finer details of the materialistic approach, and in the next chapter we can examine it fully. We have already said a few things about the naturalist project. What we should additionally distinguish under this project is between (i) the standard scientific view and (ii) the non-standard scientific view. Under (i), we should cite (1) the materialist approach, which is our chief concern. Under (ii), we can cite three approaches: (2) naturalistic dualism, (3) non-standard scientific monism, and (4) cognitive limitationism. There can of course be a non-naturalist approach to the problem of phenomenal consciousness as well. Under this approach, we should cite (5) supernaturalism and (6) idealism. What follows are brief descriptions of these six approaches.

Supernaturalism: One non-naturalist approach, as clearly hinted above, is to think that phenomenal consciousness is not part of nature. It is wholly or partially *supernatural*. Thus, we cannot acquire knowledge of phenomenal consciousness by standard philosophical/scientific (or rational/empirical) methods. In order to know about its nature we have to have access into its *supernatural nature*. Human cognitive capacity might include this access or it might not, we do not know. That is another issue. The crucial thing here is that if phenomenal consciousness is wholly or partially supernatural, standard rational/empirical methods will not suffice to acquire knowledge about its nature. We will need non-standard methods to investigate it.

Idealism: Another approach to phenomenal consciousness within non-naturalism is to see the whole reality as consisting merely of consciousness, but not matter. Consciousness is what there is, and is the ground of everything that exists. This is a monistic view, but completely the opposite of materialistic monism. According

to this approach, what we can know about the world is obviously not about something material, but rather is only the content of our own phenomenal consciousness. So, epistemological methods designed to know about the material world will not work in knowing about phenomenal consciousness. To know about phenomenal consciousness we need epistemological methods compatible with the metaphysical world view of idealism.

Naturalistic Dualism: Another view that might shape our epistemological approach to phenomenal consciousness is to see the reality as consisting of two fundamental elements: mind and matter. This should be considered a naturalist approach because mind and matter are still considered to be the basic metaphysical components of nature. If the reality consists of the facts of these two distinct substances, however, methods of acquiring knowledge of these facts will clearly differ. Ways to know about matter will not be the same as ways to know about mind. And since phenomenal consciousness is the most crucial characteristic of mind, to know about phenomenal consciousness, we will need different epistemological apparatuses than those we might be using in knowing about matter, such as empirical and rational tools.

Cognitive Limitationism: Another naturalist option is to reject any supernatural theory and hold that phenomenal consciousness is part of nature. Within this approach, however, one may still believe that we cannot acquire knowledge of phenomenal consciousness either by standard rational/empirical methods or by non-standard ones. It is because of the epistemic barriers we encounter when we try to penetrate into the relevant domain epistemologically. These epistemic barriers might result from limited human cognitive capacities (such as limited conceptual abilities), or from limited biological capacities (such as limited neurobiological abilities), or from nature itself (such as the special character of conscious phenomena not allowing scientific investigation). I would like to call

this approach “cognitive limitationism” or “limitationist naturalism.” It is possible to adopt such a view and claim that phenomenal consciousness is entirely closed to human cognition.

Non-Standard Scientific Monism: Another option is to remain a naturalist while still rejecting that we can know about the nature of phenomenal consciousness by the standard scientific/empirical methods. According to this view, non-standard scientific tools and methods are required to know about the nature of phenomenal consciousness. The motivation for such an approach usually results from the ambiguity of our conception of “material.” How one defines ‘material’ seems to depend on current scientific theories. But science is not static. Its dynamic and developing character at least occasionally forces one to add either new ontological categories or new properties to the categories already available on its list of ontologically fundamental entities. One may thus think that future science might comprise mental entities (including phenomenally conscious ones) or add mental properties to the existent categories on its list of ontologically fundamental entities. If this is to happen one day, it is not going to happen, the approach we are considering suggests, within the boundaries of standard scientific methods. It is because the current science studies only the *structure* and *function* of material entities (Chalmers 1996 and 2003). Its methods have been formulated in such a way that only the structural and functional properties of entities that are completely *non-mental* can be examined. Its current methods do not allow scientists to examine the properties of phenomenal consciousness, since the properties of phenomenal consciousness cannot be conceptualized under the standard scientific notions of “structure” and “function.” One may call this approach “non-standard scientific monism” or “flexible materialism” since the definition of ‘material’ depends on the dynamics of the current science. Nagel’s view (Nagel 1974) that in order to understand phenomenal consciousness

we need a more developed conceptual system and richer set of concepts than human beings currently possess may also be subsumed under this approach.

Materialism: Within the naturalistic approach, the most orthodox approach is to hold that phenomenal consciousness is part of nature, and that we can certainly understand and know about it by standard rational/empirical methods. Indeed, no methods that involve reference to non-natural entities—methods other than the standard philosophical/scientific ones—should be used to investigate phenomenal consciousness according to this approach. This is the canonical naturalist/materialist view, which is sometimes dubbed as "scientific naturalism." When "materialism" is defended by philosophers regarding phenomenal consciousness, their background epistemological view is something like this version of naturalism.

In addition to the term 'materialism,' philosophers of mind use interchangeably another term, 'physicalism,' though these two terms have different histories. For our purposes, we do not need to lay out their background histories, but we need to know the central assumptions behind usage of the term 'physicalism' as I will be employing it throughout the rest of this thesis. Unlike materialism, physicalism emphasizes two additional background assumptions:

(1) All physical sciences (physics, chemistry, astronomy, and Earth sciences) and biological sciences (including anatomy-physiology, zoology, botany, agriculture, and so on.) can ultimately be reduced to physics: All existent entities can ultimately be reduced to physical entities.

(2) Instead of "matter," the central notion of materialism should be "physical entity." A physical entity is what physics tells us it is: it is defined by physical theories generated by physicists within the science of physics. Hence, not only paradigmatic material things around us are physical, but also space, time, processes, states, energy, forces like gravity, etc. are physical too.

It would not be misleading to use the terms 'materialism' and 'physicalism' interchangeably to refer to the one and the same thesis as most contemporary philosophers of mind do in the literature. Those who use 'materialism' only, might prefer to attract attention to the standard materialist thesis, which, ontologically speaking, basically claims that everything that exists is material. There is nothing non-material on this thesis, where "matter" is thought to encompass the kinds of entities having standard physical properties, such as having mass, volume, momentum, energy, etc.[5] Those who employ 'physicalism' only, often have in mind a notion of "physical entity" and, by using this notion, wish to emphasize *the above two assumptions*. Regardless of this distinction, however, we may think of both materialism and physicalism as aiming at the same metaphysical doctrine that everything, including whatever is involved in any mental phenomena, is material/physical, which is the only substance of the world.

It is also possible to define 'physicalism' based on the notion of "truth"—every truth is a physical truth—instead of defining it based on what kinds of things there are. Flanagan (1992, p. 98) calls the former "linguistic physicalism" and the latter "metaphysical physicalism," and claims that linguistic physicalism is stronger than metaphysical physicalism and is less plausible. When criticizing Jackson's way of defining physicalism in terms of "physical information" or "knowledge" (1982), Lycan states that linguistic physicalism is hardly entailed by materialism about mind (2003). For Lycan, what materialism about mind asserts is simply that "human beings are made entirely of physical matter and that their properties, and facts about them, consist in arrangements of that matter" (p. 385). This, however, does not entail that every proposition about a human being must express something about physical matter. The assertion allows that there

[5] Historically speaking, George Berkeley, to give an example, defines matter as "an inert, senseless substance, in which extension, figure, and motion do actually subsist" (Berkeley 1710, pt. 9).

may be truths about human beings, such as conceptual truths, which are not about physical matter, e.g., human beings employ concepts when thinking. I think this is true in an important sense. So, I will take metaphysical physicalism to be the primary version of the doctrine unless it is necessary to refer to the notion of "the complete physical truth" employed in relation to anti-physicalist arguments in Chapter 3.

Having clarified what is meant by the terms 'materialism' and 'physicalism,' we are now ready to go into the details of the issues that are widely discussed around physicalism. In Chapter 2, I will examine those issues in connection with three metaphysical relations, which are purported to hold between the mental and the physical: identity, supervenience, and realization. The conclusion of the chapter will be that any minimalist physicalist relation connecting the mental/phenomenal to the physical has to be metaphysically necessary. In Chapter 3, the anti-physicalist arguments put forward against any minimal version of physicalism will be examined in detail, and it will be concluded that the epistemic/explanatory gap on which the anti-physicalist arguments rest is not bridgeable due to several epistemic constraints. In Chapter 4, after arguing that a posteriori necessary physicalism is the only tenable option, we will go on to examine type-B physicalism and a specific version of it called "Phenomenal Concept Strategy." Before finishing the chapter, I will introduce what I call the "problem of phenomenal subjecthood" as a severe difficulty against type-B physicalism and the above strategy. In Chapter 5, the problem of phenomenal subjecthood will be dealt with in detail in connection with several mental phenomena. And in the last chapter, we will see that the problem of phenomenal subjecthood substantially influences the metaphysical status of the phenomenal-physical gap, and because of this, type-B physicalists and proponents of phenomenal concept strategy face a severe dilemma. In the end, I will point to

the right path that should be followed in order to cope with the problem of phenomenal subjecthood.

CHAPTER 2

MINIMAL PHYSICALISM AND THREE TYPES OF RELATION: IDENTITY, SUPERVENIENCE, AND REALIZATION

Finding the place of phenomenal consciousness in nature no doubt requires a comprehensive approach to the nature of mind. As we have seen in the previous chapter, phenomenal consciousness is only one problematic issue of philosophy of mind. Other problematic issues such as intentionality, qualia, mental-physical[6] causation, and so on also threaten the naturalist/materialist approach. Arguments against physicalism not only target phenomenal consciousness but also these problematic features of mind. For this reason, in going into the details of the issues surrounding physicalism, we have to take mind as a whole. Only after a full examination of the core idea of the physicalist thesis, we can turn back to the nature of phenomenal consciousness and try to see whether we can correctly place it in nature within a physicalist framework.

Apart from the assumptions (1) and (2), considered at the end of the previous chapter, we have to have a better expression of the core idea of physicalism. The basic expression that *everything that exists is physical* is rather general and needs to be clarified with respect to several issues that have been confusing minds for a long time. Here, questions like "What is the extension of 'everything'?" or "What does 'physical' exactly mean?" i.e., "How do we determine the correct application of the predicate 'is physical'?" are focused on as general concerns for physicalism.

[6] It is very common to use the adjectival term 'psychophysical' to describe the metaphysical relations between the mental and the physical. I will not prefer to use this term, but will instead use either 'mental-physical' or 'phenomenal-physical' throughout the remaining chapters. The only reason for this preference is that I see a big semantic difference between what is regarded as psychological or behavioral and what is considered mental or phenomenal.

What we will be primarily interested in are not these questions, but rather the questions centered on the mental-physical connection, such as “What is the nature of the connection between the mental and the physical?” or “Can it be known a priori or a posteriori?” or “Is it a necessary connection or a contingent one?”. To answer such questions, a lot of different proposals are put forward in the literature. And these proposals culminated in different views adopted by philosophers of mind regarding the mental-physical connection. It seems, nevertheless, that all those different views concerning this connection can be reduced to three metaphysical relations: (1) the identity relation, (2) the supervenience relation and (3) the realization relation. Our aim in this chapter will be to specify a minimal version of physicalism with regard to these three kinds of relations—a minimal version that should be agreed on by almost all physicalists. The boundaries of this minimal version will also indicate the borderline beyond which one would not be called a physicalist anymore. We need such a minimalist characterization of physicalism since the target of the anti-physicalist arguments that we will deal with in the following chapter should be as clear as this minimalist version, so that a physicalist can defend herself on a solid ground. To this end, we need first a brief clarification of the distinction between monism and dualism. And then we can examine the three different metaphysical relations while attempting to reach that minimal version of physicalism. Afterwards we will be ready to analyze the anti-physicalist arguments in the next chapter, that are claimed to threaten even any minimal version of physicalism.

2.1. Monism vs. Dualism

In the ontological sense, “monism” roughly means that *there is only one substance, and all things that exist are forms of this single substance*. Physicalism is a monist doctrine, since it entails that everything is a form of physical substance. There can be other monist conceptions of nature, like idealism

claiming that *only ideas exist*, or neutral monism claiming that *both the physical and the mental are forms of one single neutral substance, the nature of which we may or may not know*. Dualism, on the other hand, simply claims that *the world consists of two fundamentally different substances*, and these two substances are *mind and matter*. Mental entities (mental states, events, properties, etc.) are fundamentally different from material/physical entities, according to dualism. They are fundamentally different because their properties are different from those of physical entities, and according to Leibniz's Law of Identity, two things are distinct if they do not simultaneously share the same exact properties.

Dualism has been defended since Plato, but it is actually Descartes who formulated the view in some detail with regard to the mind-body relations (Descartes 1637 and 1941). Cartesian philosophers later developed different versions of it, such as substance dualism—*mind and matter, particularly body, are different substances* (within which we have interactionism—*mind and body causally interact with each other* and epiphenomenalism—*mental events, while being the effects of bodily events, are causally inefficacious*) and property dualism—*mental properties of a physical substance are fundamentally different from the material/physical properties of that substance*. The materialist/physicalist doctrine is in fact a monist reaction to all versions of Cartesian dualism. Among those reactions that can be cited under the materialist/physicalist doctrine are behaviorism, which holds that mental states are behavioral states; the identity theory, which holds that mental states are identical to material/physical states; functionalism, which holds that mental states are functional states; and eliminativism, which holds that only brain states (physical or functional) exist, and that there exists no mental state.

The monism-dualism distinction is crucial for any version of physicalism, since anti-physicalist arguments aiming to undermine the physicalist thesis base their

line of thinking on this distinction. For example, if an anti-physicalist argument, succeeds in showing that a part of mental reality cannot be solely explained in physical terms, then the argument will challenge the physicalist to explain this dual character. For a physicalist, then, the precise line between monism and dualism is quite significant. The minimalist interpretation of physicalism is, indeed, aimed to draw this exact line between the two.

2.2. The Identity Relation

The most influential (and commonly argued) kind of relation that is purported to hold between the mental and the physical is the identity relation, according to which any mental item is *identical* to a physical item. The three of four physicalist reactions to dualism (behaviorism, the identity theory, and functionalism), which we have mentioned in the previous section, in fact, employ the identity relation. What is identified with mental items, however, changes: it can be behavioral or functional or blood-and-flesh material items depending on the motives and goals of the theories. Here, our plan is to specify a definition of the identity relation first, and then go on to examine briefly some specific issues around several versions of physicalism employing the identity relation. Consider the following definition of the identity relation:

> Definition of Identity: *A property / object / state / event x is identical to a property / object / state / event y, respectively, if and only if x and y are the one and the same property / object / state / event, respectively.*

This definition of identity is of numerical identity—there is not two but only one entity—rather than of qualitative identity—two distinct entities instantiate the same types of properties. What this definition actually tells us is that an entity can have the true identity relation only to itself, and nothing else.

Here, one point we should not miss is the fact that theories employing the identity relation are in fact reductionist theories. When you identify mental items with physical ones, you actually reduce mental items to physical ones. It is a good start, then, to begin with the reductive approach and look at how this identity relation is employed within the approach.

2.2.1. Reductive Physicalism Using the Identity Relation

How are physicalists supposed to show that mental entities are physical? There should be at least a possible way of proving or showing this. The common way of showing this is to make ontological reduction. One may reduce mental entities to physical ones ontologically by "reducing mental terms to physical ones." One may avoid ontological reduction, and reduce only mental terms to physical ones while being neutral in the ontological sense. Reduction in either sense has been, for a long time, thought to be the only way of building a physicalist thesis. Smart, for example, defended, as a physicalist, the view that *mental terms can be reduced to topic-neutral terms that are ontologically neutral on what these mental terms refer to* (Smart 1959). These terms neither presume the existence of fundamentally different mental items nor presuppose the identity of mental properties to physical ones, and, as a result, there is still room for physicalism. But here, one may rightly question the notion of reduction, and require further elucidation of it.

There are several ideas to elucidate the notion of reduction in philosophy of mind. One idea is to think in terms of "conceptual analysis," within which the conceptual content of a given mental term is analyzed in physical terms. Another idea is to construe reduction as an "inter-theoretic reduction," which is done between the theory of folk psychology and another theory containing no mental terms, such as the neuroscientific theory. One other idea is to derive mental statements from non-mental ones a priori, so that the contents of mental

statements can be reduced to the content of physical statements. One final idea is to associate mental predicates with non-mental ones implying that they both refer to the one and the same entity. The thing is that all these ideas amount to one single notion: "property identification." It is the core idea of reduction that one identifies the properties referred to by mental terms with the ones referred to by non-mental terms. The difference between these approaches results from the way one does the property identification. In conceptual analysis, one identifies mental properties with non-mental ones based on conceptual contents. In inter-theoretic reduction, one provides bridge laws that state identity of properties based on the elements of the theories. In a priori derivation, one does the same job claiming that the relevant identity statements are known a priori. In predicate association, one cannot associate mental predicates with non-mental ones without identifying properties referred separately by these mental and non-mental terms.

To illustrate, consider the conceptual analysis approach. Reducing mental terms to physical ones, for this approach, means that one analyzes the conceptual content of mental terms in physical terms. Take the mental term 'itch,' for instance. When you use this term in an introspective report like "My right leg itches," you actually mean something like "My right leg is exposed to the neurological event X" upon conceptually analyzing the content of <itchiness>.[7] Here, you identify the property referred to by the term "itchiness" with the property referred to by the term "being exposed to neurological event X." By doing this, you actually reduce the mental term "itchiness" to the non-mental (physical) term "being exposed to neurological event X."

The same property identification goes for inter-theoretic reduction, a priori derivation and predicate association. The question is whether or not a physicalist

[7] I will refer to concepts and phenomenal qualities—understood completely as mental entities—by expressions in between the arrowheads, "<...>."

should be committed to reductionism in this sense of property identification. Can physicalism embrace a non-reductive approach to mental reality, an approach that involves no property identification and admits irreducibility of mental terms to physical ones? Ontologically speaking, the answer should be "no" since physicalism is ontologically a monist view, which holds that he world contains nothing but physical entities. If one embraces properties ontologically irreducible to physical properties, one will no longer be a monist physicalist. But, metaphysically speaking, there are non-monist physicalist views too regarding mental reality. These views adopt non-reductive approaches according to which mental states/events/properties cannot be reduced to physical states/events/properties while they can still be subsumed under the physicalist doctrine. Supervenience physicalism and the view of physicalism employing the realization relation, for instance, both of which we will examine later in this chapter, is perhaps the most popular of these non-reductive physicalist approaches involving no property identification. The result is that reductive physicalism is not the best candidate for providing the core commitment of physicalism. Let us go into the details of some specific issues and distinctions that occupy minds regarding the reductive approach.

2.2.2. Type vs. Token Identity

Suppose I am having pain in my left arm at the moment. This particular pain is a token pain. When I say "Pain (or more correctly having pain) is a mental phenomenon," on the other hand, what I mean by "pain" is a type. It does not refer to a particular pain. It refers to all instances of pain phenomena. Now, the distinction between type vs. token physicalism obtains its content from the two usages above.

Type physicalism claims that mental types (properties, types of states, types of events, etc.) are identical to physical types (properties, types of states, types of

events, etc.). We can think of type physicalism as what the following formulation says:

> (1) For each mental type *M*, there is a physical type *P* such that *M* is identical to *P*. (i.e. for each person *x*, *x* has/undergoes *M* if and only if *x* has/undergoes *P*.)

Defenders of the identity theory typically hold type physicalism. We may mention the pioneers of the identity theory here, such as Place (1956), Feigl (1958), and Smart (1959). According to the identity theory, the pain I am experiencing currently in my left arm is an instance of a type (in other words, my pain experience is a member of the set that comprises all pain instances in the world), and this type is identical to a certain type of physical phenomenon in my brain, say c-fiber firing.

Token physicalism, on the other hand, does not go this far and claims that only current instances of mental phenomena (not properties, but particular mental states and events) are identical to current instances of physical phenomena (not properties, but particular physical states and events) in brains. This doctrine can be formulated as follows:

> (2) For each mental particular *m*, there is a physical particular *p* such that *m* is identical to *p*. (i.e. for each person *x*, *x* undergoes *m* if and only if *x* undergoes *p*.)

This idea has some advantage over type physicalism against arguments such as multiple realizability argument, according to which mental types can multiply be realized: while the same type of pain is realized by the physical process *X* in me, it may be realized by the physical process *Y* in someone else, or it may be realized

by some other kind of, say silicon-based, process in an extraterrestrial being (Putnam 1967). This argument was one of the chief motives behind functionalism (historically the next step after the identity theory) while it does no harm to token physicalism. One concern here is whether or not all physicalists should be at least committed to token physicalism. The answer depends on what kind of relation a physicalist adopts between the mental and the physical. If she adopts the identity relation between the two, then token physicalism seems to be the most basic ground position she can take. However, if she adopts one of the other kinds of relation, such as the supervenience and the realization relation, then she does not have to be committed to token physicalism at all. In other words, it is possible (though this is not uncontroversial) for a physicalist to deny any identification of mental entities (including particular states and events) with physical ones, while still remaining a physicalist.

2.2.3. A Priori vs. A Posteriori Physicalism

Another issue concerns the distinction between a priori and a posteriori approach to the physicalist thesis employing the identity relation. To get a better grip on the issue let us categorize for the moment all *true* statements into three kinds from the physicalist perspective: [8]

(p) Physical-to-physical statements: "Most tables are made of wooden."

(m) Mental-to-mental statements: "Sorrow is similar to sadness."

(mp) Mental-to-physical (or physical-to-mental) statements (can be called "bridge statements" as well): "Pain is identical to c-fiber firing."

Let *N* be the conjunction of *all* truths of the world, and *P* be the conjunction of all *physical* truths of the world. Now, *N* would comprise all kinds of true statements, including (p), (m) and (mp)-statements. These statements could be deduced easily from *N*, because they are the conjuncts of *N*. Now, the physicalist claim is

[8] There are other kinds of true statements such as abstract truths, mathematical truths, and indexical truths. For simplicity, let us either disregard them or put them all into (a), the class of physical-to-physical statements.

simply that *P* is equivalent to *N*; hence one must be able to deduce these three kinds of statement from *P*. In other words,

(3) *P* entails *N*.

To put it differently, the totality of physical truths is all there is to know. If one knows all physical truths there are, there is nothing else left for him to know. Here, the subsequent issue that should be considered is this. Let *Q* be the conjunction of all mental truths, including only statements of type (m), mental-to-mental statements. Assuming that (3) is true, and that *Q* is a *subset* of the conjuncts of *N*, the following physicalist claim must also be true:

(4a) *P* entails *Q*.

This means *Q* must also be deducible from *P*. Here, the key point is the derivability of (mp) statements from *P*. For, if those bridge statements asserting the mental-physical connection by means of an identity relation can be deduced from *P*, deriving (m) statements from *P* will be no problem at all. The question is how the physicalist will derive (mp)-statements from *P* (and thus (m)-statements and *Q* altogether), a priori or a posteriori? To state the matter differently, can the following material conditional be known by reason alone (a priori) or by also appealing to empirical data (a posteriori)?

(4b) If *P* then *Q*.

No doubt *P* contains no mental language. Nevertheless, *Q* (which contains mental language, viz. mental-to-mental statements) must somehow be deducible from *P*. At this point, a priori physicalists claim that *Q* follows from *P* a priori—(4b) is an a priori statement—without need to resort to any empirical information. A posteriori physicalists, on the other hand, maintain that *Q* does not follow from *P* a priori. Empirical investigation is needed, and hence it follows from *P* only a posteriori, that is to say, (4b) is an a posteriori statement. From this picture, one might be inclined to conclude that a posteriori physicalism requires less: It seems

committed to less, so it is more minimal than a priori physicalism. But there is another related issue that is often discussed together with this one, namely the necessity vs. contingency of the physicalist thesis. A brief consideration of the modal status of the physicalist thesis would reveal the fact that a posteriori physicalism employing the identity relation has a severe problem, which leads us to consider it more carefully. The problem concerns the nature of identity statements hosting two rigid designators that flank the identity sign. Let us look at the problem in a bit more detail.

2.2.4. The Problem of Modal Status: The Necessity vs. Contingency of the Physicalist Thesis

Given that all a priori truths have traditionally been considered to be necessary and all a posteriori truths to be contingent, for an a priori physicalist the following statement must be both a priori and necessary:

(4b) If *P* then *Q*.

As we have said earlier, *P* expresses the conjunctive statement of all physical truths, and *Q* expresses the conjunctive statement of all mental truths. To simplify the issue and give a concrete example to think about, consider a narrower instance of (4b):

(5) If a person *x* undergoes c-fiber firing, then *x* is in pain.

Now, the a priori physicalist maintains that (5) is knowable by a priori reasoning, and is thus a necessary truth. For an a posteriori physicalist, however, (5) is knowable a posteriori and must be contingent. A complicated modal problem arises for the a posteriori physicalist just at this point. Think about (5) again. The alleged truth of (5) comes from a background identity statement: “Pain is identical to c-fiber firing.” For the a posteriori physicalist, (5) is contingent because this identity statement is contingent. After Kripke’s work, however, most philosophers began to interpret the modal picture differently. Many philosophers are convinced that some a posteriori truths (in particular, (mp)-statements

expressing mental-physical identities like the above one) should not be contingent but necessary. Such a conviction emerged because Kripke persuaded them that identity statements accommodating two rigid designators[9] flanking the identity sign must be necessarily true if they are true at all (1971 and 1980, pp. 140-53). Consider the following two statements, the first of which is scientifically constructed, and the second of which is an (mp)-statement:

(6) Water = H_2O

(7) Pain = c-fiber firing.

Based on the externalist semantics of "rigid designators," which was previously formulated by Putnam (1975), Kripke argued that both (6) and (7) are necessarily true if they are true at all, because the two sides of the identity sign in both (6) and (7) comprise terms that refer to the same entities in all possible worlds (1971 and 1980, pp. 153-55). The problem arose not for the a priori, but for the a posteriori physicalist. For the orthodox view had been that an a posteriori statement is a contingent truth if it expresses a truth at all. But after Kripke, a posteriori physicalists had no option but claim that mental-physical identity statements are necessarily true, just as standard scientific reductions in science are. And this led to other serious problems that we will examine in detail in Chapter 4.

At the end of this section, let me draw attention to the three versions of *identity* physicalism we have examined so far: (i) a priori necessary physicalism, (ii) a posteriori contingent physicalism, and (iii) a posteriori necessary physicalism. Version (i) is still an option to take for the a priori physicalist. Version (ii), however, has been severely undermined by the Putnam-Kripke considerations we have summarized above. And version (iii), buttressed by the Kripkean view, has been receiving the greatest attention recently, as it is sometimes considered the most promising version of physicalism (Chalmers 2003 and 2007). The first two

[9] For Kripke, rigid designators are the terms that refer to the same entities in all possible worlds. Some commonly accepted rigid designators are natural kind terms and proper names.

versions, (i) and (ii), are not going to be investigated further in what follows. Version (iii), however, will be the center of our attention in the following chapters. But first let us discuss two other types of metaphysical relations employed by physicalists.

2.3. The Supervenience Relation

In the previous section, we have examined several applications of how the identity relation is used to present the core thesis of physicalism. It seems that each version of the reductive approach is aimed at ontological reduction, and is thus bound with property identification (identifying mental properties with physical properties). While type physicalism suggests property identification as well, token physicalism has a different position with respect to using the identity relation. But at the end, it too suggests the identity of particular mental tokens with particular physical tokens (not properties, but particular token states and events). We have also given a brief consideration to a priori and a posteriori derivability and the modal status of physicalism, which employs the identity relation.

The Identity relation is, however, not the only relation a physicalist can employ in formulating her thesis. In fact, the identity relation is considered quite problematic by a considerable number of philosophers. We will not examine the problems of the identity relation here because of the scope of this work, but analyze one of the important problems in the next chapter. As we indicated at the beginning of this chapter, apart from the identity relation two other relations are invoked in the literature to present the core thesis of physicalism that everything that exists is physical. While versions of physicalism employing the identity relation are reductive, versions using the supervenience and realization relations are non-reductive. Thus, the supervenience physicalism is a non-reductive approach to the problem of mental-physical connection. Consider

below the nature of this supervenience relation in two different formulations using two different concepts in each:

Definition A: Supervenience (using the notion of identity): *X-kind properties supervene on Y-kind properties if and only if, for any two objects a and b, if a and b have identical Y-kind properties (share all Y-kind properties), then they also have identical X-kind properties (share X-kind properties).*

Definition B: Supervenience (using the notion of difference): *X-kind properties supervene on Y-kind properties if and only if any two objects a and b cannot differ in their X-kind properties without also differing in their Y-kind properties.*

These two different formulations are logically equivalent actually. Definition A uses the notion of identity while definition B uses the notion of difference, but what they tell us is the same. Based on these formulations, let us see what supervenience physicalism claims.

2.3.1. Supervenience Physicalism

The notion of "supervenience" has been introduced into philosophy of mind by Davidson (1970) in these words: "...[M]ental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect" (p. 214). After him, many philosophers such as Horgan (1982 and 1993) and Kim (1993) have gone into details of the notion of supervenience. The reason why <supervenience> has received a lot of attention is that it does not require property identification as reductive physicalism does.

Based on the above definitions, supervenience physicalism simply asserts that:

(1a) (Employing Definition A) *All properties supervene on physical properties: if any two worlds have identical physical properties, i.e. share all their physical properties, they have all identical properties, i.e. share all properties.* (For any world w, any *physical* duplicate of w is a duplicate of w simpliciter.)

(1b) Employing Definition B) *All properties supervene on physical properties: if any two worlds differ in some properties, they also differ in their physical properties.* (Any two worlds cannot differ in some properties without also differing in their physical properties.)

The above claims are the general claims of supervenience physicalism. With respect to the mental-physical relation in particular, we can deduce from (1a) and (1b) the following claims:

(2a) (Employing Definition A) *Mental properties supervene on physical properties: if any two persons have identical physical properties, i.e. share all their physical properties, then they also have identical mental properties, i.e. share all their mental properties.*

(2b) (Employing Definition B): *Mental properties supervene on physical properties: if any two persons differ in their mental properties, they also differ in their physical properties.* (Any two persons cannot differ in their mental properties without also differing in their physical properties.)

It is clear from these definitions that supervenience physicalism is not committed to property identification. In fact, it is not committed to any kind of identification.

So, it is a non-reductive approach. We can therefore say that it is more minimal than any version of reductive physicalism. But one important issue is the kind of necessity involved in the above definitions. Do the conditionals in (1a) and (1b) and (2a) and (2b) hold with *metaphysical necessity* or *nomological necessity*? In other words, how do we interpret the term "cannot" in (1b) and (2b)?

Supervenience of mental properties on physical properties *with nomological necessity* is weaker and implies only nomic dependence. In other words, mental properties supervene on physical properties because the laws of nature in the actual world necessitate so. Hence, zombies (hypothetical beings that are physically identical to human beings but have no supervening mental property at all) are nomologically impossible. Such beings would violate the laws of nature. But keep in mind that the notion of "nomological necessity" is silent on those alleged possible worlds that have different laws of nature and accommodate zombies. It seems that while nomological necessity is a necessity, it is nevertheless a contingency in terms of modality: mental properties supervene on physical properties with nomological necessity, but only *contingently* as far as modality is concerned.

Supervenience of mental properties on physical properties *with metaphysical necessity* is, on the other hand, stronger and indicates strict ontological dependence: there can be no possible world that contains the same exact physical properties as our world does, but does not contain one or more of the mental properties of our world. This means mental properties in our world *necessarily* supervene on the physical properties in it. One cannot enjoy physical properties without also enjoying the relevant supervening mental properties in any possible world (but not vice versa according to the definition of supervenience). This kind of supervenience clearly does not allow zombies in any possible world.

There are two problems for the former notion of supervenience. The first one is that almost everyone, even substance dualists, grants that a kind of supervenience relation holds with nomological necessity between mental and physical properties. In other words, no one denies the idea that the existence of zombies is impossible in the actual world given the constant and rigid character of laws of nature. Besides, contingent supervenience allows for the possibility that there can be worlds in which mental properties supervene on nothing but exist on their own, and the possibility that there are zombies that instantiate no mental property at all. If so, supervenience with nomological necessity cannot distinguish a physicalist from her opponent as it is compatible with dualism. It seems that such a physicalist position does not save the core spirit of the physicalist thesis, and hence cannot be counted a genuine physicalist position. Perhaps just because of this reason some believe that supervenience physicalism must at least embrace token physicalism within itself in order to be counted a minimal version of physicalism.

Second, recall the Kripkean view (1971 and 1980, pp. 140-53), which we have mentioned earlier in Section 2.2, regarding the modal status of identity physicalism that identity statements containing rigid designators on the two sides of the identity sign are necessarily true if they are true at all. This includes identity statements that express mental-physical connections as well—(mp)-statements that we also mentioned in Section 2.2. Consider the following two statements:

(3) Pain = c-fiber firing

(4) Pain supervenes on c-fiber firing.

Now, claiming that (3) cannot be true contingently, while (4) is true contingently—if it is true at all—poses a problem. Notice that both relations are flanked by the same rigid designators that are strictly tied with their referents in

all possible worlds. So, it is reasonable to think that any world containing pain and c-fiber firings will also entertain all the metaphysical relations that hold between these two natural kinds in the actual world. Yes, the identity in (3) is a numerical identity—i.e. there exists not two, but only one entity involved—whereas the supervenience relation in (4) is supposed to hold between two different properties. But pain and c-fiber firing are natural kinds whose intrinsic properties are fixed in all possible worlds. So, conceiving a world that contains pain and c-fiber firing, which instantiate necessarily the same sets of intrinsic properties forces one to also embrace the idea that if they are metaphysically related to each other in some way in the actual world resulting from their intrinsic nature, then they must also be connected with the same metaphysical relations between each other in other possible worlds.

For the reasons given above, we will only keep on the table the version of supervenience physicalism that claims that supervenience holds with metaphysical necessity between the mental and the physical. Besides, when it comes to supervenience physicalism, the anti-physicalist arguments we will examine later can be better understood keeping this version in mind.

2.4. The Realization Relation

Supervenience physicalism has received a lot of criticism while it has attracted much attention. Among the criticisms, the crucial one is that it allows for the existence of two ontologically distinct properties leading philosophers to consider it as a view compatible with property dualism. This does not change, even if the supervening mental properties are metaphysically necessitated by the base physical properties. Some, therefore, began to look for a new type of metaphysical relation to connect the mental to the physical. Such a relation should be stronger than supervenience, but weaker than identity, so that it can neither face the threat of falling into property dualism, nor of being exposed to

the problems that the identity physicalism (particularly type-identity physicalism) is entangled in. These are the motives of philosophers in proposing the realization relation. It has not been employed in the literature as much as the identity and supervenience relations have. It is a quite newly introduced type of relation. Three names in relation to it should be especially mentioned here: Levine (2001), Melnyk (2003), and Shoemaker (2007). Let us briefly look at their proposals regarding the nature of the realization relation and consider the versions of realization physicalism they each suggest.

Levine is primarily concerned with property realization since he considers property dualism to be more compelling than other theories like substance dualism. He actually does not offer a complete analysis of the realization relation. Nevertheless, after stating the following proposal of realization physicalism, he attempts to distinguish the realization relation from other types of metaphysical relation holding among properties:

> (1) Only the fundamental properties of physics are instantiated in a basic way; all others, particularly mental properties, are instantiated by being realized by the instantiation of other properties. (2001, p. 12)

Because the definition of 'physical' is problematic, after giving a bit of consideration to whether we should define it or not, he revises (1), and proposes the following:

> (2) Only non-mental properties are instantiated in a basic way; all mental properties are instantiated by being realized by the instantiation of other, non-mental properties. (p. 21)

As to the nature the realization relation, Levine thinks that it is distinct from these three metaphysical relations: accidental correlation, nomological/causal

connection, and identity. Accidental correlation obtains between two properties when instantiations of one property co-occur with instantiations of the other one. Nomological/causal connection is just a causal relation holding between two entities, and happens when there is a law enforcing the connection. Identity is a relation an entity can have only to itself and nothing else. The realization relation, for Levine, is fundamentally distinct from the first two in that the first two imply no intimate ontological relation. The relata are not ontologically tied together; they are ontologically independent from each other. The realization relation is fundamentally distinct from identity in that the identity relation is ontologically more intimate than the realization relation. Furthermore, the realization relation is not like nomological/causal connection, because in the nomological case, there is not any constitutive relation between the parts as there is in the realization case. It is not like identity, because when two things are numerically identical, there is not two but only one thing. The reason why we talk about two things is a linguistic one: there are two *terms* referring to the one and the same thing. In the realization case, however, there are still two *things* related to each other by the constitutive realization relation that sustains a strong ontological dependence.

Unlike Levine, Melnyk attempts to give a precise definition of the realization relation:

> (3) A token x of a functional type, F, is *physically realized* iff (i) x is realized by a token of some physical type, T, *and* (ii) T meets the associated condition for F solely as a logical consequence of the distribution in the world of physical tokens and the holding of physical laws. (2003, p. 23, italics in original.)

Here, (3) is in fact a definition of the term "physically realized." The term "being realized" should be understood in this way: A token of functional type F is realized by a token of some or other type G if and only if F is tokened just in case a token of G meets an associated condition (whatever that may be) for F (p. 21).

Based on these definitions, Melnyk announces his version of realization physicalism as follows:

> (4) Every property instance is *either* an instance of a physical property *or* a physically realized instance of some functional property, every object is *either* an object of some physical object kind *or* a physically realized object of some functional object kind; every event is *either* an event of some physical event kind *or* a physically realized event of some functional event kind. (p. 26, italics in original.)

What we should understand from all these is that the actual world contains, for Melnyk, two types of properties: physical properties and physically realized properties. The latter properties are physically realized by means of the former's meeting a variety of associated conditions required for the existence of the latter. Here, the problem is the ontological status of physically realized properties. They are physically realized, but they are not physical. So, what are they ontologically? In both (3) and (4), by invoking the notion of "functional type" Melnyk clearly states that they are functional properties. But then his realization physicalism would be the realization version of a functionalist theory. He only seems to give a theory of how functional properties are produced by their physical bases. He does not provide a theory of how mental properties are functional properties, which should be done anyway by a functionalist.

Perhaps what Melnyk means by a "functional type" is a higher-level property. If so, his formulation of realization physicalism should be read this way: *Every property instance is either an instance of a lower-level physical property or an instance of a higher-level property which is physically realized by an instance of a lower-level physical property*. Even in this interpretation, one may still question the ontological status of higher-level properties—whether they are physical, or some sort of non-physical properties. If these higher-level properties are physical,

then Melnyk's physicalism collapses into identity physicalism. If they are non-physical, then Melnyk's theory is exposed to the problems supervenience physicalism faces regarding its compatibility with property dualism, which we will deal with in a moment.

While he allows other sorts of realization, Shoemaker too directs his attention to property realization. He thinks that realization physicalism is better than supervenience physicalism because of its several advantages. Consider what he says regarding the nature of the realization relation:

> (5) In general, X realizes Y just in case the existence of X is constitutively sufficient for the existence of Y—just in case Y's existence is "nothing over and above" X's existence. (2007, p. 4)

Applied to property realization, (5) will read as follows:

> (6) A property instance F is realized by another property instance G if and only if the instance of G is *constitutively* sufficient for the existence of the instance of F, i.e., the instance of F is nothing over and above the instance of G.

Construing it in this way, Shoemaker thinks that one of the advantages of the realization relation over the supervenience relation is that the former is not compatible with property dualism, whereas the latter is. If he is right, it would be a big advantage not to be compatible with property dualism, since it is the core spirit of physicalism to reject any kind of dualism concerning the metaphysical nature of mental entities (states, events or properties). In fact, any physicalist theory that leads to the existence of two ontologically distinct kinds of property will no doubt suffer from falling into property dualism. Identity physicalism saves

itself from such a threat because of the nature of the identity relation, but supervenience and realization physicalism need extra help to be saved from falling into property dualism. Whether the notion of "metaphysical necessity" provides the needed help is a question that should be considered together with the issue of modal status.

Hence, what should really concern us here regarding the above views is the contingency or necessity of the realization relation, i.e. its modal status. They actually differ on the matter. Levine thinks that the realization relation involves metaphysical necessitation, but only in one direction, *bottom-up*, as opposed to the identity relation that involves metaphysical necessitation in both directions. Accordingly, for Levine, the higher-level property is realized by the lover-level property with metaphysical necessitation from the lower-level to the higher-level: once, the realizing property is instantiated, the realized property is *necessarily* instantiated, but not vice versa. Melnyk, on the other hand, strictly announces that the realization relation he offers holds only *contingently* between the higher-level and lower-level properties, namely, between mental and physical properties. Accordingly, his realization physicalism is a contingent thesis. In other words, mental properties are physically realized by physical properties in the actual world without any metaphysical necessitation. There are possible worlds where mental properties are not realized by the relevant physical properties. Shoemaker considers the issue in a way similar to the way Levine does, and stresses the notion of constitution that has modal consequences. That the instance of the realizing property is constitutively sufficient for the existence of the instance of the realized property clearly implies that, once the realizing property is instantiated, the instance of the realized property *necessarily* comes into existence.

It seems that Melnyk's contingent (and thus a posteriori) realization physicalism does not preserve the core spirit of physicalism, since it makes zombies possible. If a theory of mental-physical connection embraces the possible existence of zombies, it means that the mental does not *necessarily* need a physical basis, or a functional one, or even any basis at all, which would ultimately bestow an independent ontological status to mental entities, if not in this actual world, at least in some possible worlds. And this yields a more threatening compatibility with property dualism than the version with metaphysical necessity. For this reason, we will just assume that the version of realization physicalism with metaphysical necessity, which Levine and Shoemaker seem to support, is more tenable than any contingent version, and deserves to be considered the minimal commitment of realization physicalism.

For those who need further reasoning to be convinced about the unacceptability of a contingent relation between the mental and the physical, let us think about the issue step by step. We are at least so far sure on the basis of Kripkean considerations that any purported identity relation between the mental and the physical must be a necessary relation. As for the supervenience and realization relations, consider the following statements:

(7) A distinct mental property *m* supervenes on a physical property *p* contingently.

(8) A distinct mental property *m* is realized by a physical property *p* contingently.

It is obvious that both of these statements are totally compatible with property dualism, since we are talking about exactly *two* distinct properties here, and they are only *contingently* related on the basis of *nomological* laws, say some causal laws. To avoid such a conclusion, defenders of supervenience and realization relations either must deny the reality of mental life altogether, and become eliminativists, or remove the contingency from the relations expressed above as a

first step. Note that they cannot preserve contingency by claiming that there is only one property involved here, because if there is just one property, then the relation obtains between two different terms for the same referent, which ultimately entails that the relation is an identity relation, and thus a necessary one. So, removing contingency means the relation holds with metaphysical necessitation, and the statements become like this:

(9) A distinct mental property *m* supervenes on a physical property *p* necessarily.

(10) A distinct mental property *m* is realized by a physical property *p* necessarily.

Now, as a second step, a new question arises: Are these last formulations of the mental-physical relation compatible with the core spirit of physicalism? I don't think so. We are still talking about *distinct* properties. A property dualist may still declare that she is completely comfortable with mental properties being necessitated by physical properties in such a way that they supervene on or are realized by physical properties in *every* possible world. As long as they are distinct from physical properties, and thus are not on the list of ultimate ontological categories of physical sciences, she will happily grant such a conclusion. It is this reason why some (including myself) tend to think that the supervenience and realization relations have to be bolstered ultimately with an identity relation of some sort (Kim 1998; and Levine and Trogdon 2009).

Furthermore, I think we have a separate reason to believe that the versions of physicalism based on these two relations must involve a hidden identity claim within the theory. Let us think about the "necessity" involved here a bit more. How should we construe "*p* necessitates *m*" in the first place? The most plausible interpretation of necessitation here is that wherever *p* exists, *m* also exists with the relevant purported relation to *p*. This is clearly a "metaphysical

necessitation." But for a metaphysical necessitation to be truly a necessitation, it has to be reduced to a conceptual necessitation. And a conceptual necessity ultimately is rooted in a logical necessity that can be established in the mental-physical case only by an identity relation. I am aware that this diagnosis requires a wider discussion of the issue, but for our present purposes I will just leave it at this point. [10]

Having come to the end of this chapter, we can now say that the three relations—identity, supervenience and realization—postulated by the physicalist have to involve *metaphysical necessity* in order to be considered the minimal commitment of the corresponding version of physicalism. At the end of each section, we have given separate reasons why a contingent relation between the mental and the physical cannot meet the minimal requirement of physicalism in the modal sense, but surely there may be other reasons to reject contingency of these three relations, such as those given in a recent paper by Levine and Trogdon (2009). Furthermore—although it is not strictly required for the purpose of this chapter— physicalist theories based on the last two relations have to involve an identity thesis within the theory in order be fully saved from being compatible with a version of property dualism. Notice that even Melnyk's theory of realization involves the identification of the mental with the functional. In the next chapter, then, it should be kept in mind that the anti-physicalist arguments are claimed to run against not contingent versions of physicalism, but rather against all metaphysically necessary mental-physical relations.

[10] One might here consider four types of necessities: "nomological," "conceptual," "metaphysical," and "logical." In my view, nomological necessity is not a real modal term. It expresses the lawful connection between two events but *only* in the *actual* world. Conceptual necessity and metaphysical necessity, on the other hand, are modal terms and utilize *indirect* ways of expressing logical necessity. They both root in logical necessity. The chief difference is that the former is more directly connected to the logical necessity than the latter, and the latter is ultimately reduced to the former. Conceptual necessity arises from logical relations between the concepts (or semantics of the terms), and metaphysical necessity ultimately results from conceptual necessity.

CHAPTER 3

ANTI-PHYSICALIST ARGUMENTS AND THE GAP BETWEEN THE PHENOMENAL AND THE PHYSICAL

In the preceding chapter, we have attempted to put on board the minimal version of physicalism in relation with the three kinds of metaphysical relation: the identity, supervenience and realization relations. Our purpose was to see the core commitment of physicalism clearly. Having seen that, we can now hope to evaluate correctly the arguments that have been put forward against physicalism. Examining the anti-physicalist arguments will enable us first to get a grip on the recalcitrant nature of mind (and phenomenality) against the materialist conception of the world, and second to see the source of the problem in integrating phenomenal consciousness into our understanding of nature. But we will not target the core problematic feature—subjectivity—of phenomenal consciousness until the next and succeeding chapters.

3.1. Arguments against Physicalism

Historically speaking, arguments against materialism even go back to ancient times, but more modern ones constitute the real concern for physicalism. Hence, we will only examine the modern arguments against physicalism, particularly the contemporary ones. It is worth mentioning first the Cartesian arguments against materialism.

Descartes' arguments for the mind-body duality get their force from Leibniz's law of indiscernibility of identicals, namely that if two things are identical, then they share (instantiate) all the same properties. The logical—contrapositive—

structure of these anti-materialist arguments is quite simple: *X* and *Y* do not instantiate all the same properties; therefore, by Leibniz's law, they are not the same thing. Because Descartes examines a good deal of distinct features that mind and body have, a number of anti-materialist arguments can be constructed out of his works, particularly from *Meditations* (1641). To give some examples, consider the following arguments. The doubt argument—I can doubt that my body exists. I cannot doubt that I, a thinking substance, exist; therefore, I am not identical with my body. The divisibility argument—my body is divisible. I am indivisible; therefore, I am not my body. The argument from extension—my body has a spatial location. My mind does not have a spatial location; therefore, my mind is not my body. The argument from introspection—I can come to know about my mind (mental states) by introspection. I cannot come to know about my brain (or any physical states) by introspection; therefore, my mind and my physical parts are distinct.

What we are primarily interested in are the contemporary arguments against physicalism, rather than Descartes' arguments for the mind-body duality. The reason is a minor one actually, but plausible enough to determine the strategy. It can be said that Descartes' anti-materialist arguments do not specifically target the entire materialist thesis. Contemporary arguments against physicalism, on the other hand, specifically aim to undermine the entire physicalist thesis. Besides, contemporary arguments have multiple aspects. And if a physicalist defeats the contemporary ones, she can easily defeat those given by Descartes as well. Let us briefly look at some of those influential contemporary anti-physicalist arguments now, which have been debated in the literature.

3.2. Contemporary Arguments against Physicalism

It will be useful to present the arguments first, and then evaluate them together by indicating both the common and separate points they each make. This will enable us to see the background intuitions these arguments rest on.

The Knowledge Argument: Several versions of this argument can be found in the literature, but the most famous form was introduced by Jackson in his paper "Epiphenomenal Qualia" (1982) and in his later paper "What Mary Didn't Know" (1986). The argument simply claims that one cannot deduce a simple phenomenal fact—what it is like to see red, for example—from the entire (even hypothetically complete) physical knowledge. Imagine a female neuroscientist, named Mary, who has been kept in a black and white room since her birth. The room contains no colorful objects; even the outer surface of her body is made black and white and shades of those. She has, again, a black and white screen in her room to communicate with the outside world. Now, suppose Mary knows every bit of physical/neuroscientific information there is to know about the world, and she has an unrestricted power of reasoning. Accordingly, she knows every detail concerning the processes of color vision. She even sometimes counsels other scientists while she monitors surgical operations on people's brains through her black and white screen. The question is this: Does she know what it is like to see red? While she is in the room, she can perfectly use color terms in sentences like "The rose is red," and "The sky is blue," but when she is released and see the sky and a real rose, would she learn a new fact or not be surprised at all? Jackson claims she would obviously learn a new fact about the world and our color vision. But this would imply the falsity of the physicalist thesis through the following reasoning:

While Mary is still in the room:

(1) Mary knows all the *physical* facts.

(2) But she does not know *some* facts (phenomenal facts).

(3) Therefore, there are facts that are not physical.

(4) Therefore, physicalism is false.

If it is sound, the argument is convincing enough. If such a hypothetical Mary whose physical knowledge is complete misses some phenomenal facts, this means she is ignorant of some phenomenal entities, which cannot belong to the physical domain. Hence, Marcy faces two ontologically different entities, phenomenal and physical. And she can acquire concepts of phenomenal entities only if she is exposed to the phenomenal.

The Conceivability/Modal Argument: Several versions of this argument have been put forward. Those include Kripke's modal argument (1980, pp. 144-55), Bealer's argument from semantic stability (1994), Chalmers' argument from two-dimensional semantics (1996, ch. 2, sec. 4; and ch. 4, sec. 1-2; and 2009) and Nida-Rümelin's argument from cognitive transparency (2007). Versions of this argument employ modal notions such as conceivability, possibility and necessity; and they operate on a common line of thinking. They first attract attention to an epistemic gap between our knowledge of phenomenal facts and that of physical facts, and claim that one can go from this epistemic gap to a metaphysical gap. And if the phenomenal and the physical are metaphysically distinct as such an ontological gap indicates, one can conclude that physicalism is false. Consider the general form of the argument below. Here, *P* stands for the following statement:

P: A physical duplicate of our world exists without any phenomenal ingredient (zombies exist).

(1) It is conceivable that *P*.

(2) If it is conceivable that *P*, then it is possible that *P*.

(3) If it is possible that *P*, then physicalism is false.

(4) Therefore, physicalism is false.

Each of the three premises above has been argued to be false in a number of ways in the literature. What we need here are not the details of the objections to the premises, but rather the sort of justification behind the whole argument. Here the notion of conceivability in premise (1) is aimed to reflect the epistemic gap we are exposed to when we think about the connection between the phenomenal and the physical. If one assumes that this epistemic gap results from an ontological gap between the entities involved, one may legitimately think that the phenomenal and the physical are distinct entities. And this possibility is expressed in premise (2). But the possibility that the phenomenal and the physical are metaphysically distinct falsifies the physicalist thesis (as what premise (3) says) if the metaphysical connection between the phenomenal and the physical has to be a necessary connection, regardless of whether it is identity, supervenience or realization. Granted, there are physicalists who defend contingent (and a posteriori) physicalism like earlier physicalists such as Feigl (1958) and Smart (1959) and the more contemporary ones such as Melnyk (2003). Nevertheless, maintaining contingent physicalism does not seem to be a tenable position. As we have briefly examined in Chapter 2, there are several reasons not to advocate contingent physicalism. Kripkean considerations concerning the modal status of identity statements and the worries of falling into dualism can be recalled here.

The Explanatory Gap Argument: The background problem this argument rests on can be illustrated in several ways. Indeed, one may think that other anti-physicalist arguments wholly or partially root in the same problem. Nevertheless, Levine's construction of the problem (1983, 1999, and 2007) is more striking than other presentations. He attracts attentions to a simple contrast between theoretical reductions in science like that of water to H_2O and the ones occurring between phenomenal and physical states like the reduction of pain to c-fiber firing. For Levine, when scientists identify water with H_2O, the identity statement

seems explanatory. The connection between water and H_2O seems totally clear. After the details of the connection are explained fully, no one would ask "But how come water is identical to H_2O?" The question would be unintelligible. The case is not the same in the reduction of phenomenal states to physical ones. Anyone could ask quite intelligibly "But how come pain is identical to c-fiber firing?" since the purported connection between pain and c-fiber firing does not seem fully explained. Something is left unexplained, and this makes the connection look arbitrary. What is worse is that the explanatory gap in question does not seem to be closable because we have no idea how to bridge the gap. We can put this line of thinking in the form of an argument as follows:

(1) While theoretical identities in science are explanatory (leaving nothing unexplained), identities between phenomenal and physical states seem arbitrary (leaving things unexplained).
(2) There is an explanatory gap between phenomenal and physical states.
(3) The gap is not bridgeable.
(4) The gap is not bridgeable only if phenomenal states are not physical.
(5) Phenomenal states are not physical.
(6) Therefore, physicalism is false.

The crucial question here should be "If we do have such a gap, why is that so?" Is it because there is an ontological gap between phenomenal and physical states, or is it because of some other reasons that do not threaten the core thesis of physicalism? If the answer is "yes" to the former, physicalism is endangered. If it is "yes" to the latter, physicalism is still saved.

As an alternative way of illustrating the explanatory gap problem, Levine cites another anti-physicalist argument that he calls "the open question argument"

(2007), which is worth mentioning here.[11] He basically applies the notion of open question to the case of the explanatory gap problem within the issue of phenomenal-physical identities. To see the conceptual relation between the notions of "explanatory gap" and "open question," consider the following two questions:

(Q1): Are substances that are micro-structurally different from water (say, XYZ), but macro-structurally (regarding their surface properties) the same as water water?

(Q2): Are creatures that are physically different from humans (say, an alien species), but functionally (say, phenomenally) the same as humans, conscious? And if so, is to be conscious for them the same as what it is to be conscious for humans?

These two questions are different for Levine. (Q1) is a totally semantic question. There is no further *non-semantic* chemical or physical fact to be discovered in order to answer the question satisfactorily. The answer will only be based on our semantic decision whether we want to call XYZ water or something else. On the other hand, (Q2) is an open question implying the possibility that there are two fundamentally different sets of properties, namely, phenomenal and physical. This is to say that there may be a deep conceptual gap between our phenomenal and physical concepts.

The Property Dualism Argument: This argument is originally presented as the third objection in Smart's paper (1959), which he ascribes to Max Black. Smart considers the objection the most challenging one that he is least confident to have met satisfactorily. When explaining the objection Smart says "...a sensation can be identified with a brain process only if it has some phenomenal property,

[11] The term "open question argument" has been used within a different context in ethics. It has been originally introduced by Moore to show that ethical properties cannot be identified with non-moral properties (1903, sec. 13).

not possessed by brain processes, whereby one-half of the identification may be, so to speak, pinned down" (p. 149). Here, in order to "pin down" a phenomenal state, he thinks, we need a property (or properties) of that state, or a mode of presentation in modern terms, by which the corresponding phenomenal term can pick out the phenomenal state itself. Otherwise, how can the referent of the phenomenal term be fixed, and equally importantly, how can the relevant identity statement (the statement that the phenomenal state *f* is identical to the physical state *p*) be informative and non-trivial? Consider the classic example of "Evening Star/Morning Star." These two terms refer to the same object, but the reason why we have two different terms and why the identity statement "The Evening Star is identical to the Morning Star" is informative and non-trivial is that they are associated with two different properties (appearing in the evening and in the morning) possessed by the same object. The case, the objection goes, is the same with phenomenal and physical states.

One might think that the argument only runs against the identification of states, events, processes, etc., but not against the identification of properties. This would be wrong (as Block (2006) and others point out) because identification of properties with other ones also requires that the relevant referring terms be associated with some properties of those properties. In other words, the problem arises even for a mind-body "property" identity theory, which basically says "mental properties are identical to physical properties," because the phenomenal property that is claimed to be identical to a physical property can be "pinned down" again only by a property of that phenomenal property—by a mode of presentation of that phenomenal property. What we have said so far can be presented as follows:

(1) Mental-physical identity statements are informative and non-trivial.

(2) Mental-physical identity statements are non-trivial, only if it the terms involved are associated with distinct properties.

(3) If the associated properties are distinct, the ones associated with mental terms are not physical.

(4) If the properties associated with mental terms are not physical, physicalism is false.

(5) Therefore, physicalism is false.

Now, recall that we have examined three metaphysical relations in the previous chapter: identity, supervenience and realization relations postulated by physicalists. And we have said that the anti-physicalist arguments run even against each minimal version of physicalism employing these three distinct relations. Consider the knowledge argument. The anti-physicalist claims that while still being in the black and white room, Mary fails to know the phenomenal character of perceptual experiences, and this intuition is not explained away even if—understood in terms of minimal physicalism—perceptual processes are identical to physical processes, or they supervene on physical processes, or they are realized by physical processes.

The conceivability/modal argument too runs against each of these three relations, since these metaphysical relations must have a necessary character as we discussed in the previous chapter. Granted, most philosophers think of physicalism as a contingent thesis. And the debates regarding the modal status of physicalism are not likely to be completed soon. But it is reasonable to believe—and there are physicalists who strictly defend the same view, like Levine and Trogdon (2009) as well as dualists like Chalmers (1996, 2002, and 2009)—that the only tenable position within physicalism is to defend the thesis that the metaphysical connection between the mental and the physical is a necessary relation. So, if zombies are possible, then these three relations all fail to be the real metaphysical glue holding the mental and the physical together.

The explanatory gap argument as well poses a threat against each of these three relations, since the unbridgeable gap between mental and physical entities will still remain untouched even if mental entities are identical to, or supervene on, or realized by physical entities. None of these relations are actually aimed to close or narrow the gap. The purpose of all is to specify the nature of the connection between the mental and the physical. Even after that specification, the gap still remains. But the gap threatens the identity relation more than it does the other two. If two things are numerically identical to each other, obviously this identity suffers from any unexplained (supposedly metaphysical) gap. On the other hand, if a distinct entity supervenes on, or is realized by another entity, an unexplained gap perhaps is tolerable to some extent.

The property dualism argument, nevertheless, specifically targets the identity theory. It is possible to think that supervenience and realization relations are protected from the threat of this argument on the grounds that there are two distinct entities involved in the postulation of these relations. Clearly if there are two distinct entities, the terms referring to these entities will likely to be associated with distinct properties. Supervenience and realization relations seem fine with distinct properties, as long as those properties conform to the purported nature of the mental-physical relation.

It should also be noted that while what is employed against physicalism by these arguments involves both the phenomenal and the non-phenomenal character of the mind, when employed in the arguments, the phenomenal character strikes us more. Mary is surprised, as the knowledge argument claims, more about the phenomenal quality <sensation red.> The most distinguishing character, among other non-phenomenal mental features, that zombies lack, according to the conceivability/modal argument, is the phenomenal features of the mind. The explanatory gap makes itself more explicit when we attempt to understand the

relation between the phenomenal character of the mind and its physical basis. And finally in the relevant identity statements the properties associated with the relevant terms seem more distinct if one side of the identity statement expresses a phenomenal property. It is this reason why phenomenal consciousness is considered to be more problematic than non-phenomenal character of mind in our effort to correctly place it in our understanding of nature. And this is the chief reason, in accordance with our main target, why we will from now on turn our attention to the phenomenal-physical relation instead of the more general mental-physical relation.

Now it is time to look at the grounds these arguments rest on, and consider why we have intuitions that lead to a phenomenal-physical duality. This exercise is likely to reveal some features of the way we think about mental entities, in particular phenomenal entities.

3.3. The Epistemic/Explanatory Gap

As one can easily notice, the anti-physicalist arguments given above employ a sort of gap between the phenomenal and the physical. Recall the issues surrounding a priori physicalism that we have looked at in Chapter 2. The anti-physicalist arguments employ the non-derivability of phenomenal truths from the complete physical truth *P*. Mary cannot deduce a phenomenal truth from her complete physical knowledge while still being in the black and white room. Zombies are conceivable because *P* does not seem to entail phenomenal truths. We feel a sort of explanatory uneasiness with phenomenal-physical identity statements because they do not seem to be tied to each other derivationally. Even though the two sides of the identity sign are supposed to pick out one and the same referent, those identity statements are informative and non-trivial because there is a derivational distance between what is expressed by phenomenal terms and what is expressed by physical terms. This non-deducibility

turns out to be an epistemic gap between the phenomenal and the physical. The anti-physicalist arguments appeal to this epistemic gap and conclude that we are exposed to such sort of a gap because there is an underlying ontological gap, meaning that the phenomenal and the physical are ontologically distinct.

Here, the terms “epistemic gap” and “explanatory gap,” are different in meaning, though they are used for the same purpose. The former attracts attention to the distinct epistemological positions of epistemic agents, while the latter, introduced by Levine, runs on the explanatory level aiming the same. As the explanatory gap argument presents, we do not suffer from any explicit unintelligibility in comprehending standard scientific identity statements such as “Water is H_2O;” “Lightening is a sudden discharge of electricity in the atmosphere;” “Temperature is average molecular kinetic energy;” “Light is electromagnetic wave;” and so on. On the other hand, we have serious difficulties in understanding phenomenal-physical identity statements such as “a red quale is such and such a neurophysiological property in the brain;” “itching is such and such a neurophysiological state;” “fear is such and such a neurophysiological state;” “imagining is such and such a neurophysiological process;” “succession of thoughts are such and such neurophysiological events,” and so on.

One may rightly think that the anti-physicalist arguments may exploit similar gaps on different levels. On the cognitive level, for example, a subject takes distinct cognitive positions towards phenomenal and physical entities. The way we know about the phenomenal is cognitively distinct from the way we know about the physical. I can know about my visual experience of red by introspecting the content of my relevant visual mental states, whereas I can know about the corresponding physical processes occurring in my brain only by third-person physical examination. This cognitive difference can be said to present a cognitive gap between the phenomenal and the physical. Or think about it on an intuitive

level. One may rightly say "I do not know if there really is an ontological gap, but one thing is clear to me: the phenomenal and the physical seem more distinct to me than other pairs involved in scientifically true identity statements," and call this "intuitive gap." It then seems that we can call the epistemic/explanatory gap the "phenomenal-physical gap," regardless of whether or not it is an ontological one. Henceforth take "PP-gap" as standing for "phenomenal-physical gap" implying both the epistemic and explanatory gaps, and other possible ones that might be intelligibly phrased on different levels.

One may also think that the gaps in question are different conceptualizations of the one and the same distinctness between the phenomenal and physical domains. If we consider these two as distinct domains, however, we might be presupposing that members of the each are distinct entities. We do not want to do that because what we are trying to find out is whether a given member of the phenomenal domain is the one and the same corresponding member of the physical domain. Nevertheless, we have good reasons to think that the phenomenal and the physical constitute at least distinct perspectival domains in the sense that we can only look into these domains from two different perspectives. We can investigate the phenomenal only from the first-person perspective while we can examine the physical from the third-person perspective.[12] Hence, a dualist may think of these two as metaphysically distinct domains, whereas a physicalist may think of these two as two different domains only in an indirect metaphoric sense: the phenomenal and the physical can be scrutinized from different perspectives.

[12] Here, what I mean by the "first-person perspective" is a subject's unique ability to examine (the unique way a subject can examine) consciously the contents of her own mental life—the ability that cannot be possessed by any other subject. What I mean by the "third-person perspective," on the other hand, is the public ability (public way) to examine something, which can be possessed by anyone.

Note that none of the three metaphysical relations proposed by physicalists can escape Levine's core contrast between the phenomenal-physical reductions and the theoretical ones in science. Consider the following three statements:

(I) The feeling of anger is identical to the brain state B.

(S) The feeling of anger supervenes on the brain state B.[13]

(R) The feeling of anger is realized by the brain state B.[14]

All these three statements suffer from a phenomenal-physical gap if the anti-physicalist is right about the PP-gap. The PP-gap is more apparent in (I) because the identity relation in (I) is a numerical one; it relates an entity exactly to itself. The other two, nonetheless, might be thought to be compatible with two ontologically distinct entities. But still, the supervenience relation in (S) and the realization relation in (R) have to be necessary relations as well. Necessary character here means the two distinct entities are so strictly tied to each other that if the base or realizing properties are instantiated in any possible world, the supervenient and realized properties have to be instantiated in the same world as well. How can such a strong modal connection be expected in the presence of the PP-gap? Mental properties are so tightly connected to physical properties, and yet truths involving mental properties cannot be deduced from truths involving physical properties, or statements expressing the supervenience or realization relation between the phenomenal and the physical seem more mysterious than the statements expressing the same relations between other types of properties, say aesthetic and physical properties, or economic and physical properties.

[13] Let us understand necessary supervenience as follows: *A*-properties supervene on *B*-properties with metaphysical necessity if and only if there can be no possible world in which *B*-properties of an x are the same as the original ones while *A*-properties of the same x are different from the original ones, or *B*-properties are instantiated by x while *A*-properties are not.

[14] Let us understand necessary realization as follows: A property *A* is realized by another property *B* with metaphysical necessity if and only if there can be no world in which *B* is instantiated by an x while *A* is not, or *A* is instantiated by x without *B* being instantiated.

3.4. Is the Phenomenal-physical Gap Bridgeable?

Does the PP-gap emerge because we do not know *enough* yet about the phenomenal and the physical? In other words, is it true that the more we know about these two, as our philosophical and scientific understanding improves, the smaller the PP-gap will be? If the answer is "yes," it seems the anti-physicalist arguments will gradually lose their power against physicalism as some philosophers thought (Nagel 1974 and Churchland 1996). But it is not the case at all. The phenomenal and the physical are at least epistemologically distinct domains. Phenomenal entities are non-public. They cannot be observed publicly as non-phenomenal entities can. This simple fact constitutes the following two epistemological constraints that make hard—and sometimes impossible—any complete examinations on the nature of the PP-gap:

> (*NC*) *The Nomological-Physical Constraint: It is nomologically impossible to observe a phenomenal token from the third-person perspective.*

Even a phenomenal subject herself cannot observe a phenomenal token of her own from the third-person perspective. Think of a situation in which a person's skull is opened up during a surgical operation while the person is still awake and conscious. If the person under the surgery attempts to see her own phenomenal state from the third-person perspective, say, by a mechanism of mirrors, what she can observe is nothing but physical states in the brain. It is not likely that any scientific improvement can make it possible to observe a phenomenal token from a perspective other than the person's own first-person perspective. Hence, no phenomenal token can be observed from a third-person perspective as a matter of nomological impossibility. There is another epistemological constraint on the examination of the phenomenal:

(*LC*) *The Logical-Phenomenal Constraint: It is logically impossible for a phenomenal token to be experienced by more than one phenomenal subject.*[15]

A first-person perspective is a person's own perspective. The person uses this perspective when she looks at her internal mental happenings, i.e., when she introspects. And there is an intimate connection between the experience and the experiencing subject. Perhaps this intimate connection is a constitutive one, i.e., one is a component of the other, an issue we will deal with in Chapter 4 and 5. Here what (*LC*) suggests is something more to (*NC*): For every phenomenal token, logically there can only be *one* phenomenal subject. A phenomenal subject cannot observe, even from her own first-person perspective, someone else's phenomenal tokens. Although there will always be room for the possibility that the phenomenal and the physical are numerically identical, as long as we face these two constraints, the PP-gap will always remain. No matter how much we know and learn about the natures of the phenomenal and the physical, the PP-gap will always present itself. This is a point that provides the anti-physicalist arguments with an extra power.

Here what crucially matters for the physicalist is whether the PP-gap is *potentially* bridgeable. Do the above two constraint also suggest that the PP-gap is not bridgeable in principle? In other words, is there any possible way to convincingly show that the phenomenal and the physical are identical? There are many reasons for believing in one or another version of the physicalist thesis. There are also reasons for being sympathetic to an anti-physicalist approach, especially on an intuitive ground that we will mention in the following section. Nevertheless, the anti-physicalist has one strategic advantage. The physicalist has at least the

[15] What I mean by "phenomenal subject" is exactly the subject that entertains phenomenal tokens in a given phenomenal domain whatever her ontological status is. This notion will be examined in detail in Chapter 5.

burden of showing that the implications of the physicalist thesis can *potentially* be proved to be true. Potential provability in this sense is a burden that the physicalist, but not the anti-physicalist, is supposed to shoulder. Consider the following constraint:

> (*PC*) *The Provability Constraint for Referential Terms: It is <u>not provable</u> that any two referential terms refer to one and the same entity* (*are co-referential*)*, unless there is a <u>potential</u> way to verify that there is not two but one entity to which these terms refer.*

To understand what (*PC*) says, consider the classic "Morning Star/Evening Star" example. Once people realized that these two terms refer to the same planet, they did not have an actual way of proving this, but there was a potential way to show that. An imaginary astronaut could have travelled to the orbit of Earth and determined that what people called Morning star and Evening star are the one and the same planet. There was a potential way of proving or disproving the co-referentiality of the two terms. Or imagine a case where someone claims that a planet called X in a distant galaxy is nothing but a mere reflection of a near planet called Y in the same galaxy, and that they look different because X is, for some reason, the reflection of Y's unseen side. This fancy claim may be true or false, but it is a fact that the claim can potentially be proved or disproved. It is both logically and nomologically possible that someone or some creatures prove or disprove the claim in question. Nonetheless, we do not have this chance in the phenomenal-physical case as (*NC*) and (*LC*) dictate. The phenomenal-physical identities cannot satisfy the antecedent of (*PC*)—the "unless clause." Hence, the consequent: However much we know about the phenomenal and the physical, we will not be able to make sure that these two are nothing but the one and the same entity in the metaphysical sense. So, unfortunately, for the physicalist, the gap will always remain. There is no potential way, both in the logical and

nomological sense, to bridge the gap. This and other facts are the reasons why philosophers such as McGinn (1989) are quite pessimistic about our ability to fully comprehend the relation between the phenomenal and the physical.

If the PP-gap is unbridgeable, the philosophical consequence of this is a serious one: A priori physicalism is not tenable since there is no way to deduce mental-to-mental truths or mental-to-physical truths from the complete physical truth *P* because of the PP-gap. If a priori physicalism is not tenable, does this mean a physicalist should give up reduction as well? This is not certain yet. For being unbridgeable does not necessarily entail ontological distinctness; it only entails improvability of a physicalist thesis. The phenomenal can still be identical to the physical in the metaphysical sense. So it depends on how a physicalist approaches to the PP-gap. If she can, for example, show that even if the PP-gap is not bridgeable, there are still good reasons to believe that a phenomenal-to-physical reduction holds, or the PP-gap itself can be explained in physical terms, there will still be rooms for reductive physicalism.

3.5. The Phenomenal-physical Gap and the Intuition of Distinctness

So far, we basically asked two questions regarding the PP-gap: What nature does the PP-gap have, and how does it emerge? Having been persuaded that the PP-gap is not bridgeable in principle, we will now ask a third question: Why do we then have such a gap? To see why we are exposed to such a gap, the best thing we can do is to investigate the sources (or possible causes) of the gap. In other words, we should ask why we immediately think that the phenomenal and the physical are quite distinct? Two reactions are possible here: (1) It is because of their ontologically distinct natures—which would ultimately falsify physicalism. (2) It is not because of their ontologically distinct natures, but because of something else that explains why our immediate thinking presents an ontological distinctness between the phenomenal and the physical. David Papineau (2002, p.

94; 2007; and 2008) is one of those philosophers who sustain (2). He thinks that we have an intuition telling us stubbornly that the phenomenal and the physical are ontologically distinct entities: they cannot be the one and the same entity. This is an intuition even a hard-core materialist cannot help but suffer from. Nevertheless, we have to learn to live with it, because it results from the nature of our way of thinking about phenomenal entities. And the good news is that it is not a problem for physicalism at all.

Papineau admits that the PP-gap that the intuition of distinctness stimulates is not bridgeable. But we can explain in physical terms why it is unbridgeable. And if we can do that, physicalism will be saved from anti-physicalist arguments. The intuition of distinctness is just an intuition. It should not lead one to disregard other good reasons for believing that phenomenal entities are physical/functional entities. For we have other misleading intuitions both in our ordinary lives and in our scientific thinking, such as the intuition that however fast one travels, one does not remain younger than one's twin sister: he gets aged in normal ways. But in spite of this intuition, we believe in Theory of General Relativity. If so, we can disregard the intuition of distinctness as well in favor of other good reasons for believing the physicalist thesis.

As to the source of the intuition of distinctness, Papineau maintains that the intuition results from a gap between our concepts of the phenomenal and the physical. He even offers a theory of phenomenal concepts that he claims explains away the intuition. What he claims is basically that we employ phenomenal and physical concepts in our thinking through quite different cognitive processes. These different processes make them look fundamentally different. There is not two but only one entity, but different conceptualizations yield different concepts which in fact refer to one and the same entity. Although there is no ontological duality, there is, thus, a conceptual duality, which ultimately yields a conceptual

(but not ontological) gap between the phenomenal and the physical. The intuition of distinctness is merely an extrapolation of this conceptual gap.

Leaving aside the details of Papineau's theory of phenomenal concepts (2007) and other similar theories to the following chapter, the proper strategy here, I believe, should not be what Papineau and some others choose to do. Rather, it should be to make sure first, to a satisfactory degree, that reaction (1)—that the intuition results from the ontologically distinct natures of phenomenal items—is a false option. This requires understanding both sides (phenomenality and physicality) to a reasonable degree. We have a good understanding of the physical, but do we have a satisfactory understanding of the phenomenal? If we do not—and Papineau talks as if we do—how can we be sure that the intuition of distinctness does not result from the nature of the phenomenal?

For any physicalist who acknowledges that there is really a gap between the phenomenal and the physical, there are basically three options to follow in accounting for the PP-gap: the PP-gap results from the nature of the phenomenal, or from the nature of "us" (whatever this *us* is), or from the nature of the way *we* know about the phenomenal—through the concepts mediating between us and the phenomenal. Given the fact that we do not have a satisfactory level of understanding of the first two, the phenomenal and the *us*, looking for the root of the PP-gap on the conceptual level is not much likely to yield the right conclusions. We have not asked sufficiently many questions about the phenomenal yet. We have not completed even our basic understanding of different dimensions of phenomenality. We have not yet fully conceptualized the phenomenality. Our notions of phenomenal entities are still primitive. Postulations of new concepts, their correct categorization, discovering new dimensions, and so on are required before any attempt to scrutinize the relation between the phenomenal entities and their concepts can be fruitfully conducted.

We will attempt to take some of these steps in Chapter 5, but we shall first look into reaction (2) in more details in the next chapter. It is the reaction whose basic idea is that the gap between the phenomenal and the physical results not from ontologically distinct natures of these two, but from something else that explains the PP-gap in such terms that can be accommodated within the physicalist approach.

CHAPTER 4

TYPE-B PHYSICALISM, PHENOMENAL CONCEPT STRATEGY AND THE PROBLEM OF PHENOMENAL SUBJECTHOOD

In the preceding chapter, we have examined a number of anti-physicalist arguments and the grounds these arguments rest on. The chief ground was the gap between the phenomenal and the physical. We have also seen the basic difference between the physicalist and the anti-physicalist reactions to the PP-gap. The anti-physicalist reaction was simple: the PP-gap arises because there is an underlying ontological gap. The physicalist reaction, on the other hand, varies in kinds. We will not examine all of those kinds here. Instead, we will assess one version of the reaction that respects the phenomenal reality unlike the others. And we will do that particularly in relation with the subjectivity of phenomenal consciousness.

One thing should also be kept in mind in what follows. Our central focus, in the following chapters, will be the identity physicalism adopting the type-identity relation (not the supervenience and realization physicalism). Besides, as we have pointed out in Chapter 2, for the identity relation to truly hold between the phenomenal and the physical in the metaphysical sense, it has to be a necessary relation by Kripkean considerations. Moreover, as argued in Chapter 3, any version of necessary physicalism that concedes the reality of PP-gap has to have an a posteriori character because if there is really an unbridgeable PP-gap—if phenomenal entities cannot be reductively explained in physical terms—phenomenal-physical identity statements can only be known a posteriori. The a posteriori character of the phenomenal-physical identities, on the other hand,

will appear, in this chapter, as something crucial that has to be explicated in detail by the proponents of a posteriori necessary physicalism. So, throughout this chapter, we will be assuming that a posteriori necessary physicalism is the best candidate for a comprehensive and coherent physicalist story of phenomenal consciousness.

4.1. Two Types of Physicalist Response to the Phenomenal-Physical Gap

Once the PP-gap is recognized particularly by the identity physicalist to be a genuine puzzle that has to be treated properly, there appear only a few possible ways of reacting to it. Chalmers' categorization of these reactions is a good one (2003 and 2007). Two of those reactions, which he calls type-A and type-B physicalism, are what we need here. Type-A physicalism denies completely that there is an explanatory gap at all. The so-called explanatory gap has no significant basis according to this view. Proper treatment of the phenomenal and the physical would prove that the alleged gap is just an illusion. Phenomenal-physical identity statements are no different from other theoretical identity statements in science. If, for example, the statement "Sadness is such and such a physical or functional state in the brain" seems gappy as opposed to scientific reductions like "Temperature is average molecular kinetic energy," it is not because there really is a gap of any sort, but because of our improper treatment of phenomenal-physical identities. Analytic functionalism and philosophical behaviorism can be cited here as versions of type-A physicalism.

Another reaction, which Chalmers calls Type-B physicalism embraces the gap, as opposed to type-A physicalism, and even admits that the gap is not bridgeable at all. Yet, the type-B physicalist boldly denies that the PP-gap falsifies the core thesis of physicalism. How so? The basic idea is this: While there is an unbridgeable gap between the phenomenal and the physical, and no version of identity physicalism can provide a satisfactory account denying the existence of

the PP-gap, there is still a way of saving physicalism. We will never be in a situation where the PP-gap vanishes, but we can still give a perfect account of *why we face the PP-gap* in physical terms, or at least in physically embraceable (non-phenomenal) terms. Although we cannot deduce phenomenal truths from the complete physical truth *P*, giving such an account would actually mean *deducing the gap itself* from the complete physical truth *P*.

There is another intermediate way of reacting to the PP-gap that should be briefly reminded here, since we have already examined it shortly in Section 3.4 without giving it a name. This version of physicalism is simply a positive answer to the question "Does the PP-gap emerge because we do not know enough yet about the nature of either the phenomenal or the physical?" Proponents of this view admit the reality of the PP-gap, but suggest that it appears only momentarily, because, they believe, the PP-gap will be "less gappy" as our philosophical and scientific understanding improves, and eventually will vanish completely. As I argued earlier in the previous chapter, however, the three constraints (*NC*), (*LC*) and (*PC*) set out clearly that the PP-gap is not due to our lack of knowledge. Human cognition is exposed to several kinds of undeniable epistemic limits including biological and conceptual ones. (*LC*) and (*PC*) derive not only from our epistemic limits, but also from the subjectivity of phenomenal consciousness that involves such mental phenomena as phenomenal I-ness, agency, and peculiarity, which I will argue for and examine in detail in Chapter 5. Besides, the consequent of (*PC*), deduced from (*NC*) and (*LC*), that phenomenal-physical identity statements cannot be proved to be true, establishes that the PP-gap will never disappear.

There is a conceptual relation between a priori and type-A physicalism; and between a posteriori and type-B physicalism. An a priori physicalist has to be a type-A physicalist, because there is no way for her to embrace the PP-gap. A type-

A physicalist, on the other hand does not have to be an a priori physicalist since it is possible for her, for example, to be an eliminativist who denies the whole reality of phenomenal consciousness altogether. At the same time, since an a posteriori physicalist does not have to concede the necessity of the physicalist thesis—since she can defend a contingent version of physicalism—she does not have to be a type-B physicalist. But a type-B physicalist has to adopt the a posteriori character of phenomenal-physical identity statements because she admits the reality of the PP-gap.

Type-A and type-B physicalisms are actually incompatible interpretations of anti-physicalist arguments, though their purpose is the same. For a type-A physicalist, a fully correct analysis of Mary's case would show that Mary wouldn't learn a new truth that she did not know before leaving the black and white room. Zombies are not conceivable if we properly take into account the full features of such fancy creatures. The explanatory gap argument is not sound because there is not a gap of any sort. Phenomenal-physical identity statements are informative, but it is not because two metaphysically distinct properties are involved. A type-B physicalist, on the other hand, considers these issues completely in an opposite way. Since she respects the phenomenal reality, for a type-B physicalist, Mary may learn something new after leaving the room; zombies are conceivable; there really is an epistemic/explanatory gap; and phenomenal-physical identity statements are non-trivial and the relevant terms are really associated with distinct properties.

In spite of these admissions, type-B physicalists deny any possibility of derivation form the above cases to an ontological gap between the phenomenal and the physical. How so? To show that an ontological gap cannot be deduced in those cases, they usually appeal to identity statements in standard scientific reductions such as "Water is H_2O" and "Light is electromagnetic wave." Just like 'water' and

'H_2O' are two co-referential terms referring to one and the same chemical, 'pain' and 'c-fiber firing,' too, are two co-referential terms referring to one and the same physical/functional state in the brain. Just like "Light is electromagnetic wave" cannot be known a priori, "Itch is the neurological state N" cannot be known a priori. But both are still necessary truths. Phenomenal-physical identity statements are a posteriori necessities just like standard identity statements in science.

The notion of "a posteriori necessity" is an innovation of Kripke (1971 and 1980, pp. 140-53). But even Kripke points out a serious problem with the analogy between the phenomenal-physical identity statements and the standard reductions in science. Compare the following two statements:

(1) Water is H_2O

(2) Pain is c-fiber firing.

It has never been difficult to explain the a posteriori character of scientifically discovered identity truths such as (1), argued Kripke. For when scientists discovered that water is identical to H_2O, they came to know this identity by means of the contingent properties of water: being odorless, colorless, transparent, and so on. There are two concepts associated with two different modes of presentation here: one (those contingent properties) associated with the concept "water" and the other (those essential properties) associated with the concept "H_20." But both concepts pick out the same referent—the same chemical substance. What scientists discovered was that the chemical substance H_2O is the substance that instantiates those contingent properties that we have been ascribing for a long time to what we call "water." But H_2O *might not have instantiated* those contingent properties: those are not essential properties that are instantiated by H_2O in all possible worlds. We came to know about those contingent properties and became familiar with them only in this actual world. Hence, the reason why (1) is known a posteriori is very clear. What is stated in (1)

is a necessary truth, but we know it a posteriori because of a contingent mode of presentation of water associated with contingent properties of water.

Discovering the truth of (2) a posteriori, on the other hand, is quite problematic, argued Kripke. The statement (2) also contains two rigid designators "pain" and "c-fiber firing," but here we do not have the needed contingent properties to explain the a posteriori character of (2), since all the properties pain instantiates are its essential properties. Pain does not have accidental/contingent properties. Whatever properties it instantiates are rigidly attached to it in all possible worlds. We cannot conceive a world in which pain still has its real nature (pain is still pain), but lacks some of its properties we enjoy in the actual world. The reason is obvious: what we know about pain is nothing but those properties. Saying that those properties are contingent would imply that pain has another aspect of its nature—the essential aspect, which we are completely unaware of. But there is nothing available in the phenomenal domain other than those properties, which constitute the essential nature of pain. Therefore, the type-B physicalist has to find another way of explaining the a posteriori character of necessarily true identity statements expressing the phenomenal-physical connection. A specific strategy is specifically utilized exactly for this purpose.

4.2. The Phenomenal Concept Strategy

We have already pointed out the basic rationale behind type-B physicalism. So the task is:

(a) To embrace the reality of phenomenal consciousness.

(b) To admit that the phenomenal-physical identities exhibit an unbridgeable epistemic/explanatory gap.

(c) To explain fully why the epistemic/explanatory gap exists, in physical (or at least phenomenal-free) terms without leaving behind any residue of the gap or

creating a new one, or to deduce the gap itself from the complete physical truth *P*.

(d) To do task (c) in such a way that we can understand why a necessary phenomenal-physical identity statement is known a posteriori.

For a type-B physicalist, in order to fulfill these tasks, options are limited on her metaphysical undertaking. Consider the below illustration:

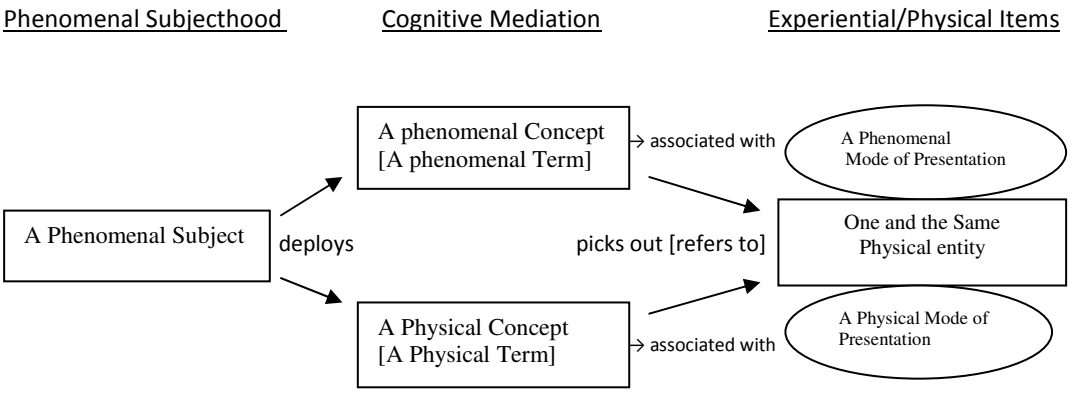

Figure 4.1: A Type-B Physicalist's Way of Thinking about Phenomenal Consciousness

The type-B physicalist both has to point to the source of the PP-gap and has to explain the a posteriori character of phenomenal-physical identity statements somewhere on this metaphysical picture while still respecting the reality of phenomenal consciousness. Pointing to the source of the PP-gap actually means ascribing contingent properties to some elements on this picture within the thesis to be defended. Those cannot be the elements on the right-hand side of the figure because what we know about phenomenal/physical items are only essential properties, and our physical concepts may only pick out their referents through essential properties. One may, on the other hand, attempt to detect the source of the PP-gap by referring to the nature of cognitive mediation, particularly to the nature of concepts, or to the nature of deploying-relation, or to the nature of picking out/referring-relation. About the last one, however, the

Kripkean considerations concerning rigid designators pose problems. The first two are open possibilities. Finally, one may also attempt to detect the root of the PP-gap in the nature of phenomenal subjecthood, an option which we will discuss in detail later. For now, it should be clear that there can be several strategies open to type-B physicalism in order to fulfill the three tasks above.

A particular strategy, which is the one adopted by what is considered to be the most promising version of type-B physicalism, suggests that the root of the PP-gap resides in cognitive mediation, particularly in the nature of phenomenal concepts. Following Stoljar (2005), the strategy is named the "phenomenal concept strategy." Proponents of the phenomenal concept strategy (henceforth PCS) admit that we have in our hands a kind of PP-gap that has to be explicated in detail. But from such a gap one cannot deduce an ontological gap since the PP-gap is merely the projection of a *conceptual gap*. A defender of this strategy typically holds that phenomenal concepts—concepts that we employ when we think about phenomenally conscious states—are fundamentally different from physical or functional concepts—concepts that we employ when we think about non-phenomenal entities. This difference between the two kinds of concepts, i.e. this conceptual duality creates a conceptual gap. The conceptual duality, however, does not necessarily give rise to an ontological duality. According to the defender of PCS, phenomenal sates, events, and properties are still identical to physical/functional states, events, properties, etc. Or at least, phenomenal properties supervene on or realized by physical/functional properties with *metaphysical necessity* in the sense that there could not be any phenomenal property somewhere without there being physical/functional properties there. It is true that anyone who thinks about phenomenal-physical identity statements could not help but face a conceptual duality, but there is still room to reject

ontological duality on the grounds of conceptual duality; thus, physicalism is still saved.[16]

To give an example, PCS defenders believe that the terms 'fear' and 'brain state F' refer to the same physical/functional state though they express different concepts. There are two different terms corresponding to two distinct concepts, but there is only one (physical/functional) entity with which all these terms and concepts are associated. There are two different terms because there are two modes of presentation of the same physical or functional state/process/property.

Modes of presentations are properties or sets of properties. These are contingent properties on one side of the identity statements in those cases of standard scientific reductions. In the cases of phenomenal-physical identities, on the other hand, the properties serving as modes of presentations on both phenomenal and physical sides are essential ones. Accordingly, the properties associated with the concept "pain" and the properties associated with the concept "c-fiber firing" are all essential properties, and they are all instantiated by one and the same physical/functional state. How do PCS proponents account for the a posteriori character of these identity statements then? Their answer is simple: there is a special referential relation between phenomenal concepts and the associated essential properties that serve as modes of presentation. This special relation is nothing like other sorts of relation between non-phenomenal concepts and non-phenomenal entities. And because this special relation is a contingent relation—in that such a relation occurs in the actual world and might not occur in other possible worlds—the phenomenal-physical identity statements can be known only a posteriori.

[16] The ontological status of concepts, their acquisition, possession, and individuation conditions are quite controversial issues. Although I prefer to stay neutral to the issues, it is obvious that talks of concepts within the framework of PCS at least requires one to be a realist about concepts, otherwise the discussion cannot get started. So, I will take concepts to be the simple constituents of thoughts, whose existence cannot be denied coherently.

Here, PCS defenders specifically aim to attract attention to the special nature of phenomenal concepts. Due to this special nature, phenomenal concepts are conceptually isolated from all other sorts of concepts that we employ in thinking. What makes phenomenal concepts and the way they refer to their referents so distinct and special differs from one theory of phenomenal concepts to another. I am not planning to examine in detail any theory of phenomenal concepts here, but rather detect the general architecture offered for these concepts. This will allow us to see how the fundamental cognitive difference between phenomenal concepts and physical concepts leaves room for saving the core thesis of physicalism. To my knowledge, the following can be considered a list of the most commonly known theories of phenomenal concepts in the literature:

Phenomenal concepts:

→ *are causal-recognitional concepts* (*Direct Reference Account*): Brian Loar (1990 and 1997) and Michael Tye (2003, but he very recently denied any need for phenomenal concepts to explain the PP-gap (2009)).

→ *are pure-recognitional concepts* (*Pure-Recognitional Account*): Peter Carruthers (2003).

→ *have distinct conceptual role* (*Dual-Conceptual Role Account*): Christopher Hill (1997, developing a suggestion made by Thomas Nagel (1974)).

→ *are indexical concepts* (*Demonstrative Account*)*:* John Perry ((2001) based on his another paper (1979)), John O'Dea (2002) and Janet Levin (2007).

→ *are quotational concepts* (*Constitutional Account*): Christopher Hill and Brian McLaughlin (1999), Katalin Balog (1999, 2006, and 2009), David Papineau (2002 and 2007), Ned Block (2006), and on the dualist side David Chalmers (2003).

→ *have no evidential intermediary*: Scott Sturgeon (1994).

→ *can be explained by the Narrow Representationalist Account of qualitative Experience*: Georges Rey (1998).

→ *have non-ascriptive modes of presentation as opposed to other sorts of non-phenomenal concepts that have ascriptive modes of presentation*: Joseph Levine (2001, pp. 53-4).

→ *can be explained by an information-theoretic analysis of the special relation between phenomenal and perceptual concepts*: Aydede and Güzeldere (2005).

Loar's paper "Phenomenal Concepts" (1990) can be considered to be the pioneering step in attempting to give a physical explanation to the PP-gap while still respecting the phenomenal reality. Loar thinks that phenomenal concepts are *special recognitional concepts* that are different from other sorts of recognitional concepts. Unlike other recognitional concepts, the referent of a phenomenal concept serves as its own mode of presentation, meaning it is a non-contingent mode of presentation. This is kind of bizarre because in the usual cases, referents do not serve as modes of presentation. Only those properties instantiated (contingently or necessarily) by the referents are associated with the concepts and terms, and serve as modes of presentation. Two co-referential terms—and accordingly two co-referential concepts—refer to one and the same entity by being associated with distinct sets of properties instantiated by the same referent. In the case of phenomenal concepts, however, Loar argues, referents—which are also properties—themselves serve as their own modes of presentation. So the difference between the concept of "water" and that of "pain" is that the former picks out its referent by means of a contingent mode of presentation (a set of properties such as liquidity, colorlessness, tastelessness, and so on) whereas the latter picks out its referent directly because the referent itself (which is a property or a set of properties) serve as its own mode of presentation.

Balog (2009) points out that other theories of phenomenal concepts can be grouped under two main directions, both inspired by Loar's idea above. One direction emphasizes "direct reference," and the other "involvement of the

experience within the concept." Direct reference accounts include, for example, Tye's theory of phenomenal concepts (2003) according to which phenomenal concepts refer to phenomenal properties directly without there being a mode of presentation. This is different from Loar's original proposal in that Loar's account still allows for there being a mode of presentation, which happens to be non-contingent. In direct reference accounts, a phenomenal concept and its referent remain distinct in the deployment of the concept by the subject. Balog argues that this is a disadvantage for such accounts because it does not explain the special intimate connection between the phenomenal concept and its referent—a type of cognitive connection that we do not encounter in cases of non-phenomenal concepts (which we will examine in detail in the next section). I agree with the above diagnosis, and because of that I will focus on the other direction.

The other direction includes accounts based on the involvement of the experience itself within the concept. The constitutional account can be cited under this direction, according to which a phenomenal concept refers to a phenomenal entity without a cognitive distance between them in that the referent itself is constitutive to the concept. My concept of my "current stomachache," for example, is wholly or partially constituted by the referent itself, my current stomachache. This seems quite clever because it satisfies most of the intuitions we have about our cognitive relation to phenomenally conscious states. Nevertheless, it faces a severe dilemma regarding phenomenal subjecthood, which we will target throughout the following two chapters.

For our present purposes in this chapter, what we need is the general architecture of the strategy: phenomenal concepts have special natures that make them fundamentally different, in the cognitive sense, from other sorts of concepts like physical, functional, formal concepts, etc. This special nature is what

gives rise to the PP-gap. But it does not pose any threat to the core thesis of physicalism—that everything is physical, or every truth is a physical truth.

About this general structure, Chalmers introduces an argument that, he claims, puts pressure on the proponents of PCS to choose one of the horns of an anti-physicalist dilemma, either of which undermines physicalism. He argues that on the one hand, if a theory of phenomenal concepts explains the PP-gap in a phenomenally satisfactory way, the theory (of phenomenal concepts) itself cannot be explained in physical terms. If, on the other hand, a theory of phenomenal concepts can be itself explained in physical terms, it cannot explain the PP-gap in a phenomenally satisfactory way. Any theory of phenomenal concepts within PCS, he argues, has to suffer the consequences of one of these two options.

There are other critics who raise significant objections to PCS. In fact, options vary to criticize PCS. One may ground her objection, for instance, on the nature of phenomenal properties, or on the nature of phenomenal concepts, or on the nature of the relation between phenomenal properties and concepts. Or one may criticize PCS on the ground of the nature of the relation between phenomenal subjects and the phenomenal concepts they deploy. All these options have in fact been employed by several philosophers. Nevertheless, it is surprising that the notion of *phenomenal subjecthood*, assumed one way or another by the proponents of PCS, has never been criticized. There may, of course, be other ways of criticizing PCS from a physicalist perspective; for example, by denying the existence of any phenomenal concept (Ball 2009) and denying the very special nature attributed to the phenomenal concepts (Tye 2009), and so on, which I will not examine here at all. As I said earlier, I will, instead, criticize the strategy from the standpoint of the subjective character of phenomenal consciousness, or the "phenomenal subjecthood" as I prefer to call it.

4.3. The Unique Features of Phenomenal Concepts

In order to get an initial grip on the problem of subjectivity for PCS, we need to focus first on the very nature of phenomenal concepts. Without aiming to adopt any theory of phenomenal concepts motivated by PCS, several philosophers attempted to detect several features of phenomenal concepts. Levine, for example, diagnosed that the content of phenomenal concepts (or our grasp of phenomenal properties) are "substantive" and "determinate"—phenomenal concepts are presentationally thick concepts (2001), which proponents of PCS cannot explain satisfactorily (2007). Our grasp of phenomenal properties is substantive in the sense that we grasp *directly the substantive nature* of phenomenal properties, not as "know-not-what" (2001, p. 84). This happens because phenomenal properties serve as their own mode of presentation. A quick reminder is needed here. In the case of other sorts of concepts, like the concept of "water," there are mediating properties serving as modes of presentation by means of which concepts pick out their referent. Accordingly, <water> picks out its referent in virtue of some properties (of H_2O) that serve as a mode of presentation. The property (or properties) serving as a mode of presentation like being transparent, odorless, tasteless, etc. mediate between the concept "water" and its referent. In the case of a phenomenal concept, however, there is no mediating property between the concept and the referent. The referent itself serves as its own mode of presentation. This is why the contents of phenomenal concepts are said to be substantive—they give substantive information about their referents.

Our grasp of phenomenal properties, for Levine, is also determinate in the sense that we grasp a phenomenal property as a specific quality as distinguished from other sorts of phenomenal qualities. This means that what we grasp when we think about a phenomenal property through a phenomenal concept is an intrinsic

quality, not a relational one. We can contrast this type of grasping to that of a non-phenomenal property like "weight." The latter is clearly a relational property of a physical object as our grasp of it necessarily relies on other properties (of the same or different physical objects) such as the gravitational effect on the object.

Balog gives a list of desiderata for a theory of phenomenal concepts to satisfy—desiderata that constitutes most commonly accepted features of phenomenal concepts in the literature in their relation to phenomenal consciousness (2009, sec. 3). The list includes the following: *Acquaintance*: We are *acquainted* with our phenomenal states in the sense that we are directly aware of our phenomenal states without any intermediary rational processes. *Asymmetric epistemology*: We are directly aware of our phenomenal states in a way no one else can be. Other people can know about our phenomenal states only through the third-person observations. *Infallibility/incorrigibility intuition*: We seem to be infallible in our phenomenal judgments; thus they are incorrigible. *Experience thesis*: Only subjects who is undergoing (or has undergone before) a phenomenal state can acquire the corresponding phenomenal concept. *Fineness of grain*: Our color experiences can be so detailed that we can experience millions of different shades of a given color. But our phenomenal concepts of a given color are quite limited.

Trogdon (2010) summarizes the features of what he calls "cognitive immediacy" based on Russell's notion of acquaintance and Levine's notion of substantive and determinate content. He draws attention to two main features of our cognitive relation to the phenomenal, one emphasizing the "directness," and the other emphasizing the "thickness" (substantive and determinate content) of our grasp of a phenomenal property. We have already clarified the thickness of phenomenal concepts. The directness intuition is concerned with Russell's notion of "knowledge by acquaintance." Indeed, most of the features cited above of

phenomenal properties, or of our cognitive relation to phenomenal properties, or of the referential relation between phenomenal concepts and properties one way or another root in Russell's notion of acquaintance. It is a good starting point then to look at the notion of "acquaintance" in going into details of the problem of phenomenal subjecthood.

Russell himself explains the content of <acquaintance> by saying "I say that I am acquainted with an object when I have a direct cognitive relation to that object, i.e. when I am directly aware of the object itself." (1910, p. 108) And he adds in his *The Problems of Philosophy* that "We shall say that we have *acquaintance* with anything of which we are directly aware, without the intermediary of any process of inference or any knowledge of truths." (1912, p. 25) The application of these lines to the phenomenal-physical case should be like this: If I am acquainted with an object, say the quale blue I am experiencing now by looking at the blue pencil box on my table, there is no *cognitive distance* between "me" and the "blue quale" I am acquainted with. This is in contrast with "knowledge by description" in which I know about an object through an indirect cognitive relation (through intermediary rational processes like inferences); say the table on which I am writing these lines at the moment. I know about the table in virtue of some inferences like "It seems to me that there is a table in front of me, on which I am writing. Therefore there is a table in front of me, which I am writing on." This is taken to be as that there is a cognitive distance between "me" and the "table," which does not exist between "me" and the "blue quale."

Now, recall the metaphysical picture suggested by the type-B physicalist in Figure 4.1. One may of course have an alternative picture for the metaphysics of phenomenal cognition, but two things will be inevitable in any picture: the phenomenal subject and what she experiences—the experiential content itself. If we want to think at the most ground level, and only consider these two items, all

the features of phenomenal experiences/concepts pointed out above will come down to three aspects of the cognitive relation between a phenomenal subject and what she experiences: (i) cognitive directness, (ii) cognitive privateness, and (iii) cognitive thickness. It appears that being substantive and determinate, and fineness of grain fall under (iii); acquaintance falls under (i); asymmetric epistemology, infallibility/incorrigibility intuition, and experience thesis fall under (ii). I believe other likely features that essentially characterize phenomenal experiences/concepts would fall under one of these three aspects. But these aspects belong to the cognitive relation of the experiencers to what they experience, which tells us something very crucial about the phenomenology involved. Phenomenal experiences/concepts require their intrinsic character mostly from their possessors, or experiencers, the phenomenal subjects who entertain them. They owe their intrinsic character to the substantial existence of their possessors. Let us look at what I mean by this in more detail within another context below.

4.4. PCS and the Problem of Phenomenal Subjecthood

Here, the first thing one should notice about PCS is that it is a qualia-centered strategy. Within this strategy, accounts of phenomenal concepts are qualia-based accounts. Nevertheless, phenomenality of consciousness *necessarily* involves a subjective character as well. And this subjective character has been neglected for some or other reasons. I agree with Levine's following diagnosis here:

> While the problem of providing an explanation for qualitative character... has been the focus of most of the literature on conscious experience ... the deepest problem lies with understanding subjectivity. In fact ... the explanatory gap between physical properties and qualitative properties is a symptom of the subjectivity of consciousness. (2001, p. 7)

What Levine understands from "the subjectivity of consciousness" is the cognitive relation a phenomenally conscious being has to its conscious states (2001, sec. 6.6). I, on the other hand, will construe subjectivity differently, in a more substantive way. Subjectivity of phenomenal consciousness, as I will construe it, primarily involves the very ontological nature of a *phenomenal subject herself* rather than this subject's instantiating the cognitive relation to whatever she experiences. And I will diagnose several fundamental problems with regard to the subjectivity of phenomenal consciousness as construed that way.

Now, one can easily notice that all of those intrinsic features we have been mentioning at the end of the previous section are either of what is experienced or of the relation (cognitive mediation—concepts) holding between what is experienced and "us." But recall that the original task for a proponent of type-B physicalism is to give a physical explanation of the PP-gap while not denying the phenomenal reality as a *whole*. If what gives rise to our discerning the PP-gap is somehow substantially connected with "what constitutes *us* (the phenomenal subject)," then attempting to detect the source of the PP-gap merely in the nature of experiential items and/or in the nature of cognitive mediation will not lead us to the real causes/factors why we encounter such a gap. To put it straightforwardly, I am here simply saying that the PP-gap may result not only from the nature of phenomenal properties or the nature of concepts (of these properties) deployed by the phenomenal subjects of phenomenal properties, but also from the very nature of the *phenomenal subjecthood*. If this is actually true, PCS will have serious problems. I will argue that it is true. But there is more. As has been stated several times before, phenomenality of consciousness possesses a twofold character: the qualitative character and the subjective character. As expressed in the quotation from Levine at the outset of this section in a similar thought, the qualitative character, I believe, is ontologically dependent on the subjective character—though I don't think Levine has an ontological dependence

in mind. In other words, without there being a phenomenal subject, there would not exist any qualia either. This is not all. There are some other substantial reasons for thinking that the subject term in any introspective phenomenal reports, such as "*I* am feeling pain right now," or "*I* am experiencing a red quale now" has to have a real ontological correspondence in the relevant phenomenal domain. If all these are true, proponents of PCS will face a serious dilemma regarding the phenomenal subjecthood.

To specify clearly where PCS would fail if the above is true, recall the tasks proponents of PCS originally have undertaken:

(a) Show that what causes us to discern the PP-gap results *completely* from a fundamental difference between our phenomenal and physical/functional concepts. Here, 'completely' means "without leaving behind any residue of the PP-gap or creating a new one."

(b) Explicate the fundamental difference between the two sorts of concepts in physical (or at least phenomenal-free) terms in such a way that we can understand why phenomenal-physical identity statements are known a posteriori.

It is clear that task (b) relies on task (a), and proponents of PCS will fail to fulfill task (a) if (i) phenomenal subjecthood is among the ingredients that are responsible for the PP-gap in a substantial way, and/or (ii) there is an ontological dependence (mutual or one way) between phenomenal subjects (the subjective character) and the qualitative character, and/or (iii) there is a real ontological correspondence in a given phenomenal domain for the subject terms of phenomenal reports.

One may, at this point, immediately raise an objection and say that phenomenal subjecthood does not even require any explanation, let alone explicating it in

relation with other components in the phenomenal domain. I will reckon with this potential objection and, because of this, will not have any presumption about the ontological status of phenomenal subjecthood. It is then better to follow a step-by-step process to have a clear mind regarding phenomenal subjecthood without presuming anything ontological about it. Here is the starting question: Does phenomenal subjecthood really require explanation (for the physicalists or the proponents of PCS)? My answer is "Yes," and I believe the negative response is not tenable. Let me argue for this. First, consider the very content of the notion of "experience," which is a defining concept of "phenomenal consciousness." It, at the most basic level, consists of two types components in thought: an *experiencer* and what is experienced. I believe it is quite clear that no one can conceive of an experience without there being an experiencer of it. An experiencer, in the phenomenal sense, is what I will call a "phenomenal subject." So, it seems that we have a desideratum concerning the content of <experience>:

> (*ES*) *Desideratum for the Existence of Phenomenal Subjects*: *For every phenomenal token (for every experience), there is a phenomenal subject (an experiencer) who experiences that phenomenal token. A phenomenal subject is an essential component of a phenomenal domain.*

We will return to (*ES*), and other desiderata regarding phenomenal subjecthood in the following chapter. For our present purposes, it is clear that a phenomenal subject is an essential component of <experience>.

Second, in providing an explanation for the PP-gap, proponents of PCS are expected to offer a contingent thesis about phenomenal concepts. We do not have to take a position on the controversial issues concerning the metaphysics and epistemology of concepts in general and phenomenal concepts in particular—issues like concept possession, concept acquisition, individuation

conditions and reference determination. Nor do we have to be a concept realist here, but one thing is quite clear: whether or not they are mind-dependent entities, concepts inevitably need a concept possessor. We cannot think of a concept being deployed without a thinker. It is hard to imagine a theory of phenomenal concepts without the notion of "concept possessor" utilized in the theory, whose referent we do not know about fully.

Third, as we have already pointed out several times before, phenomenality of consciousness has a twofold structure: qualitative and subjective characters. The PP-gap arises because of the nature of phenomenality involving not only qualia but also subjectivity. I agree with Levine's construal of subjectivity as expressed below:

> There are two important dimensions to my having this reddish experience. First, as mentioned above, there is something it's like for me to have this experience. Not only is it a matter of some state (my experience) having some feature (being reddish) but, being an experience, its being reddish is "for me," a way it's like *for me*, in a way that being red is like nothing for—in fact is not in any way "for"—my diskette case. Let's call this the *subjectivity* of conscious experience. Nagel (1974) himself emphasized this feature by noting that conscious experience involves our having a "point of view." (2001, pp. 6-7, italics in original.)

But I am asking a further question: What is "me" here? And I am insisting that this is a question that needs to be answered for the physicalist to adequately respond to the PP-gap. The reason is that the PP-gap is not something that arises merely between subjectless qualia and their purported physical correspondences. It arises between phenomenality of consciousness and physical states. And phenomenality would be missing one of its essential ingredients if it is construed as involving merely a qualitative character without a subjective character. If it is so, this is against the comprehensiveness requirement for the physicalist.

Remember that in order to be convincing, physicalism and other approaches to phenomenal consciousness must present a *comprehensive and coherent* story leaving behind nothing unexplained—in the phenomenal sense—regarding phenomenal consciousness.

Notice that in the above quotation, Levine construes a phenomenal subject as something passive without emphasizing its active subjecthood—think about what he can possibly mean by the phrase "for me" in his sentence "...there is something it's like for me to have this experience." A phenomenal subject, however, can be a "phenomenal *agent*" too. I may *intentionally and deliberately attempt* to experience the reddishness of Levine's diskette case, acting as an agent. This is a topic that we will examine in the following chapter.

Fourth, in relation with subjectivity, there is more to say about phenomenality and qualitative character here. I argue that the qualitative character ontologically depends on the subjective character. In other words, without a phenomenal subject there cannot exist a quale in any possible world. This is different from the ontological dependence between experiences and their experiencers mentioned a few pages ago. The ontological dependence between an experience and its experiencer is more like a conceptual dependence. The content of <experience>—which consists of both subjectivity and qualitativity—conceptually contains <experiencer> in it. There seems no conceptual dependence of this sort between subjectivity and qualitativity, however. But I claim there is still an ontological dependence: Qualia ontologically depend on the subjective character. To test this idea, consider a zombie counterpart of me that, nevertheless, possesses qualitative features. This zombie-me is different from the standard zombies we are familiar with from the literature in that it possess all qualitative features of phenomenal consciousness, but it nonetheless misses one thing: "the sense of being a subject," which we may also call "sense of I-ness." Despite the

fact that most of the qualia I am experiencing also occur in its phenomenal domain, this zombie-me lacks the sense of I-ness I always feel deep inside of me whenever I am phenomenally conscious. This subjectivity-lacking (or subjectless) zombie and I simultaneously utter sentences that include the first-person pronoun "I," such as "I am experiencing a red quale now," but nevertheless "I" in my usage has a full meaning and most probably refers to a real referent, whereas the same pronoun in my zombie counterpart's utterance has no meaning at all and refers to nothing. Well, I believe this is completely inconceivable provided that what we understand by a quale should be shaped within a phenomenal construal.[17] No one can coherently conceive such a zombie. Conceiving a zombie like that would be conceiving a shadow in the absence of the object it is the shadow of.

The question then is this: Why can we not conceive of such a subjectless but qualia-possessing zombie? My answer is quite simple: It is because of our conception of a quale. A quale is a quale only because it is *experienced* by a phenomenal subject having a sense of I-ness. There cannot be a quale "floating around" alone that has not ever been experienced by a phenomenal subject. If so, then there could also be a sort of hidden conceptual tie, if not a dependence, between a <quale> and a <phenomenal subject.>

There is a bad habit in some circles of treating what is experienced (i.e. the experience itself) as if it is independent of its experiencer or of treating qualia as if they can exist with a total independence from a phenomenal subject. Proponents of PCS think of experienced entities as if they were third-person

[17] Recall that we distinguished between the qualitative and the phenomenal construal of experiences in Chapter 1 in relation to blindsight patients. According to the qualitative construal of experiences, there can be experiences, and thus qualia, that lack the relevant phenomenal subject's awareness. I do not adopt this construal, but even if one adopts this construal, there is still no reason to claim that qualia exist with a total independence from a phenomenal subject, because absence of awareness of a phenomenal subject in blindsight patients is only momentary.

objects. They construe “pain,” for example, like tables, chairs, carpets, etc. But this creates a false base on which talks of phenomenality are being grounded. It is an essential part of the nature of phenomenality that it is completely private. Here the privateness of an experience means that it cannot be experienced by more than one phenomenal subject. It is inconceivable that two distinct phenomenal subjects can experience one and the same phenomenal token at any time. I will argue for this further in the next chapter.

All these tell us one thing: “phenomenal subjecthood” clearly requires explanation because it is conceptually and ontologically connected to other essential components in the metaphysical picture adopted by PCS proponents. It is strictly tied to qualitativity, cognitive mediation and phenomenality in general. It is clear that without a clear understanding of phenomenal subjecthood in phenomenal terms first, the PP-gap cannot be adequately explained in physical terms.

All in all, a type-B physicalist might still object and say that even though phenomenal subjecthood requires to be explained fully, a “phenomenal subject” has no real ontological correspondence in the phenomenal domain other than a linguistic existence. We use the term only in sentences; it has merely a linguistic ontological status. So it has no effect on the PP-gap. Such an eliminativist view is not tenable either. But a defense of realism for phenomenal subjects is a big issue that has to be examined in its entirety. We will do that in the next chapter. For now, it should be clear that type-B physicalism in general and PCS in particular cannot adequately respond to the PP-gap without offering a persuasive thesis of phenomenal subjecthood.

CHAPTER 5

SUBSTANTIAL EXISTENCE OF PHENOMENAL SUBJECTS

We have now come to a point where we need an expansion in our understanding phenomenality of our conscious experiences in phenomenal terms. Proponents of PCS, it seems, inadequately treat phenomenality disregarding some of its essential ingredients such as the ontological nature of phenomenal subjects. This is likely to create severe problems in their theories of phenomenal concepts because treating the qualitative character of phenomenal consciousness as an isolated feature, like a third-person quality, from other aspects of phenomenality, particularly from the subjective character, either will leave some residue of the PP-gap, or present a new one. Taking the subjective character of phenomenal consciousness as if it has almost no phenomenal role, or a very limited role in exposing us to the PP-gap is the chief basis for the predicament the proponents of PCS face. What we primarily aim in this chapter, then, is to explore other aspects of phenomenal consciousness that would provide us with a good comprehension of the nature of phenomenal subjecthood. Recall that we ended the previous chapter asking the question "Does a phenomenal subject have more than a mere linguistic ontological status?" Exploring the traces of phenomenal subjects in the phenomenal domain will also enable us to see why eliminativism for phenomenal subjects is not tenable, or why phenomenal subjects have a non-linguistic ontological status. Only after such considerations can we determine whether or not PCS and physicalism in general can adequately respond to the PP-gap in physical terms.

5.1. Consciousness and Phenomenal Subjecthood

As explained in the "Introduction," the term 'consciousness' is used in many different ways. One of the reasons for these different usages is that the phenomenon of consciousness has numerous different aspects. Different employments of the term 'consciousness' is good in one sense, because such a complicated and most mysterious phenomenon needs to be analyzed from all possible aspects, and that is likely to provide a useful way of understanding the true nature of consciousness. Employing 'consciousness' in different meanings is confusing in another sense because it complicates the issue. It is then quite natural for philosophers to make a lot of distinctions regarding the nature of consciousness. What we are particularly interested in here are the central distinctions that are likely to give us some clues for the ontological status of phenomenal subjects and the nature of phenomenal subjecthood.

So, philosophers usually distinguish between a creature's being conscious and that creature's being in a conscious state (state vs. creature consciousness). They subsequently need a clear separation between conscious and non-conscious states (conscious vs. non-conscious states). This section aims to draw attention to the relation of these distinctions to the subjective character of consciousness in the ontological sense. After all, our main target, phenomenal consciousness, is not a fully discovered mental phenomenon, and the concept that the term "phenomenal consciousness" expresses is almost a primitive one whose content is quite cloudy just like the obscure content of Nagel's worn out term "what-it-is-likeness." This means that taking a step back and thinking about the subjective feature of consciousness in general is required. If so, understanding the ontological status of subjectivity within the framework of those distinctions bears a real significance for our endeavor, which will hopefully also provide an argument for the substantial ontological status of phenomenal subjects.

5.1.1. Creature vs. State Consciousness

Consider David Rosenthal's (1997) elucidation of the distinction between state and creature consciousness:

> We know in general terms what it is for a creature to be conscious; it is conscious if it is awake and sentient. So there is no special mystery about what creature consciousness is.
>
> By contrast, we lack even the sketch of a generally accepted answer to what kind of property the consciousness of mental states is. This lack of a schematic answer makes it difficult to say even what it could be for a mental state to be conscious. And the absence of even a general idea of what such consciousness consists in has led some to see it as primitive, unanalyzable property of mental states...
>
> As with creature consciousness, there is no particular mystery about introspective consciousness. Introspection is the attentive, deliberately focused consciousness of our mental states. So whatever puzzles we have about its nature are not special to it...
>
> The consciousness of states is in a certain way intermediate between the other two phenomena involving consciousness. Since only conscious creatures can be in conscious states, but the mental states of a conscious creature may not all be conscious states, state consciousness presupposes creature consciousness. Similarly, all states of which we are introspectively aware are conscious states, though not conversely. (1997, pp. 730-1)

Most philosophers agree with Rosenthal on his diagnosis that creature and introspective consciousness seem, to some extent, less problematic (but not unproblematic) than state consciousness, and that what it is for a mental state to be conscious is more obscure than what it is for a creature to be conscious. And this is one of the chief reasons why analytic philosophers produce plenty of thoughts concerning state consciousness. Another reason perhaps is the physicalist strategy to divide and conquer the conscious. Understanding and

explaining state consciousness in physical terms will in some sense ensure understanding and explaining consciousness as a property of creatures.

Nonetheless, such a strategy is circular. We cannot understand state consciousness without appealing to creature consciousness. Nor can we explain it in physical terms without first acquiring an adequate level of comprehension of the latter. Besides, the reason why state consciousness seems metaphysically problematic is not because the consciousness of a state has some properties that have nothing to do with the consciousness of the relevant creature. Rather, it is because we miss some crucial features of creature consciousness. First, on the metaphysical level, creature consciousness is ontologically prior to state consciousness. A state's being conscious ontologically depends on a creature's being conscious. In other words, without there being a conscious creature, there would not be a single conscious state in any possible world. To test the truth of this, try to imagine a robot that possesses conscious states of some sort but nevertheless lacks consciousness as a creature. The robot is in conscious states but is not creature-conscious. This is inconceivable. No one can conceive of such a robot or any system. Rosenthal, too, expresses this clearly when he says "state consciousness presupposes creature consciousness."

Second, on the semantic level, the above metaphysical claim leads us to think that there is a tight conceptual connection between the two sorts of consciousness. This conceptual connection, I claim, is a one-way dependence. Our understanding of state consciousness strictly relies on our understanding of creature consciousness. To understand the problematic features of state consciousness, which some philosophers consider primitive and unanalyzable (as does Rosenthal), we have to understand first the nature of creature consciousness (which is regarded rather clear by Rosenthal).

Third, there is another essential aspect of creature consciousness. At this point, I am aware of the fact that one might want to consider the conceptual and the metaphysical tie between state and creature consciousness as a mutual relationship as some philosophers do. According to this view, not only state consciousness relies on creature consciousness in both senses, but also a creature can be conscious only if it is at least in one conscious mental state. Some may even attempt to define creature consciousness in terms of state consciousness: creature consciousness is the totality of conscious states in a given system. This is not tenable, however. Recall the problematic status of state consciousness that Rosenthal is concerned with. There is a simple explanation for why state consciousness appears problematic whereas we intuitively think we understand better creature and introspective consciousness. State consciousness does not completely mirror the essential conceptual content of our notion of consciousness in general. A creature is conscious only if it involves a sort of subjecthood, i.e. it is the subject of some experiences. Without a creature's possessing certain features of subjecthood, a sense of I-ness, can we really make sense of Rosenthal's "awakeness" and "sentience" in the quotation I gave a few pages ago? I do not believe we can.

A mental state's being conscious seems rather puzzling for Rosenthal and others exactly because *we do not know how to detect and discern a complete and undivided experiential subjecthood in each conscious mental state*. Since we do not know how to do this, we rather wrongfully treat qualitative states as being third-person objects like table, chairs, etc. And this brings about further problems for our understanding of phenomenality of consciousness, and thus the PP-gap.

One may raise an objection here concerning the ontological priority of creature consciousness, and thus of phenomenally conscious subject, and say that in the case of animals and infants it is more troublesome to think that creature

consciousness appears first as opposed to thinking that an early version of state consciousness emerges before a full-blooded creature consciousness does. Certainly experimental data bear much importance here. But we can still think about the issue on a purely conceptual level. Here are some relevant questions: Can animals and infants entertain a perceptual state without themselves being a perceiver; or intentional states without being a thinker, believer, or knower; or emotional states without being a feeler? Well, it is not obvious that they can. Or at least if they can entertain such states, on a most basic intuitive level, we are tempted to think that without a sort of conscious subject, it does not make a clear sense to say that those states are conscious states.

Do these considerations entail that the referents of personal pronouns in the subject position of our introspective phenomenal reports have a substantial ontological status? Not yet actually. But at least we are now closer to judging whether or not there is a real phenomenal subject to which the pronoun "I" refers in those reports. The following section, I think, will provide such a firm ground needed to believe that there is.

5.1.2. Conscious vs. Non-Conscious States

Are all mental states conscious? Surely the answer requires that we know what "being conscious" means in general and for mental states in particular. But that is partly what we are after in our investigation of the phenomenal-physical relationship. And as explained in Chapter 1, it is hard to have a non-circular definition of consciousness, let alone having a canonical one. Nevertheless, we may employ a pre-theoretical meaning of the notion of state consciousness in answering such a question. We may take "being conscious" of a mental state *m*, for example, as "a subject S's being aware of *m*." Indeed, this is the best available pre-theoretical interpretation of "being conscious."

Thinking pre-theoretically then, most, including myself, believe that there are certainly non-conscious mental states. We briefly touched on this issue in "Introduction" chapter, but we have to think about it again for our present purpose. Rosenthal (1997, p. 731) gives several examples of mental states that we have to count as non-conscious. We sometimes see that some people desire something without being aware of their desires. We sometimes notice that we are sad or angry or have some thought only after a person near us makes us realize that we are in that emotional state or have that thought. There are examples of subliminal perception and peripheral vision that clearly imply our being unaware of them. We even sometimes do not notice, for a short intervals of time, minor bodily sensations such as pain or headache, but we still think that we experienced a single, persistent pain or headache during the entire period.

All those cases show that both intentional and qualitative mental states can go unnoticed by the subject. But notice that examples show us more. They specifically imply that mental states that are non-conscious at a time t_1 can be conscious at another time t_2. In fact, this is exactly how we know that those mental states were non-conscious at time t_1. Otherwise, if non-conscious mental states were non-conscious permanently, how could we possibly know at time t_2 that they were non-conscious at time t_1? If so, we can derive a significant conclusion from this as Rosenthal (1986 and 1997) does: Consciousness is not intrinsic to mental states.

Consciousness is an extrinsic property of mental states. But does this mean it is also a relational property? Surely "extrinsic" does not always mean "relational." But in the case of conscious mental entities, we do not have any other way of interpreting extrinsicality. "Being extrinsic" in a mental domain means "being relational." A simple consideration of, say, pain would reveal this feature. If at a given moment being conscious is extrinsic to a specific instance of pain in a

phenomenal domain, then it has to be the case that that instance of pain is conscious in the relational sense. Pain cannot have a non-relational extrinsic property because it cannot have a temporarily instantiated property other than a relational one by means of which we know of it. All other properties it permanently possesses are essential properties—properties that makes the pain pain.

I think we have an argument then for the substantial ontological status of phenomenal subjects:

(1) The best available pre-theoretical meaning of 'being conscious' for a mental state is "a subject's being aware of it." [Premise]

(2) There are non-conscious mental states. [Premise]

(3) A mental state *m* that is non-conscious at time t_1 can be conscious at another time t_2. [Premise]

(4) Being conscious is not intrinsic to *m*. [(2), (3)]

(5) Being conscious is a relational property of *m* holding between *m* and a subject. [(1), (4)]

(6) '*m* is conscious' means "a subject *S* is aware of *m*." [Restatement of 5]

(7) In a mental domain, there cannot be a substantial ontological *relation* between two *mental/phenomenal* entities one of which does not possess a substantial ontological status. [Premise]

(8) Therefore, the subject *S* (as a mental/phenomenal being) cannot be ontologically *vacant*. [(5), (6), (7)]

We have already discussed premise (1), (2), and (3). And premise (4), (5), and (6) are derived from the preceding ones. As to premise (7), I think its truth is intuitively clear. We cannot think of a substantial relation between a genuinely existing thing and an "absence" of the relatum in the mental/phenomenal domain, especially when this relation is a genuine subject-verb-object relation.

We therefore have to admit that the ontological status of phenomenal subjects must be substantial—not merely conceptual, linguistic, or formal.

5.2. Two Dimensions of Phenomenality: Subjectivity and Qualitativity

As we have briefly mentioned in Chapter 1, intentionality and qualia are among those hotly debated features of mentality. Most philosophers have been inclined to think of these two as metaphysically separate features of mind. But then several people rightly attempted to show these two should not be treated as isolated features of the mind, but rather must be examined treating the two as rigidly connected to each other (Gulick 1995, p. 271; Horgan and Tiension 2002). The same holds for the notion of phenomenal consciousness. It has been analyzed as if its essential aspects are not substantially related to each other. As presented earlier, subjectivity and qualitativity are the two widely accepted aspects of phenomenal consciousness. There might be other aspects of phenomenal consciousness that are yet undiscovered, but the key point here is that subjectivity and qualitativity must be strictly interrelated to each other. The reason that compels us to believe such interrelatedness lies in the fact that *one's phenomenal consciousness cannot be divided unless there is not one but more than one phenomenal domain corresponding to different phenomenal subjects*.

Having established the ontological status of phenomenal subjects as substantial items, what we are after in the next section will be two things: the nature of the tight metaphysical connection between phenomenal subjects and what they experience—qualia—and setting some desiderata concerning the nature of phenomenal subjects drawn from the preceding considerations. Let us begin with the firm metaphysical relation that phenomenal qualitativity bears to phenomenal subjects.

5.2.1. The Connection between Subjectivity and Qualitativity

We have already discussed briefly the two different construals of the notion of "experience" in Section 1.1. On the qualitative construal, depending on one's approach to access-consciousness, blindsight and similar cases are either non-conscious or access-conscious experiences that nevertheless lack phenomenal consciousness. On the phenomenal construal, however, blindsight and similar cases are not experiences at all. But anyone who thinks of an experience as "essentially involving a subjective feel" has to respond to what is happening in those extraordinary cases. Recall that the blindsight cases present some evidence that subjects somehow feel visually what is happening in the blind area without *being aware of seeing it*. If there is a non-conscious seeing, then there is a non-conscious experience. So a friend of the phenomenal construal (such as myself) has only one option in responding to these cases: Blindsight patients are not conscious of those blind areas in their visual field in spite of the experimental results that imply that the relevant qualitative features are cognitively processed by subjects.

What significance does this bear on what we have defended so far? Remember that in Section 4.4 we have said that qualitativity—which we define as the *distinctive feel* but not the *subjective feel*—ontologically presupposes a phenomenal subject. In other words, without a phenomenal subject, there cannot be a quale in any possible world just like there cannot be a shadow without the object it is the shadow of. Do blindsight cases—qualia without a subject's awareness—constitute a counter-example against this claim? My answer is "No." Blindsight cases do not present subjectivity-lacking (or subjectless) qualia, but perhaps merely present inaccessible qualitative states that are nevertheless cognitively related to some other accessible and phenomenally conscious states or experiences entertained by the subjects. There are still persisting phenomenal subjects involved in those cases. The qualia that

are claimed to be present in those cases do not exist in a total independence of a persisting phenomenal subject.

One may raise an objection here and rightly say that not only qualitativity ontologically presupposes/necessitates a phenomenal subject, but also phenomenal subjecthood in some sense requires/necessitates qualitative entities. In other words, we cannot conceive of a zombie that possesses a complete sense of I-ness but has nonetheless never experienced a single quale. This is intuitively true. As many philosophers including Hume (1793, bk. 1, pt. 4, sec. 6) and Russell (1910, p. 110) rightly pointed out, whenever we look into our mental content trying to see the pure sense of our "I-ness" in a complete isolation of anything else, we fail to catch such a thing. We certainly feel it, i.e. sort of experience our core existence or have a kind of self-experience, but we never catch it alone. It is always with something else present in the mind. It is always inherent in phenomenal qualities, but it seems it is not separable from other ingredients of phenomenality particularly from the distinctive feel. Do all these mean our "sense of I-ness"—the most likely candidate for filling in the position of phenomenal subjecthood—metaphysically necessitates the qualitative ingredients in order to exist? My answer is "no" on the metaphysical level but "yes" on the nomological level. In this actual world, I do not think we can ever observe a phenomenal subject in a complete isolation of anything else. But think about it metaphysically. I see no apparent contradiction in conceiving of a possible world in which purely bodiless conscious entities exist without experiencing anything other than the very existence of themselves. There is no logical contradiction in conceiving such entities. Here is another way to test the correctness of this intuition. Just try to remove one by one the content of your current perceptual states, bodily sensations, emotional and sensual states (desires, hopes, wishes, etc.), intentional states (thoughts, beliefs, doubts, etc.), memories, imageries, etc. When we try to do this, I believe we would not feel any

substantial decrease in our sense of "being a phenomenal subject." Even when I begin to think of myself emptying my memories one by one from my phenomenal domain, I never feel I begin to lose my sense of I-ness gradually. From this exercise, the conclusion I reach is quite simple: There is a mutual nomological necessitation between phenomenal subjects and qualia, but a one-way metaphysical necessitation from qualia to phenomenal subjects, as the inconceivability of subjectless qualia-possessing zombie example shows.

Apart from these considerations, there is another vitally significant mental phenomenon that explicitly implies a firm intrinsic relation between qualia and phenomenal subjecthood. It is the unity of (phenomenal) consciousness, which we always experience within our phenomenal domain both synchronically and diachronically. Several philosophers defended such a unity of phenomenal consciousness, including Bayne and Chalmers (2003), Bayne (2008), Dainton (2000, 2004, and 2008), and Alter (2010); and some others, including Tye (2003), rejected it. I will mostly stick to Barry Dainton's terminology on this issue because of his comprehensive work adopting a purely phenomenological approach to the topic—an approach that seems to be free from materialistic worries. So, following Dainton, we may call synchronic unity of phenomenal consciousness "phenomenal unity" and diachronic unity of phenomenal consciousness "phenomenal continuity." Let us look at the former first.

5.2.2. Phenomenal Unity

Torin Alter's precise formulation of phenomenal unity will be very useful to start here. What he calls "phenomenal unity thesis" (PUT) is as follows:

(PUT) Necessarily, when a subject has multiple states of phenomenal consciousness simultaneously, there is something it is like for the subject to be in all those states at once. (2010, p. 19)[18]

What (PUT) says can be best understood if we think about a problem that we can call the "Problem of Phenomenal Binding" (PPB). For the sake of understanding the problem, suppose a mind is nothing but a bundle of phenomenal items—or qualitative items to isolate qualitativity from subjectivity and see the relation between them—at a given moment. There is a special reason why we use the term "bundle." It implies a *non-random firm connectedness* between qualitative items. One's mind obviously cannot be like a basket full of *independently* existing and functioning qualitative items. We could not help but intuitively think that all qualitative items in a mind are rather tied to each other in a special way. There must be some special metaphysical glue that phenomenally *binds* all those qualitative items to each other at a given moment. But what is the nature of this glue? What is it to be such binding glue in the first place?

As can be easily guessed, my answer to PPB is that what binds all those qualitative items in a given phenomenal domain at a given moment is the very phenomenal subject itself. Qualia are metaphysically connected to each other through being experienced by a single unique phenomenal subject. What makes qualitative items phenomenal is nothing but this phenomenal subjecthood. A number of philosophers including Lockwood (1989), Hurley (1998), and Dainton (2000, 2004, and 2008) put forward a primitive and unanalyzable notion, "co-consciousness," as an answer to PPB. For example, Dainton's proposal for phenomenal binding is strictly based on such a notion (2000, p. 26 and sec. 3.7; 2004, p. 368; and 2008, ch. 2). Simultaneous experiences, for him, have this co-

[18] Bayne and Chalmers give an equivalent but slightly differently stated version of the phenomenal unity thesis: **"**Necessarily, any set of phenomenal states of a subject at a time is phenomenally unified." (2003, sec. 3)

consciousness relation to each other because “they are experienced together.” In one sense, co-consciousness is a *binding agent*, as he claims, that connects simultaneous phenomenal items with each other. It provides, in another sense, the primary support for (PUT). I have serious reservations about the idea that co-consciousness relationship constitutes an “agent” or a “subject” properly, but for the sake of not digressing we cannot go into details here.

It is actually just at this point that we should think about (PUT). Phenomenal items are not only unified by a special relation of co-consciousness or phenomenal subjecthood, but also the very unity between the phenomenal entities that are experienced together at a moment is itself experienced. And this is what (PUT) simply says. A phenomenal subject may experience multiple experiential states at a given time such as tasting a cup of Turkish coffee while feeling a backache and hearing a strange melody coming from the radio at the same time. Standard philosophical thinking in the analytic tradition has been that the subject experiences all these phenomenal states separately from each other. But (PUT) suggests that there is an *additional* phenomenal state that is experienced together with these individual states: experiencing the taste of Turkish coffee together with the feeling of backache and the auditory perception of a strange melody. If (PUT) is true, it is obvious that any version of physicalism must have an account of it in physical terms, and it is clear that such an account would have very limited options other than appealing to substantially existing phenomenal subjects.

5.2.3. Phenomenal Continuity

While phenomenal unity concerns the unity of consciousness *at a time*, phenomenal continuity concerns the unity of consciousness *over time*. Unity of consciousness over time can mean two things: Our experiences last in time, and we do not merely experience instant-like states, but also “enduring states,” i.e.

durations. The latter is our focus here. Think about the experience of hearing a lovely song. We do not merely hear the individual notes, but rather we experience (i) the change occurring in between the notes, and (ii) the whole sequence of the notes as a song. Or think about watching a movie in a theatre. What we see on the screen is in fact caused by series of frames passing in front of a light source. Obviously we can talk about individual phenomenology for each image if we look at them individually. But it seems that there is also an additional phenomenology of (i) the change between the images and of (ii) the whole succession of them—the movie itself. Here the phenomenon (i) implies that successive phenomenal states are phenomenally connected to each other. Consider the following thesis of phenomenology of change:

(*PG*) *Thesis of Phenomenology of Change: a phenomenal state* s_1 *at time* t_1 *and the succeeding state* s_2 *at time* t_2 *are phenomenally connected to each other.*

This is rather interesting because it seems that a phenomenal state s_1 as a "self-contained moment" is logically independent from the succeeding phenomenal state s_2, which is also a self-contained moment, and vice versa. But the entire phenomenality of our consciousness also exhibits that a change between two phenomenal states may perfectly possess additional phenomenology. It prima facie seems that these two facts contradict with each other. Nonetheless, as Dainton shows (2004, pp. 376-7) by refuting an argument based on the above logical independence, the phenomenology of a self-contained phenomenal state being experienced independently by a subject would not be the same as the phenomenology of a self-contained phenomenal state of the same type being experienced in between two other phenomenal states of the same features. In other words, the phenomenology of hearing a note being experienced

independently cannot be the same as the phenomenology of hearing the same note between hearing the preceding and the succeeding notes.

What is even more interesting and challenging is the phenomenon (ii). We can put that more precisely as follows. Let '*PCT*' stand for "phenomenal continuity thesis":

> Weak (*PCT*): *When a subject experiences some qualitatively related phenomenal states successively over a period of time, there is something it is like for the subject to be in those qualitatively related successive states altogether during that entire period of time.*

> Strong (*PCT*): *When a subject experiences some phenomenal states successively over a period of time, there is something it is like for the subject to be in those successive states altogether during that entire period of time.*

What strong (*PCT*) additionally expresses is that each sequenced moment in a duration for which there is a phenomenology that we experience does not have to be qualitatively similar to the preceding and following moments, just like what happens in the case of listening to a song or watching a movie. Think again about the example of drinking Turkish coffee while feeling a backache and listening to a strange melody from the radio. But this time suppose that you do these successively, not simultaneously. Would there be an additional phenomenology when doing these three things successively? Perhaps in some special cases there would be. For our present purpose, the whole idea is already clear: Even if we only believe in weak (*PCT*), there should be an explanation of the additional phenomenology for the whole duration or succession first in phenomenal terms and then in physical/functional terms. It is fairly clear that any plausible account of "phenomenal duration" has to appeal not only to the nature of qualitative

items in play but also to the experiential status of a substantially existing phenomenal subject who experiences the phenomenal duration in question. Otherwise, there is no reason to think that there would be an extra phenomenology for the duration or succession, which is additional to the phenomenology of each phenomenal states involved.

5.3. The Nature of Phenomenal Subjectivity

I think it is now clear from what we have said so far that: (i) a phenomenal subject is conceptually necessitated by the phenomenality of consciousness; (ii) phenomenal subjecthood is metaphysically necessitated by qualitativity; (iii) phenomenal subjects have substantial ontological status; and (iv) several mental phenomena such as phenomenal unity and continuity cannot be explained adequately without appealing to substantially existing phenomenal subjects. If all these are true, I think we have got some desiderata in an account of phenomenal subjectivity. Let us state them as follows:

(*ES*) *Desideratum for the Existence of Phenomenal Subjects*: *For every phenomenal token (for every individual experience), there is a phenomenal subject (an experiencer) who experiences that phenomenal token. A phenomenal subject is an essential component of a phenomenal domain.*

(*RE*) *Desideratum for the Relation of Experiencing*: *There is a non-linguistic and fully metaphysical subject-verb-object relation between phenomenal subjects and qualitative items.*

(*SS*) *Desideratum for the Substantiality of Phenomenal Subjects*: *A phenomenal subject must be something ontologically substantial.*

(*ES*) simply says that our phenomenal domain is not something like an empty theater whose screen displays constant flow of pictures. (*RE*) strictly suggests that there is a real subject-verb-object relation between phenomenal subjects and the experiences. The qualitative items are really experienced by a subject. The *distinctive feel* is really felt by a subject and by means of that becomes *subjective feel*. There is indeed an "experiencing." Without a subject (an experiencer) there would not be an event of experiencing. And based on these, (*SS*) declares the chief claim of this chapter: Phenomenal subjects do substantially exist; they are not present merely on a conceptual, linguistic, or formal level.

I have already argued for these desiderata in several places in this and the preceding chapters. Still, it is possible for a skeptic to doubt the truth of the above three. In the following, I will present three more mental happenings that cannot be denied without implausibility, which, at the same time, I believe, will provide further support for the belief in substantially existing phenomenal subjects.

5.3.1. Phenomenal Peculiarity

In Section 3.4, I presented two epistemological constraints: nomological constraint (*NC*) and logical constraint (*LC*). (*NC*) simply expressed one's nomological inability to *observe* a phenomenal token in a phenomenal domain from the third-person perspective. And (*LC*) expressed one's inability to *experience* someone else's phenomenal token from one's own first-person perspective. It is time now to ask why? But let us first re-formulate (*LC*) in order to see in a clearer way our deep intuition involved here:

> (*NP*) *Desideratum for Non-Publicity of Phenomenal Tokens*: *It is <u>not</u> conceivable that two distinct phenomenal subjects can experience one and the same phenomenal token at any time.*

This is really a bizarre feature of phenomenality, which we do not find in anything else. I and you can perfectly observe an external object, simultaneously or at different times, and mutually share the belief that *we are observing the one and the same object*. But we cannot even conceive that I and you experience one and the same phenomenal token simultaneously. In trying to conceive of such a case, at a moment, we might think you and I can experience one and the same phenomenal token, say a given pain token, but actually right at that moment another intuition of us says there is no longer one and the same pain token but instead two distinct pain tokens, one I personally experience, and the other you personally experience. If we try hard keeping the one-ness of the pain token, then this time the duality of phenomenal subjects begins to disappear, and we are left with a single phenomenal subject.

Why can we not conceive of such a situation? The answer cannot be that because phenomenal domains are unique. That would be circular. Phenomenal domains are unique just because of this bizarre feature. There must be some other way to explicate this peculiarity of phenomenal tokens. I do not believe we can explicate this peculiarity other than by appealing to the indivisibility of a phenomenal subject. The only plausible source of this peculiarity intuition should be our sense of I-ness. Whatever ontological status it has and whatever existence it refers to, the referent of our sense of I-ness must be what gives rise to the indivisibility feature of a phenomenal subject, and thus to the peculiarity of phenomenal tokens.

5.3.2. Phenomenal Agency

A conceptually more demanding mental event is the agency occurring in phenomenal domains. I admit that a phenomenal subject does not have to be an agent since phenomenology does not require agency. There is no difficulty in

conceiving a phenomenal subject who is merely and completely a passive observer. Experiencing does not necessitate an *active* experiencer; a passive experiencer too can perfectly experience. Nevertheless, we find such agency in our phenomenal life. On many occasions, we think that we act as an agent in the phenomenal realm. There is a good way to intuit such agency occurring purely in the phenomenal domain. Suppose you are a hard-core eliminativist concerning the existence of a substantial phenomenal subject. You have limited options for your candidate for the subject position in mental actions: It can be the body, brain, part of the brain or a function of the brain. It is true that in a lot of mental actions, one of these options might fill in the subject position intelligibly enough. But there are some mental actions whose subject position cannot be filled in intelligibly by any of those candidates. Consider the following utterances regarding some of my current mental actions:

(1) The body sitting on my chair is desiring to eat an ice-cream now.

(2) The body sitting on my chair is emotioning uneasiness now.

(3) The body sitting on my chair believes that philosophy heals.

(4) The body sitting on my chair is perceiving the brownness of my table now.

(5) The body sitting on my chair is feeling the pressure on my legs now.

(6) *I* am *introspecting*, or *self-reflecting on*, or *self-attending* to my current mental content now.

In utterances (1)-(5) above, you may perhaps intelligibly replace 'body' with the alternative options given above—'brain,' 'a part of brain,' or 'a function of brain.' But in the utterance (6), it is intelligible to replace 'I' neither with 'the body sitting on my chair' nor with a proper form of the other three alternatives. Mental actions such as introspecting, self-reflecting, and self-attending are agency-requiring actions. They not only require agents, but also require purely mental agents. How can we make sense of the sentence "The brain occupying a certain

part of the body sitting on my chair is introspecting a certain phenomenal state"? How can a brain, or a part of brain, or a function of brain fulfill this purely mental action? If these actions require mental agents, it is obvious that they require mental subjects, and since these actions involve phenomenology, they require phenomenal subjects. Let us then phrase a desideratum concerning agency as another support for the substantiality of phenomenal subjects as follows:

(*PA*) *Desideratum for Phenomenal Agency: There is true phenomenal agency: At least in some cases a phenomenal subject acts as an agent.*

5.3.3. Phenomenal I-ness

Let us start with the case of perception. Think about the perceiver-perception distinction. We are perceiving a lot of things right now, and we have perceived quite a number of many other things starting from this morning and yesterday. In fact, what we mostly do every day is to perceive things. And we think that we are the perceiver of all these perceptions. There is the perceiver that is "we" and the perceptions that are perceived by us. Nevertheless, there is also this Humean idea that "we," phenomenally speaking, are nothing but a collection of perceptions. Most philosophers agree with Hume on the observation that when we pay attention to what is going on in our mental life, we observe nothing but current perceptions—which, for Hume, include emotions and bodily sensations—and memories of past perceptions. If we try to catch what we call "self," we do not catch anything other than perceptions (1739, bk. 1, pt. 4, sec. 6). From this observation, Hume concluded that there is no "self" that is additional to the perceptions occurring in one's mental life. Each of us is merely a bundle of perceptions, phenomenally speaking.

There are serious problems with this Humean construal of self. Three of them pointed out by David Armstrong (1999, pp. 32-9) will be enough for our present

purposes here. First, the mind is not just a mental activity. If it were, in sound sleep, in the womb, and so on, where there is no mental activity, we would not exist. Hume believed that we indeed do not exist in such cases. But this is hard to believe, and rather counter-intuitive. Second, if we are just bundles of perceptions, then those perceptions must have independent existence (and perhaps be independent substances), which nonsensically entails that perceptual qualities, bodily sensations, intentional items, such as thoughts and beliefs, and emotions can exist outside of a mind. Third, it is almost impossible to establish convincingly the *unity* of these bundles neither synchronically nor diachronically.

Apart from these problems, I have another objection against this Humean conception of self, one that follows from the considerations we have been having so far. What is this "self" in the first place? How are we supposed to construe it? (a) Is it the mental subject (the perceiver, feeler, desirer, thinker, and so on) that experiences all of those mental happenings? Or (b) is it something more like a spiritual substance whose substratum constitutes the core of a person, and thus the identity of a person at a time and over time? Surely, it is possible for one to have in mind both of them when one use the term "self." But Hume particularly means (and targets) the latter because he wants to carry out a job uncompleted by Berkeley who attacked the Lockean defense of material substance. For Hume, not only the idea of material substance as an unknowable substratum, but also the idea of spiritual substance as an inaccessible substratum that we call "self" is wholly wrong. Our present purpose does not allow us to go into the details of this heavy historical debate, but there is something that crucially concerns us here: In the case of the Humean construal, eliminating self understood as such a spiritual substratum gives rise to eliminating mental subjects as well, which creates the problem on which my objection is based. It is that there cannot be a perceptual state without a perceiving subject. Since perceptual states involve phenomenology, there cannot be a perceptual state without a genuine

phenomenal perceiver. If there is no perceiver, then there is no perception. But when he denies the existence of a self inhering in our phenomenal domain, Hume also denies any mental/phenomenal subject as well, because for him "we" are nothing but a collection of perceptions.

Consider the issue from the perspective of developmental psychology. It is a big controversy whether an infant at early stages can be said to be a phenomenally conscious creature. From what I have defended so far, my position is very clear: On the one hand, if we think that there are genuine perceptual processes occurring in any creature, then there is necessarily a genuine perceiver, and thus the creature is phenomenally conscious. If we think that there is no genuine perceptual process involved, then there is no phenomenal subject, and thus the creature lacks phenomenal consciousness.

On the other hand, we can think of the issue the other way around. If we think that infants at early stages have a sense of I-ness, then they are phenomenally conscious. Think about Descartes' famous statement "I think, therefore I exist." What does "I" refer to in this statement? Disregarding what Descartes himself meant by "I" here—the mind itself—we can still ask whether it can be the body, the brain, a part of the brain, or a function of the brain. None of these can be the referent of "I" because it is a phenomenal report and says nothing about the physical world. Can it be a psychological being constituted by behaviors? Again the answer is "no" because what is involved in the statement is a mental/phenomenal action—thinking—and has nothing to do with psychological behaviors. Finally, can the referent of "I" be empty corresponding to nothing? No, because it is taken to be a crucial evidence for the first step of the inference to existence. I think what "I" refers to in such sentences is a phenomenal subject that fulfills the subjecthood position of the relevant phenomenal action, viz. thinking. And I take this to be a good reason to think that our sense of I-ness is an

indication of the existence of a phenomenal subject. Returning to the issue of infants at early stages, having a sense of I-ness, therefore, is meant to entail nomologically that infants actually experience and are capable of experiencing because they are the phenomenal subjects of some phenomenal events. If they have a sense of I-ness, then they are phenomenally conscious. If they do not, they lack phenomenal consciousness.

Note also the presupposition relationship between a phenomenal subject and the sense of I-ness. A phenomenal subject, as we already know, is the subject that entertains phenomenal tokens in a given phenomenal domain. Hence, it can possess the form of a human, an infant, an animal, an alien, or an inorganic being provided that it is the subject of an experience. It is basically the experiencer of an experience but not in a definitionally circular way. It underlies the proper place in a phenomenal subject-verb-object relation. The "sense of I-ness" we human beings possess, on the other hand, is something that we feel deep inside in our existence. It presupposes a phenomenal subject, however. Or it is an indication of the existence of a phenomenal subject. We do have this sense of I-ness because, I believe, we are the experiencer/phenomenal subject of experiences in the first place. But I do not think a phenomenal subject presupposes the sense of I-ness. We can intelligibly say that some animals exhibit phenomenal subjecthood, and the body of those animals house a phenomenal subject, but this does not necessitate that those animals have a sense of I-ness as well.

Some philosophers including Searle (2004, ch. 11) and Dainton (2008, sec. 8.2) employ the term "sense of self" in a similar fashion. The reason why I am not willing to adopt such a phrase is that the term "self" has a heavy baggage. Not only has it psychological connotations that distract attention, but also it has been worn out in a lot of different contexts, which misdirects philosophical intuitions.

Besides, the term itself is more oblique than the term “I-ness.” The term “I-ness” serves its philosophical purpose more directly.

Now, I agree that whenever we turn our attention inward to catch or to be acquainted with what is called “self” or what I call “I-ness,” we end up catching the sense of I-ness *with* a mental content. Catching it always in conjunction with something else does not sustain alone an eliminativist approach to the possibility of a genuine phenomenal referent for the term “I-ness.” At least it does not provide any support for denying the substantial existence of phenomenal subjects. Furthermore, even if one may take a Humean eliminativist position on the ontology of self or I-ness, I do not think one can easily deny the existence of the “sense of I-ness.” An eliminativist must at least grant that we human beings do experience such a sense. And because of the results we reached from our considerations of the ontology of a phenomenal subject, it must not be illusory: it must have a real phenomenal correspondence in the phenomenal domain. Even if it were an illusion, a Humean eliminativist would have to explain the source of this illusion. It seems that the most likely candidate for such a phenomenal correspondence is the phenomenal subject itself.

Finally, one may wonder if other possible phenomenally conscious creatures experience such a sense of I-ness? We may never know that. What we know for sure, on the other hand, is that if they have phenomenal consciousness, then there is a phenomenal subject involved in them. Although a phenomenal subject does not require such a sense of I-ness, it is reasonable to think that creatures such as infants at early stages and some animals do feel such a sense of I-ness in a *weaker* sense: They may well distinguish themselves from other things around and be aware of themselves as a distinct body, in which case their sense of I-ness will be resulting from the awareness of their bodies as a distinct thing. They, nevertheless, could not be aware of themselves as an independent inner mental

life in the *strong* sense. Let us phrase in a more succinct way what we have reached so far in this section through considerations of "self," "sense of I-ness," and their relation to a phenomenal subject:

(*SI*) *Desideratum for the Sense of I-ness: We do not know if there is a self-like entity as a spiritual or phenomenal substance, but we know that we (phenomenal subjects) experience a sense of I-ness that must have a phenomenal correspondence in our phenomenal domains and, because of this, presupposes a phenomenal subject.*

CHAPTER 6

A DILEMMA FOR PHYSICALISTS IN RESPONDING TO THE PROBLEM OF PHENOMENAL SUBJECTHOOD

Let me summarize what we have examined so far and where we are now. In Chapter 1, we have seen that phenomenal consciousness poses a serious threat against the scientific conception of the world within naturalism. Among 6 main approaches, materialism/physicalism is the most commonly embraced view of phenomenal consciousness. Although cognitive limitationism is the most tenable view, as I have argued, proponents of other views including materialism/physicalism have to offer a comprehensive and coherent story of their own in order to establish the possible truth of their theories. Accordingly, in Chapter 2, I have defended that physicalism must be a necessary thesis: Any relation type between the physical and the mental/phenomenal must be a necessary relation in the sense that there cannot be a possible world containing these two entities but lacking the relation that holds in the actual world. Hence, contingent physicalism is untenable. Furthermore, in Chapter 3, we have additionally seen that physicalism must admit that there is an unbridgeable epistemic/explanatory gap (or a PP-gap) between the phenomenal and the physical, which ultimately makes the physicalist thesis an a posteriori truth. There we left open the question of why the phenomenal-physical relation can be known a posteriori while it is a necessary relation. In Chapter 4, we specifically targeted type-B physicalism as a project to answer the above question. Answering this question means giving an explanation in physical terms for why we are faced with such an unbridgeable gap, which, in other words, would ultimately enable one to deduce the PP-gap itself from the complete physical truth *P*. Within type-B

physicalism, the phenomenal concept strategy (PCS) was our core target because proponents of this strategy aim to present a specific theory to explain the epistemic/explanatory gap in physical terms. This theory is supposed to show why the existence of the claimed ontological gap between the phenomenal and the physical cannot be concluded from the anti-physicalist arguments or the epistemic/explanatory gap. At the end of Chapter 4, however, we have seen that the strategy disregards two crucial points in the evaluation of the epistemic/explanatory gap: the substantial ontological status of phenomenal subjects and the reality of the subject-verb-object relation in the phenomenal domain. And finally in Chapter 5, I have attempted to show why and how a phenomenal subject and her relation to phenomenal entities must be substantial by appealing to some central features of phenomenal consciousness.

Here in this chapter, the remaining part of our objective is pretty much clear: under the light of what we examined in the preceding chapter regarding the metaphysical sources of phenomenal subjecthood, we should explicate the predicament of type-B physicalism and PCS in a clearer way and point to whether there is an open route to go for the proponents of PCS. I will indeed point out that the dilemma they are faced with is harder than anyone expects.

6.1. Two Models for the Metaphysical Structure of "Experiencing"

Throughout the following I will assume two things that the reader is already familiar with based on what has been established in the previous chapter: (i) Phenomenal subjects are ontologically substantial beings, and (ii) there is a substantial subject-verb-object (hereafter s-v-o) relation between phenomenal subjects and qualitative entities in a phenomenal domain. Now it is clear that if they are true, the effect of these two claims will be rather significant for physicalism. First, type-B physicalists will have some additional job to do. Recall the metaphysical position of a type-B physicalist. He admits that the phenomenal

is real; that the PP-gap is real and unbridgeable; but that the gap itself can be given an explanation in physical terms. But now the PP-gap is either bigger than how it is usually construed, or there are two other separate gaps that should also be given a physical explanation. Second, although the predicament a type-B physicalist faces is now harder to handle, proponents of PCS encounter a severer dilemma in their physical explanation of the PP-gap. To see these results clearly we should first look at the metaphysical picture of the phenomenal domain with respect to (i) and (ii) because otherwise we cannot see why the PP-gap is bigger, or why there are two other separate gaps, or why the dilemma proponents of PCS face is rather severe.

Let us think about the metaphysically basic phenomenal items in a phenomenally conscious domain again. As we have established so far, there is (1) the phenomenal (experiencing) subject; (2) there are the entities that are experienced; and (3) there is the experiential relation between these two.[19] What we really wonder is the nature of the way these three are harmonized in constituting the phenomenon of "experiencing." Notice that the classical physicalist treatment of phenomenal consciousness is based on attempts to show how what is experienced (but not the phenomenon of "experiencing") is numerically identical to something physical or functional. But this treatment is exactly what we should object to because it implicitly *makes central* merely *the items that are experienced by a phenomenal subject*. To put differently, phenomenal consciousness does not consist merely of *what is experienced* in a phenomenal domain, but rather consists of the phenomenon of "experiencing" *as a whole*. Construing phenomenal consciousness as rather impoverished in such a narrow way is the core reason why the reality of phenomenal subjecthood is commonly ignored by physicalists. If, however, the phenomenon of *experiencing*

[19] Of course there are other items like intentional ones or other types of relations like mental-to-mental causal relations and so forth in a mental domain. But we are primarily interested here in the basic items of a *phenomenal* domain, not a mental domain.

were taken as the central notion in investigating the phenomenal-physical connection, then it would be more likely to be noticed that experiencing were divided into the three metaphysical elements we have been drawing attention to, and thus philosophers would be more aware of the reality of phenomenal subjects and phenomenal s-v-o relation.

One might here warn us that we are falling into the fallacy of homunculus argument. A fallacious homunculus argument is characterized by its presuming the existence of another little mind within a mind when the target of the explanation is already the mind itself, or of another perceiver within the perceptual domain when the target of the explanation is already the perception itself, or of a conscious creature within the domain of consciousness when the explanatory target is already the consciousness itself, and so forth. Are we making such a fallacy when we are metaphysically distinguishing between a phenomenal subject, what is experienced, and the experiential relation? Far from it indeed. A phenomenal subject, as we have been construing it, is not another phenomenally conscious being within a phenomenally conscious domain. It is just an essential part of the phenomenality of consciousness. *It does not have its own phenomenality*. Phenomenality cannot be completed by any of the three phenomenal items *alone*. It is completed by all the three together. If there is really a fallacy in thinking about phenomenal consciousness, I think it is the *fallacy of construing phenomenal consciousness as a bunch of separated qualia of a mental domain—separated in that they are ripped off from the substantial phenomenal relation with phenomenal subjects*.

One might, however, further question the nature of a phenomenal subject's being in the subject position of the phenomenal s-v-o relation. If a phenomenal subject is not another conscious being in a conscious domain, how is it that the subject can be in such a subjecthood relation? What is the nature of this

phenomenal s-v-o relation anyway? Well, first of all, the question clearly assumes that a subject in a s-v-o relation has to be something *capable of having some kind of effect* on the object, if not being something possessing a kind of awareness. Having the capacity to affect the item in the object position, however, is not strictly required for the item in the subject position. Consider the below propositions that contain typical s-v-o relations:

(4) The sun melts ice.

(5) Winds shake leaves.

(6) Shadows follow objects.

It seems that the subject of (4) has the effect of changing the object's *intrinsic* properties; the subject of (5) has the effect of changing the object's *extrinsic* properties; but the subject of (6) has no effect of changing any property of the thing in the object position. Second, even if most s-v-o relations obtain in the way the relations in (4) and (5) do, this still does not mean a phenomenal subject by itself has to have its own phenomenality in order to affect what is being experienced in the experiential relation. Having an effect on something in the sense of changing its intrinsic or extrinsic properties is not strictly forbidden for a phenomenal subject in order for it to have its proper place in the phenomenon of "experiencing." Third, as to the real nature of the phenomenal s-v-o relation, I do not think anyone knows enough about that nature, and I have neither claimed to know about its nature, nor committed myself to anything specific about that nature. What I have been trying to do so far on this issue was merely to show that phenomenal subjects must exist substantially; are related to qualitative items in a unique experiential way, whose nature we are ignorant of; and are the most likely candidates to fill in the referent position of our sense of I-ness that we can hardly deny.

To repeat the question, what kind of relation is the phenomenal s-v-o relation? Does it consist in changing properties as in the case of (4) and (5), or is it an

impotent relation changing no property of the object at all as in the case of (6), or is it some other kind of s-v-o relation that we do not encounter in any other case? Indeed, we do not know enough about the nature of this phenomenal s-v-o relation, mostly because we do not know how to construe the metaphysical place of a phenomenal subject in a given phenomenal domain. Is a phenomenal subject structurally distinct from qualitative items in the sense of constitution, or is it structurally constitutive to and constituted by qualitative items? In other words, how should we understand the internal structure of "experiencing?" As it appears, options are limited. Since we divided a phenomenal state into three metaphysically basic elements—an experiencing subject, what this subject experiences, and the experiential s-v-o relation—it seems that logically we can have only two basic models: a constitutive one and a non-constitutive one. Consider the below illustrations corresponding to these two models:

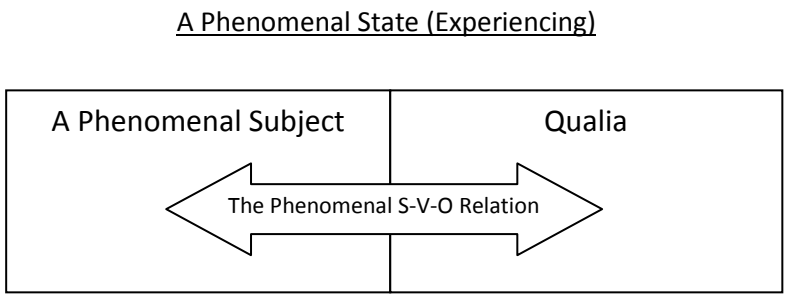

Figure 6.1: The non-constitutive model for the structural relation between a phenomenal subject and qualia

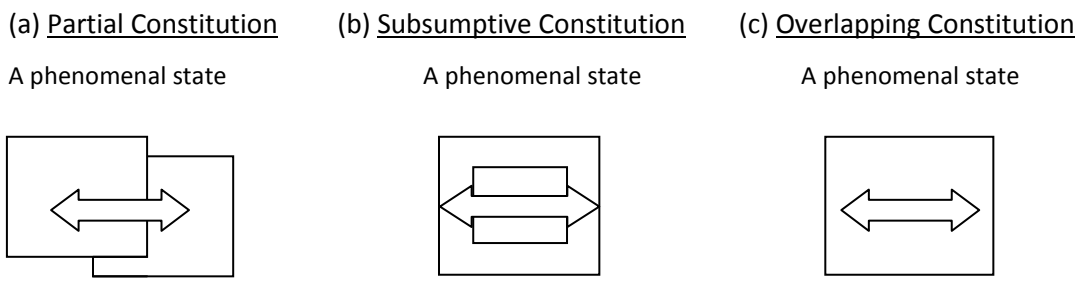

Figure 6.2: Three versions of the constitutive model for the structural relation between a phenomenal subject and qualia in a phenomenal state

In Figure 6.1, a phenomenal subject and qualitative items are structurally distinct from each other in a given phenomenal state understood as the whole event of experiencing. In Figure 6.2, on the other hand, a phenomenal subject and experiential items are constitutive of each other. It seems that there can logically be three versions of constitution on this model. A phenomenal subject can be (a) *partially* constitutive to and constituted by experiential items; or (b) *wholly* constituted by *some* of the experiential items; or (c) constituted exactly by the *totality* of experiential items at a time synchronically and in a time diachronically.

To visualize these four models, we can employ a building metaphor. For the non-constitutive model, we can think of a phenomenal state as a building possessing two main halls having some access to each other. For the partial constitution we can think of a phenomenal subject and the totality of experiential entities as two buildings sharing a room by being merged into each other. For the subsumptive constitution we can think of a phenomenal state as a building possessing one main hall (the phenomenal subject) that has doors to each of the surrounding rooms. And for the overlapping model we can think of each experiential item as the bricks of a building, and the phenomenal subject as the building itself.[20]

[20] There are authors who think of a "self" in a manner similar to the overlapping model. Dainton, for example, claims that a self is the totality of experience producing powers in a stream of consciousness (Dainton 2000 and 2008).

Apart from these four logical models, there do not seem any other possible options to reflect the structural relations between the three metaphysical elements. If so, do any of these models give us any insight to understand how the phenomenal s-v-o relation holds, which would eventually lead us to interpret the PP-gap more insightfully? I do not think they do. On the contrary, type-B physicalists and proponents of PCS confront a harder predicament in explaining the PP-gap. Let us see these in order.

6.2. Type-B Physicalism Revisited

Why should we be interested in any models illustrating the metaphysical picture of a phenomenal state in the first place? We should be because our chief concern and a physicalist's main target is the PP-gap, and it is clear that the status of the PP-gap becomes either more or less problematic depending on the correct model, whichever it is. Note also that our present aim is not to determine which model is correct, but rather to show the new status of the PP-gap, and how the core thesis of type-B physicalism and PCS is affected by the new status. For these reasons, I will be completely neutral among the four models throughout the following.

Recall the main task of type-B physicalism: Embrace the reality of phenomenal consciousness; admit that we really face an unbridgeable epistemic/explanatory gap regarding identity statements such as "a phenomenal state *m* is identical to a physical state *p*;" explain this gap in physical terms, or deduce it from the complete physical truth *P*; and do this last in such a way that we can understand why a necessary identity statement is known a posteriori. Apart from claiming that phenomenal-physical identity statements are epistemically primitive, and

thus cannot be analyzed and explained via more basic principles,[21] a type-B physicalist will have to put the following proposals on the table for a re-consideration of the problematic status of the PP-gap:

On the Constitutive Model:

(C) (Old Understanding): A phenomenal state *m*, whatever metaphysical structure it has, is identical to a physical state *p*.

(C*) (New Understanding): A phenomenal state *m**, which necessarily involves a phenomenal subject and the phenomenal s-v-o relation, is identical to a physical state *p*.

On the Non-Constitutive Model:

(Na) Within a phenomenal state, a phenomenal subject S^f is identical to a physical/functional subject S^p.

(Nb) Within a phenomenal state, the phenomenal s-v-o relation R^f (the relation of "experiencing") is identical to a physical/functional s-v-o relation R^p.

The old construal of the phenomenal-physical identity statements is something similar to (C). When we think of a phenomenal state *m* as something similar to what is represented by one of the versions in Figure 6.2, however, our construal of *m* will expand radically, and the identity statement "*m** = *p*" will be something similar to (C*). Clearly, there is a fundamental difference between (C) and (C*) in that the gap the latter presents is bigger and more problematic than the gap the former presents. On the epistemic level, for example, the PP-gap as presented by (C*) is harder to deduce from the complete physical truth *P*. On the explanatory level, one is more compelled to ask "Why and how does *p* feel like *m**?" On the

[21] I think it is not reasonable to count this claim among the options because such a claim would be adhoc and would pursue merely the goal of saving physicalism. For this point and the relevant others, see Chalmers (1999 and 2003, sec. 4 and 5).

cognitive level, one has harder difficulties in grasping the identity statement "*m** = *p.*" And on the intuitive level, one has a stronger tendency to deny the truth of "*m** = *p*."

The situation is not comforting on the non-constitutive model either. Both (Na) and (Nb) individually exhibit an additional gap or at least an unexpected separate part of the existing gap. Since a phenomenal subject and the phenomenal s-v-o relation substantially and separately exist within a phenomenal state, one is also compelled to find separately existing physical counterparts for these two within a physical state. Can there really be a physical counterpart represented by S^p and R^p for a phenomenal subject and the phenomenal s-v-o relation in a physical state? We currently do not know that. But a prior question should be "Should a physicalist really seek for a physical counterpart for S^f and R^f?" If yes, why? I think it is a human cognitive tendency to seek for counterparts in such cases. Metaphorically speaking, suppose that you are visiting a close friend in his farm. While you enter the house you have the opportunity to play with and pet a lamb in the dooryard. After spending hours inside, your friend unexpectedly comes in with a big open tray with a mass of roasted meat on the top. You surprisingly ask "What is this?" He says it is the lamb you were petting before entering the house. You are surprised, and quite naturally seek for evidence for what your friend says. Just because of that you look for the counterpart organs while you were enjoying the roasted lamb. At least you may intelligibly ask "Where is the heart of the lamb?" Of course, the phenomenal-physical relationships are unique and more mysterious than any cases that we experience in our daily life. And a physicalist does not seem to have any explanatory obligation to point out or hint at a physical counterpart for S^f and R^f in *p*. Nevertheless, we should at least understand why a physical/functional state of the brain does not have to contain separate units corresponding to S^f and R^f on the non-constitutive model while the counterpart phenomenal state involves a metaphysically separate phenomenal

subject and the phenomenal s-v-o relation. Recall that for a type-B physicalist, who strictly adopts the identity relation, metaphysically there is only one state at hand, which is completely physical. A phenomenal state is in effect another aspect of one and the same physical state. That is why the phenomenal-physical identity statements are true necessarily if they are true at all. But then it is strikingly hard to imagine a physical state necessarily involving a phenomenal subject and the phenomenal s-v-o relation within the state.

Additionally, remember the inconceivability of the subjectless qualia-possessing zombie. That inconceivability tells us that even if a neuroscientist sympathetic to type-B physicalism completes his theory of phenomenal-physical identities, something crucially essential will be felt to remain unexplained exactly because of the physicalist's qualia-centered construal of a phenomenal state. And this is plainly against the comprehensiveness principle for physicalism.

6.3. Phenomenal Concept Strategy Revisited

Once a proponent of PCS grants that either a constitutive or the non-constitutive model should be adopted in order to embrace the reality of a phenomenal subject and the phenomenal s-v-o relation, things get more complicated for her. To explain such a complication, recall the basic task of a PCS defender: Show that the PP-gap results completely from a fundamental difference between our grasping of phenomenal and physical/functional concepts; and explain this difference in physical or at least phenomenal-free terms in such a way that we can understand why a phenomenal-physical identity statement is known a posteriori. It is apparent that such a difficult task requires assuming a subject because there must be a thinking or deploying agent in order for such concepts to be deployed and in order for the unique features of phenomenal concepts to be

activated.[22] Otherwise it would be nonsensical—in the framework of PCS—to think of concepts as agent-independent entities that happen to exist on their own.

Let us think about a deploying agent a little bit. When we say "The way we deploy a phenomenal concept is different from the way we deploy a physical concept," the agent "we,"—or "I" in the first-person case—does not seem to be ontologically vacant within the framework of PCS. Then the question is "Is it an agent existing *inside* or *outside* the relevant phenomenal domain?" It should not be outside because if it were, it would mean that there is something more serious and problematic outside a phenomenal domain—the thinking or deploying subject. Our private sense of I-ness or the feeling of being a subject is already an essential part of the phenomenality of consciousness. If the thinking or deploying agent were not part of the phenomenality, then additional metaphysical questions would arise regarding the ontological status of this thinking or deploying agent. And no proponents of PCS would want to face such metaphysical questions.

Nevertheless, one might still tend to presume that the thinking or deploying agent exists outside the phenomenal domain without being involved in the phenomenality of consciousness. It is, one might further claim, the body, or the brain, or a part of the brain, or a function of the brain. Now, put one of these alternatives in the subject position of the above sentence: "The way the body/brain/part of brain/function of brain deploys a phenomenal concept is fundamentally different from the way it deploys a physical concept." Problems immediately arise. First what is to be explained by a theory is assumed to be

[22] As I stated in Chapter 4, I stay neutral to the debates regarding the ontological status of concepts, their acquisition, possession, and identity conditions. But nevertheless I prefer to be realist at least about their existence. And that requires, at a minimum, taking concepts as simple constitutients of thoughts.

already explained within the theory according to this construal. What we are after is the ontological nature of that deploying subject. We cannot presume at the beginning that it is physical. Second, according this construal, phenomenal entities are treated like third-person objects. But recall our discussion of phenomenal peculiarity in Section 5.3.1. A phenomenal token of mine cannot be experienced by anyone other than me exactly because I am directly involved in the phenomenality of that token. But if I as a physical subject experience that phenomenal token, why cannot others do the same as physical beings? Third, it contradicts with our intuition of what-it-is-likeness. "What is it like to be a bat?" is an intelligible question when being a bat is understood as being a phenomenal subject, not being a physical body, brain, part of a brain, or a function of a brain. When we try to conceive of us being a bat, we are not supposed to conceive of ourselves in the physical body of a bat carrying our own phenomenal subjecthood.

If the thinking or deploying subject exists *within* the phenomenal domain, it follows that it has to have some phenomenology. Indeed, it has to be the phenomenal subject itself. But then a proponent of PCS has to choose one of the models we discussed above. Unfortunately, on neither models, PCS seems to be successful. Here I will take the version of PCS that is based on the involvement of the experience within the phenomenal concept as the central version because of its explanatory power that we explicated a little bit in Section 4.2, and because of the other version's drawbacks in explaining the special intimate cognitive connection between a phenomenal subject and what she experiences.

Recall the constitutional account of PCS, according to which when a subject deploys a phenomenal concept, the referent (the experience itself) is constitutive to the concept. A phenomenal concept is fundamentally different from other sorts of concepts only in that the referent itself is wholly or partially contained

within the concept. Now on the constitutive model of phenomenal subjects, it is rather unintelligible to think that a phenomenal subject *S* deploys the concept *c* of a phenomenal state *m* while *S* itself is an essential part of *m* (and related to the remaining part of *m* by the phenomenal s-v-o relation), and *c* is wholly or partially constituted by *m*. How can *S* deploy *c* of *m* while *m* is contained somehow in *c*, and *S* is contained in *m*? To give a concrete example, when "I" deploy the concept of <my current sore throat>, this concept within itself involves wholly or partially "my current sore throat." But remember that the deploying agent is itself, on the constitutive model, an essential part of "my current sore throat," and I, the deploying agent, currently am in the phenomenal s-v-o relation with "my current sore throat" (I experience it). How can "I" be a part of something (*m*) that is already the total ingredient of something (*c*) possessed or deployed by "me" while I am in the phenomenal s-v-o relation with *m*? There seems to be two fundamental problems here: (1) the unintelligibility of the phenomenal s-v-o relation between a phenomenal subject and what she experiences while they are constitutive to each other and (2) the unintelligibility of containment of a whole within some part of it—the containment of the deploying agent within what she deploys (the phenomenal concept). The first nonsensicality arises from the constitutive model of phenomenal subjecthood. The second one arises from the constitutional account of phenomenal concepts itself. I currently see no hope in dissolving these problems as long as one adopts the constitutive model for the ontology of phenomenal subjects and the involvement of an experience within a phenomenal concept.

On the non-constitutive model, the predicament that the constitutional account of phenomenal concepts faces is not any easier to overcome. It is true that since a phenomenal subject is separate from qualitative entities within a phenomenal state, it is easier to establish a meaningful s-v-o relation between the subject and what she experiences in the qualitative sense. Hence problem (1) seems to be

avoided on this model. But in turn, this separation produces an additional problem on this model. When a phenomenal subject *S* deploys a phenomenal concept *c* of a phenomenal state *m*, *c* does not seem to be a concept representing the phenomenal subject *S* as well; rather it seems to represent only the qualitative aspect of *m*. The concept *c* seems to representationally exclude *S*. Because I am not part of my current pain state *m*, my concept *c* of *m* seems to represent only the qualitative aspect of *m*. If it is so, the phenomenology of subject *S* is not explained on this model, or more correctly, the PP-gap between the phenomenology of subject *S* and its physical correspondence is left unexplained. There remains an additional gap and the constitutional account of PCS is not a comprehensive physicalist theory. This is exactly what I meant when I said that PCS is a qualia-centered strategy in that its proponents merely target the qualitative aspects of a phenomenal state.

One may raise an objection here and say that it is not strictly required for *c* to also represent *S* since the subject *S* can deploy a separate phenomenal concept *c* of the subjective aspect of *m*: the phenomenal concept c^S of "being the phenomenal subject of my current pain state *m*," for example. So concept *c* represents the qualitative aspect of *m*, and concept c^S represents the aspect of phenomenal subjecthood inhering in *m*, and thus *m* is completely represented. But in this case the problem (2) above arises within the framework of the constitutional account. The phenomenal subject *S* deploys the concept c^S of herself while she is *constitutionally* a part of c^S. This is, as I said above, rather unintelligible. How can "I" deploy or possess the concept of "being the phenomenal subject of my current pain" that is already *constituted* by "me"? How can "I" be the total ingredient of something that I possess, and is completely separate from "me"?

The status of the phenomenal s-v-o relation is somewhat similar to that of phenomenal subjecthood on this non-constitutive model. Since a phenomenal subject is separate from what she experiences within a phenomenal state, as we have seen, the phenomenal s-v-o relation seems to hold without a problem between the subject and the experiential entity on this model. But in turn, there seems to be a second additional PP-gap between this s-v-o relation and its physical counterpart. It is natural for a physicalist to think that there must be a physical relation corresponding to the phenomenal s-v-o relation—in fact the latter is itself that physical relation. And a proponent of constitutional account of PCS will naturally say that the concept c^R is constituted by the phenomenal s-v-o relation itself. But a relation is not something independent from its relata. It carries along its relata. If the actual relation is contained within the concept c^R, it means the phenomenal subject *S* is partly contained in c^R too. And such a conception faces the same containment problem mentioned in the preceding paragraph.

It seems that PCS, particularly the constitutional account, faces a severe dilemma. On the constitutive model, both the unintelligibility of the containment of phenomenal subjects within phenomenal concepts, and the unintelligibility of the phenomenal s-v-o relation presents a serious problem. On the non-constitutive model, the phenomenal s-v-o relation seems to hold intelligibly, but the containment of phenomenal subjects within phenomenal concepts remains a serious problem, both in the case of concepts of pure phenomenal subjecthood, and in the case of the phenomenal s-v-o relation.

6.4. Paths to Go

Having clarified the difficulties that type-B physicalism and PCS encounter in explaining the PP-gap with respect to substantiality of phenomenal subjects and the phenomenal s-v-o relation, one might still attempt to see if there is any

possible way to avoid the problems. There seems to be none actually. Denying the substantial existence of phenomenal subjects or the s-v-o relation would be denying our most basic intuitions concerning "experiencing," which include phenomenal unity, continuity, peculiarity, agency, and the sense of I-ness. And once the substantiality of the above two is granted, the severer character of the PP-gap cannot be denied any more.

Nevertheless, one should not conclude from the above that this new character of the PP-gap cannot be overcome. There must be some way to handle the bigger PP-gap or the additional ones within the framework of type-B physicalism. I think the first step for any physicalist here should be to work out phenomenal subjecthood and the s-v-o relation completely in phenomenal terms. This is required especially for three reasons. First, a type-B physicalist's classic treatment of "phenomenal state" is qualia-centered. When it comes to the phenomenal-physical relation, type-B physicalists wrongly take a phenomenal state to merely involve some qualitative character of a mental state, understanding "experience" thoroughly within the qualitative construal. As I have stated earlier in several places, however, that approach is mistaken and what we must be after should be the nature of "experiencing," not that of "experience" understood merely within the qualitative construal. The phenomenon of "experiencing," as I propose to understand it, involves more than "experience" in that it includes both substantially existing phenomenal subjects and the phenomenal s-v-o relation holding between the subjects and what they experience.

Second, as a symptom of the mentioned mistaken approach resulting from the qualitative construal, type-B physicalists treat qualitative entities as if they are third-person entities like tables, chairs, and so on. This also results in the ontological separation of the phenomenal subject from the phenomenon of experiencing, and thus from one half of its components, the qualia. My pain state

m, however, is not something that can be entertained by other people as well, or by me as something outside my phenomenal domain. There is a sense that my pain is part of *me*, or *I* am part of my pain. And the reality of "paining" (my experiencing the pain in me) is so strong that no one else can intervene into that unbreakable experiencing relation. I cannot be something outside the reality of my phenomenal states. Those phenomenal states in some sense must be what constitute me.

Third, many philosophers have expressed their concern that there are *more* to experience than merely being physical or having some functional roles (McGinn 1991; Searle 1992; Strawson 1994; Chalmers 1996; Dainton 2008). This is true actually, and is the chief reason why we face such an epistemic/explanatory gap, which is very hard to explain purely in physical terms. I believe, and have attempted to show in this thesis, that phenomenal subjecthood and the experiencing relation are among those that are meant by "more."

A type-B physicalist, however, might think that a physicalist story of phenomenal consciousness does not have to be bothered by the substantiality of phenomenal subjects and the s-v-o relation. As we learn more and more about the physical aspects of phenomenal consciousness, theories will come up offering an acceptable physical/functional basis for the above two, and we will have a deeper understanding of the nature of the phenomenal-physical relation. So there is no absolute requirement for a type-B physicalist to work out the nature of phenomenal subjecthood in detail in phenomenal terms. Such an approach would be mistaken too, because without exactly knowing what it is to be a phenomenal subject and what it is to be in a phenomenal s-v-o relation in the first place, it is almost impossible to offer a plausible physical/functional basis for such an existence and relation. Unless we grasp the notion of phenomenal subjecthood in purely phenomenal terms with an adequate (if not full) comprehension, there

seems to be no way to architect a comprehensive and coherent account of the relation between the phenomenal subjecthood and the corresponding physical/functional processes.

As a final statement, I would like summarize what I have been defending so far in the following two sentences:

(1) *Without an adequate understanding of the ontology of a phenomenal subject, its subject-verb-object relation to qualitativity, and the metaphysical character of "phenomenal I-ness," we cannot fully comprehend the phenomenality of consciousness.*

(2) *Without comprehending the phenomenality of consciousness fully in phenomenal terms, it is not likely that we can understand the phenomenal-physical connection in a satisfactory way.*

BIBLIOGRAPHY

Alter, Torin (2010). “A Defense of the Necessary Unity of Phenomenal Consciousness.” *Pacific Philosophical Quarterly* 91:19–37.

Armstrong, David M. (1999). *The Mind-Body Problem: An Opinionated Introduction*. Boulder and Oxford: Westview Press.

Aydede, Murat & Güven Güzeldere (2005). “Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness.” *Nous* 39 (2):197-255.

Ball, Derek (2009). “There Are No Phenomenal Concepts.” *Mind* 118 (472):935-62.

Bayne, Tim (2008). “The Unity of Consciousness and the Split-Brain Syndrome.” *The Journal of Philosophy* 105 (6):277-300.

Byne Tim & David J. Chalmers (2003). “What is the Unity of Consciousness?” In *The Unity of Consciousness: Binding, Integration, Dissociation*, edited by A. Cleeremans. Oxford: Oxford University Press.

Bealer, George (1994). "Mental Properties." *Journal of Philosophy* 91:185-208.

Berkeley, George (1710). *Principles of Human Knowledge*. London: Penguin, 2004.

Block, Ned (1995). “On a Confusion about the Function of Consciousness.” *Behavioral and Brain Sciences* 18:227-47.

Block, Ned (2006). “Max Black's Objection to Mind-Body Identity.” *Oxford Review of Metaphysics* 3.

Balog, Katalin (1999). “Conceivability, Possibility, and the Mind-Body Problem.” *Philosophical Review* 108:497-528.

Balog, Katalin (2009). “Phenomenal Concepts.” In *The Oxford Handbook of Philosophy of Mind*, edited by McLaughlin and A. Beckermann. New York: Oxford University Press.

Balog, Katalin (forthcoming). “Acquaintance and the Mind-Body Problem.” In *The Mental, the Physical* edited by Christopher Hill & Simone Gozzano. Cambridge: Cambridge University Press.

Carruthers, Peter (2000). *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

Carruthers, Peter (2003). “Phenomenal Concepts and Higher-Order Experiences.” *Philosophy and Phenomenological Research* 68 (2):316-36.

Chalmers, David J. (1996). *The Conscious Mind*. New York: Oxford University Press.

Chalmers, David J. (1999). “Materialism and the Metaphysics of Modality.” *Philosophy and Phenomenological Research* 59:473-93.

Chalmers, David J. (2002). “Does Conceivability Entail Possibility?” In *Conceivability and Possibility*, edited by T. Gendler and J. Hawthorne, 145-200. Oxford: Oxford University Press.

Chalmers, David J. (2003). “Consciousness and Its Place in Nature.” In *Blackwell Guide to the Philosophy of Mind*, edited by S. Stich and T. Warfield. Blackwell.

Chalmers, David J. (2007). “Phenomenal Concepts and the Explanatory Gap.” In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. New York: Oxford University Press.

Chalmers, David J. (2009). “The Two-Dimensional Argument Against Materialism.” In *The Oxford Handbook of the Philosophy of Mind*, edited by McLaughlin and A. Beckermann. New York: Oxford University Press.

Churchland, Paul S. (1996). “The Hornswoggle Problem.” *Journal of Consciousness Studies* 3 (5-6):402-408.

Dainton, Bary (2000). *Stream of Consciousness: Unity and Continuity in Conscious Experience*. New York: Routledge, Taylor & Francis Group.

Dainton, Bary (2004). “The Self and the Phenomenal.” *Ratio* 17 (4):365-89.

Dainton, Bary (2008). *The Phenomenal Self*. New York: Oxford University Press.

Davidson, David. (1970). “Mental Events.” In D. Davidson, *Essays on Actions and Events*, 207-223. Oxford: Oxford University Press.

Descartes, Rene. (1637). *Discourse on the Method for Conducting One's Reason Well and for Seeking Truth in the Sciences*, translated by Donald A. Cress, (1998). Indianapolis: Hackett Publishing Company.

Descartes, Rene. (1641). *Meditations on the First Philosophy*, translated by Donald A. Cress. In *Classics of Western Philosophy*, sixth edition, edited by Steven M. Cahn, (2002). Indianapolis: Hackett Publishing Company.

Flanagan Owen J. (1992). *Consciousness Reconsidered*. Cambridge, MA: The MIT Press.

Feigl, H. (1958). "The "Mental" and the "Physical."" In *Concepts, Theories and the Mind-Body Problem*, edited by H. Feigl, M. Scriven, and G. Maxwell. Minneapolis: Minnesota Studies in the Philosophy of Science, Vol. 2.

Gulick, Robert V. (1995). "How Should We Understand the Relation between Intentionality and Phenomenal Consciousness." Philosophical Perspectives 9:271-289.

Hill, Christopher S. (1997). "Imaginability, Conceivability, Possibility, and the Mind-Body Problem." *Philosophical Studies* 87 (1):61-85.

Hill, Christopher S. & Brian P. McLaughlin (1999). "There Are Fewer Things in Reality than Are Dreamt of in Chalmers' Philosophy." *Philosophy and Phenomenological Research* 59:445-54.

Horgan, Terence E. (1982). "Supervenience and Microphysics." *Pacific Philosophical Quarterly* 63:29-43.

Horgan, Terence E. (1993). "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind* 102:555-86.

Horgan, Terence E. and John L. Tiension (2002). "The Intentionality of Phenomenology and The Phenomenology of Intentionality." In *Philosophy of Mind: Classical and Contemporary Readings*, edited by David J. Chalmers, 520-533. New York: Oxford University Press.

Hume, David (1739). A treatise of Human Nature. Reprinted (1969), edited by Ernest C. Mossner. London: Penguin Books.

Hurley, Susan L. (1998). *Consciousness in Action: Clarifications*. Cambridge, MA: Harvard University Press.

Jackson, Frank (1982). “Epiphenomenal Qualia.” *Philosophical Quarterly* 32 (April):127-36.

Jackson, Frank (1986). “What Mary didn't Know.” *Journal of Philosophy* 83:291-5.

Kim, Jaegwon (2003). “The American Origins of Philosophical Naturalism.” *Journal of Philosophical Research*, APA Centennial Volume:83-98.

Kim, Jaegwon (1993). *Mind and Supervenience: Selected Philosophical Essays*. Cambridge: Cambridge University Press.

Kim, Jaegwon (1998). *Mind in a Physical World*. Cambridge, MA: Bradford Books/The MIT Press.

Kripke, Saul (1971). “Identity and Necessity.” in *Identity and Individuation*, edited by M.K. Munitz, 135-64. New York: New York University Press.

Kripke, Saul (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Levin, Janet (2007). “What is a Phenomenal Concept?” In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. New York: Oxford University Press.

Levine, Joseph (1983). “Materialism and Qualia: The Explanatory Gap*.” Pacific Philosophical Quarterly* 64 (October):354-61.

Levine, Joseph (1998). “Conceivability and the Metaphysics of Mind.” *Nous* 32 (4):449-80.

Levine, Joseph (1999). “Conceivability, Identity and the Explanatory Gap.” In *Towards a Science of Consciousness: The Third Tuscon Discussions and Debates*, edited by S. Hameroff, A. Kaszniak and D. Chalmers. Cambridge, MA: The MIT Press.

Levine, Joseph (2001). *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press.

Levine, Joseph (2007). “Phenomenal Concepts and the Materialist Constraint.” In *Phenomenal Concepts and Phenomenal Knowledge: Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. New York: Oxford University Press.

Levine, Joseph & Kelly Trogdon (2009). “The Modal Status of Materialism.” *Philosophical Studies* 145 (3):351-62.

Loar, Brian (1990). “Phenomenal states.” *Philosophical Perspectives* 4:81-108.

Loar, Brian (1997). “Phenomenal States.” (Second Version). In *The Nature of Consciousness*, edited by N. Block, O. Flanagan and G. Güzeldere, 597-616. Cambridge, MA: The MIT Press.

Lockwood, Michael (1989). *Mind, Brain and the Quantum*. Oxford: Blackwell Publishers.

Lycan, William G. (2003). “Perspectival Representation and the Knowledge Argument.” In *Consciousness: New Philosophical Perspectives*, edited by Quentin Smith & Aleksandar Jokic, 384-95. New York: Oxford University Press.

McGinn, Colin (1989). “Can We Solve the Mind-Body Problem?” *Mind* 98 (391):349-66. Reprinted (1997) in *The Nature of Consciousness*, edited by N. Block, O. Flanagan and G. Güzeldere, 529-542. Cambridge, MA: The MIT Press.

McGinn, Colin (1991). *The Problem of Consciousness: Essays towards a Resolution*. Oxford: Blackwell Publishing.

Melnyk, Andrew (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. New York: Cambridge University Press.

Moore, George E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.

Nagel, Thomas (1974). “What Is It like to Be a Bat?” *Philosophical Review* 83 (October):435-50.

Nida-Rümelin, Martin (2007). “Grasping phenomenal properties.” In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. New York: Oxford University Press.

O'Dea, John (2002). “The Indexical Nature of Sensory Concepts.” *Philosophical Papers* 32 (2):169-81.

Papineau, David (2002). *Thinking about Consciousness*. Oxford: Oxford University Press.

Papineau, David (2007). “Phenomenal and Perceptual Concepts.” In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, edited by T. Alter and S. Walter. New York: Oxford University Press.

Papineau, David (2008). “The Explanatory Gap and Dualist Intuitions.” In *Frontiers of Consciousness: Chichele Lectures*, edited by L. Weistrantz and M. Davies, 55-68. New York: Oxford University Press.

Perry, John (1979). “The Problem of the Essential Indexicals.” *Nous* 13:3-21.

Perry, John (2001). *Knowledge, Possibility, and Consciousness*. Cambridge, MA: The MIT Press.

Place, U. T. (1956). “Is consciousness a brain process?” *British Journal of Psychology* 47:44-50.

Putnam, Hillary (1975). “The Meaning of ‘Meaning.’” *Minnesota Studies in the Philosophy of Science* 7:131-93.

Putnam, Hillary (1967). “Psychological Predicates.” In *Art, Mind, and Religion*, edited by W.H. Capitan and D.D. Merrill, 37-48. Pittsburgh: University of Pittsburgh Press.

Rey, Georges (1998). “A Narrow Representationalist Account of Qualitative Experience.” *Philosophical Perspectives* 12:435-58.

Rosenthal, David M. (1986). “Two Concepts of Consciousness.” *Philosophical Studies* 49:329-59.

Rosenthal, David M. (1997). “A Theory of Consciousness.” In *The Nature of Consciousness*, edited by N. Block, O. Flanagan and G. Güzeldere, 597-616. Cambridge, MA: The MIT Press.

Russell, Bertrand (1910). “Knowledge by Acquaintance and Knowledge by Description.” *Proceedings of the Aristotelian Society* 11:108-28.

Russell, Bertrand (1912). *The Problems of Philosophy*. Reprinted (1998) New York and Oxford: Oxford University Press.

Searle, John R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: The MIT Press.

Searle, John R. (2004). *Mind: A Brief Introduction*. New York and Oxford: Oxford University Press.

Shoemaker, Sydney (2007). *Physical Realization*. New York: Oxford University Press.

Smart, J. J. C. (1959). “Sensations and Brain Processes.” *Philosophical Review* 68:141-56.

Stoljar, Daniel (2001). “Two Conceptions of the Physical.” *Philosophy and Phenomenological Research* 62:253-81.

Stoljar, Daniel (2005). “Physicalism and Phenomenal Concepts.” *Mind and Language* 20 (2):296-302.

Strawson, Galen (1994). *Mental Reality*. Cambridge, MA: The MIT Press.

Strawson, Galen (2000). “Realistic Materialist Monism.” In *Toward a Science of Consciousness III* edited by S. Hameroff, A. Kazsniak, and D. Chalmers. Cambridge, MA: The MIT Press.

Sturgeon, Scott (1994). “The Epistemic View of Subjectivity.” *Journal of Philosophy* 91 (5):221-35.

Tye, Michael (2003). “A Theory of Phenomenal Concepts.” In *Minds and Persons*, edited by A. O'Hear. Cambridge: Cambridge University Press.

Tye, Michael (2009). *Consciousness Revisited*. Cambridge, MA: The MIT Press.

APPENDIX A: CURRICULUM VITAE

PERSONAL INFORMATION

Name, Surname: Murat ARICI
Nationality: Turkish (TC)
Date and Place of Birth**:** 16.03.1974, Konya/TURKEY
Email: pd.muratarici@gmail.com

EDUCATION

· 2003-2011: (Turkey, Ankara) Middle East Technical University (METU), Philosophy Department – Ph.D. degree in philosophy (philosophy of mind). Supervisor: Assoc. Prof. Erdinç Sayan.

· 2008-2009: (USA, North Carolina) University of North Carolina, Philosophy Department – Research on the dissertation topic. Supervisor: Prof. William Lycan.

· 2003-2006: (USA, Florida) University of Florida, Philosophy Department – Graduate studies for Ph.D. in philosophy.

· 2000-2003: (Turkey, Ankara) Middle East Technical University, Philosophy Department – Master's degree in philosophy (epistemology). Thesis Title: *The Connection between Justification and Truth*. Supervisor: Prof. David Grünberg.

· 1994-1998: (Turkey, İstanbul) Marmara University, Divinity Faculty – Bachelor's degree in philosophy/Islamic theology.

ACADEMIC EMPLOYMENT

· 2006-2011: (Turkey, Ankara) METU, Philosophy Department - Research assistant.

· 2003-2006: (USA, Florida) University of Florida, Philosophy Department – Teaching assistant.

· 2000-2003: (Turkey, Ankara) METU, Philosophy Department – Research assistant.

· 2010-Present: (Turkey, Ankara) MEPHIS: METU Philosophical Studies, a newly established electronic journal of philosophy under the host of METU, Philosophy Department. Link: http://www.metuphil.com/index.php?loc=board – Editor and one of the founders.

TEACHING EXPERIENCE

· 2007-2008: Teaching 235 Introduction to Logic for the fall semester at METU (Turkey, Ankara).
· 2006-2007: Teaching 407 philosophy of mind course for the spring semester at METU (Turkey, Ankara).
· 2003-2006: Assisting in Teaching 2100 Introductory Logic course for 6 semesters at University of Florida (USA, Florida).
· 2003-Summer: Teaching 302 Systematic Philosophy for the summer semester at METU (Turkey, Ankara).
· 2002-2003: Teaching 235 Introduction to Symbolic Logic for 3 semesters at METU (Turkey, Ankara).
· 2000-Summer: Teaching history of Turkish revolution to candidates of municipality vice presidents in a course arranged by Foundation of the Civil Service in Çanakkale/TURKEY.

ACADEMIC WORKS

· 2011: Arıcı M. (2011). "Can a Non-empirical and Non-Rational Way of Forming Belief Be a Sustainable Epistemic Reaction to Epistemic Barriers?" In *Conference Proceedings*, *the Asian Conference on Ethics, Religion, and Philosophy 2011*, 206-14. Osaka, JAPAN: Iafor (The International Academic Forum).
· 2010: Arıcı M. (2011). "Ideological Freedom, the Problem of Axiological Tension, and the Grounds for Respecting Others' Ideologies." In *First International Philosophy Congress "Liberty, Equality, and Fraternity,"* 591-8. Bursa, TURKEY: ASA Kitabevi.
· 2009: Arıcı M. (2009). *A Theory on the Truth Connection: An Agent-Dependant and Inter-Subjective Approach to the Connection between Justification and Truth*. Saarbrücken, GERMANY: VDM Verlag Dr. Müller.
· 2002: Proof Reading and Correction for Grünberg T. (2001). *Symbolic Logic*. Ankara: METU Press.
· 2002: Proof Reading and Correction for Grünberg T. (2000). *Sembolik Mantık El Kitabı, Volume I-II-III* [Handbook of Symbolic Logic]. Ankara: METU Press.

AREAS INTERESTED

· Philosophy of Mind, Epistemology, Metaphysics, Philosophy of Language, Metaphilosophy

COMPETENCIES AND SKILLS

· **Foreign Languages**
English: Advanced level in all four skills. Fluent in speaking.
Arabic: Advanced level in reading. Intermediate level in writing and speaking.
· **Hobbies:** Paragliding and soccer.

APPENDIX B: TURKISH SUMMARY

1. Giriş: Fenomenal Bilinç Neden Gizemli Bir Bilmece?

Yüzyıllar boyunca filozoflar tabiat ve insan doğası hakkında insan zihninin cevaplamaya güç yetiremediği binlerce soru sordular. Bunlardan pek çoğu en azından günümüz düşünürlerinin kavrayışı ile çözülmüş gözüküyor. Fakat bazıları hala kafa karıştırmaya devam ediyor. Modern bilimin de desteğiyle insanoğlu bu gün etrafımızdaki maddi dünyanın gizemli gözüken pek çok yönünü kavramış durumda. Bu da doğal olarak ona, evreni kavramak konusunda haklı bir kendine güven kazandırdı. Fakat Fenomenal bilinç bu güveni zedeleyen en önemli bilmecelerden bir tanesi. Problem oldukça basit: Nasıl olur da sadece maddi yapısını ve bu maddi yapının fiziksel fonksiyonlarını gözlemleyebildiğimiz insan beyni, fenomenal bilinç gibi göz alıcı ve gizemli bir özelliğe sahip olabilir?

Fenomenal bilincin tatmin edici bir tanımını vermek pek mümkün gözükmüyor. Çünkü böyle bir tanım, üzerinde yeterince uzlaşılmamış bir bilinç teorisini içermek zorunda kalır. Fakat yine de bazı betimleyici terimler günümüz düşünürleri tarafından kullanılmakta. “Deneyimsel nitelik,” “öznel his,” “ham (duyusal) his” ve Thomas Nagel’ın meşhur, zihni veya zihinsel durumları deneyimlemenin “ne gibi bir şey olmaklığı” terimi bunlar arasında sayılabilir. Burada “bilinç” ile “fenomenal bilinç” arasındaki kavram farkına dikkat etmek gerekir. Örnek vermek gerekirse beni, bu satırları yazdığım bilgisayarı ve önünde oturduğum masayı ele alabiliriz. Masa, açıkça bilinci olmayan bir nesne. Ben, açıkça fenomenal bilinci olan bir varlığım. Denilebilir ki bilgisayar insan zihninin bilinçli işlevlerine benzer bazı işlevleri yapma kapasitesine sahip. Dolayısı ile bilgisayar bu anlamda bilinçli bir aygıt olarak görülebilir. Bu bakış açısı doğru ya da yanlış olabilir. Önemli olan burada bilgisayar ve benzeri aygıtların fenomenal bilince sahip olmadığını bilmek.

Çünkü bilinçli bir varlık olarak ben "deneyimleme" kapasitesine sahibim. Önümdeki bilgisayar ise açıkça bu kapasiteden yoksun bir aygıt.

O halde modern bilim ve felsefe bu deneyimleme özelliğine sahip fenomenal bilinci kendi terimleriyle nasıl açıklamalı? Burada problemi iki açıdan ele almak mümkün: metafiziksel ve epistemolojik. Metafiziksel açıdan problem aslında gayet aşikâr: Bu gün pek çok bilim insanı ve düşünür evrende gözlemlenebilir her olgunun içerdiği unsurların fiziksel ve biyolojik bilimlerin belirlediği en temel elementler listesinde yer aldığını ya da alması gerektiğini düşünmekte. Başka bir deyişle bu listenin kapsadığı en temel ontolojik elementlerin dışında ve ötesinde evrende var olan bir şey bulunmamaktadır. Ve doğal olarak bu elementler sadece materyal/fiziksel elementlerdir. Soru şu: Fenomenal bilinç bu listede yer alıyor mu? Yani o da materyal/fiziksel midir? İlk anda neden olmasın denilebilir. Fakat pek çok felsefeci fenomenal bilincin, adı geçen listedeki elementlerin arasında yer alamayacak nitelikler içerdiğini düşünmekte. Bu özellikler fenomenal bilincin öznel ve niteliksel yönleriyle ilgili özelliklerdir.

Epistemolojik açıdan durum pek de farklı değil. Evreni anlamak için bilim, gözlem ve deney gibi meşru kabul ettiği bir takım yöntemler kullanır. Eğer fenomenal bilinç materyal/fiziksel ise, onun da bilgisini bilimin bu meşru yöntemleriyle elde edebilmeliyiz. Fakat daha fenomenal bilincin en temel özelliklerini kavramakta zorluk çekerken ve onun doğasını modern bilimin terimleriyle anlamanın uygun bir yolunu dahi bilemezken bunu nasıl gerçekleştirebiliriz? Bu da bize göstermektedir ki fenomenal bilinci anlamak için modern bilimin yöntemlerinin dışında ve ötesinde felsefi yöntemlere ihtiyaç duymaktayız.

Fenomenal bilincin doğasını anlamaya yönelik bazı temel epistemolojik yaklaşımları şöyle sıralayabiliriz: *Doğaüstücülük*: Fenomenal bilinç doğanın bir parçası değildir ve bu yüzden doğayı bilmek için kullanılan klasik bilimsel ve felsefi

metotlarla bilinemez. *İdealizm*: Bütün doğa aslında sadece bilinçten oluşmuştur ve fenomenal bilinci ancak bu metafizik görüşe uygun epistemolojik yöntemlerle bilebiliriz. *Natüralist İkicilik*: Doğa bilinç ve madde olmak üzere iki ayrı materyalden oluşmuştur ve bu ikisini ancak ayrı ayrı yöntemlerle bilebiliriz. *Standard Dışı Bilimsel Tekçilik*: Fenomenal bilinç doğanın bir parçasıdır fakat onun tabiatını standart rasyonel/ampirik yöntemlerle anlamak mümkün değildir. *Bilişsel Sınırlamacılık*: Fenomenal bilinç doğa'nın bir parçasıdır. Fakat onun tabiatını ne standart rasyonel/ampirik yöntemlerle ne de bunların dışında yöntemlerle bilmemiz mümkündür. Bunun sebebi insanın bilişsel doğasını sınırlayan asla aşamayacağımız epistemik engellerdir. *Materyalizm/Fizikalizm*: Fenomenal bilincin doğasını, modern bilimin ve çağdaş felsefenin kullanageldiği standart yöntemlerle bilebiliriz ve bunun önünde hiç bir engel yoktur.

Anglo-sakson felsefe geleneğinde fenomenal bilinçle ilgili en popüler yaklaşım, klasik ismiyle materyalizm, çağdaş ismiyle fizikalizmdir. Bu iki terimi, pek çok yazar birbiri yerine kullansa da ikisi arasındaki temel fark şudur. Klasik bir tanımla materyalizm evrende var olan her şeyin maddeden oluştuğunu iddia ederken fizikalizm modern fiziğin etkisiyle evrende var olan her şeyin, fiziğin fiziksel kabul ettiği nesnelerden oluştuğunu ileri sürmektedir. Bu anlamda, fiziksel olan (ki uzay, zaman, süreçler, enerji, yerçekimi gibi kuvvetler vs. yi içine alır) maddi olandan daha geniş bir alanı kapsamaktadır. Ayrıca fizikalist doktrinin bir diğer önemli varsayımı da bütün fiziksel ve biyolojik bilimlerin tamamının fizik bilimine indirgenebileceği kabulüdür.

Bu çalışmada ana hedefim materyalizm ya da fizikalizmin fenomenal bilinci açıklamada karşı karşıya kaldığı başlıca sorunları ortaya koymak, bu sorunların ana kaynağı olarak görülen epistemik/açıklayıcılık gediğini (benim adlandırdığım şekliyle fenomenal-fiziksel gediğini (FF-gap)) incelemek ve bu gediği fiziksel terimlerle açıklayabileceği öne sürülen "fenomenal kavram stratejisinin" (FKS)

yüzleşmek zorunda olduğu fenomenal öznelik problemini ortaya koymak olacaktır.

2. Fizikalizm Ve Üç Tür İlişki: Aynılık, Bağımlılık ve Gerçekleştirme

Fizikalizm en temel tanımıyla evrende var olan her şeyin fiziksel olduğunu ve fiziğin terimleriyle açıklanabileceğini iddia etmektedir. Bu tezi, fenomenal ile fiziksel olanın arasındaki bağın doğasına göre değişik şekillerde formüle etmek mümkün. Çağdaş anglo-sakson felsefede fenomenal ile fiziksel olanın arasındaki bağa dair birbirinin alternatifi üç tür ilişki biçimi öne sürülmektedir. Bu bölümde sırasıyla bu üç ilişki türünü tanımlayacak ve olumsal fizikalizmin savunulamaz olduğunu iddia edeceğim.

a) Aynılık İlişkisi

Aynılık ilişkisi felsefenin her alanında kullanılan bir ilişki türüdür. Zihin-beden ilişkisinden bağımsız olarak aynılık ilişkisini en genel anlamıyla aşağıdaki gibi tanımlayabiliriz:

> *Aynılık İlişkisi: Bir nesne, nitelik, olay ya da durum x, başka bir nesne, nitelik, olay ya da durum y ile aynıdır ancak ve ancak x ve y tek ve aynı nesne, nitelik, olay ya da durum ise.*

Bu tanıma göre aynılık fizikalizminin temel iddiası şudur: *Zihinsel durumlar fiziksel durumlarla bir ve aynıdır.*

b) Bağlılık İlişkisi

Yine zihin-beden ilişkisinden bağımsız olarak en temel tanımıyla bağlılık ilişkisini şöyle tanımlayabiliriz:

Bağlılık İlişkisi: X-tür nitelikler, Y-tür niteliklere bağlıdır ancak ve ancak herhangi iki nesne a ve b için, eğer a ve b aynı Y-tür niteliklere sahipse (*bütün Y-tür nitelikleri ortaksa*)*, o zaman a ve b aynı X-tür niteliklere sahiptir* (*X-tür nitelikleri ortaktır*)*.*

Bu tanıma göre bağlılık fizikalizminin temel iddiası şudur: *Var olan her durum fiziksel durumlara bağlıdır—Eğer herhangi iki mümkün dünya tümüyle aynı fiziksel niteliklere sahipse, bu iki dünyanın tüm özellikleri aynıdır.*

Eğer bu küresel iddiayı zihin-beden ilişkisine uygulayacak olursak karşımıza şöyle bir iddia çıkar: *Zihinsel nitelikler fiziksel nitelerlere bağlıdır—Eğer herhangi iki insan tümüyle aynı fiziksel niteliklere sahipse bu iki insan aynı zihinsel niteliklere sahiptir.*

c) Gerçekleştirme (Ortaya Çıkarma) İlişkisi

Aynılık ve bağlılık ilişkisi ile ilgili öne sürülen bir takım sorunlardan sonra Levine, Melnyk ve Shoemaker gibi felsefeciler zihin-beden ilişkisinde "gerçekleştirme ilişkisi" diye adlandırdıkları bir ilişki türünü savundular. Genel anlamıyla bu ilişki türünü Shoemaker'dan alıntı yaparak aşağıdaki gibi tanımlayabiliriz:

> *Gerçekleştirme İlişkisi*: Genel olarak X, Y'yi gerçekleştirir ancak ve ancak X'in varlığı Y'nin var olmasında yapısal olarak yeterli ise—ancak ve ancak Y'nin varlığı X'in varlığının üzerinde ve ötesinde hiç bir şey değil ise. (2007, s. 4)

Bu tanıma göre gerçekleştirme fizikalizmini şu şekilde formüle edebiliriz: *F niteliğinin bir örneği başka bir nitelik G'nin bir örneği vasıtasıyla gerçekleştirilir ancak ve ancak G niteliğinin örneği yapısal olarak F niteliğinin örneğinin gerçekleştirilmesinde yeterli ise* (*F niteliğinin örneği, G niteliğinin örneğinin üzerinde ve ötesinde değil ise*)*.*

Herhangi bir fizikalist tez, bu üç tür ilişki biçiminden hangisini merkeze alırsa alsın zorunlulukçu bir tez olmak zorundadır. Bunun için iki neden gösterilebilir. (i) Putnam-Kripke'ci yorumlamalar ve (ii) nitelik ikiciliği ile uyumlu olma riski. Aynılık ilişkisini temel alan olumsal fizikalizm, Kripke'ci yorumlamaların açıkça gösterdiği üzere en temel semantik ilkeleri çiğnemektedir. Bağlılık ve gerçekleştirme ilişkisini temel alan olumsal fizikalizm türleri de eğer dikkatle incelenirse nitelik ikiciliği ile uyumlu olmaktalar ki bu da gerçek bir fizikalist tezin ruhuna aykırı düşmektedir.

3. Anti-Fizikalist Argümanlar Ve Fenomenal ve Fiziksel Olanın Arasındaki Gedik

Bu bölümde bizi ilgilendiren argümanlar klasik Kartezyen anti-materyalist argümanlar değil, çağdaş anglo-sakson felsefesinde fizikalizme karşı öne sürülmüş argümanlardır. Bunların en önemlilerini şu şekilde sıralayabiliriz: bilgi argümanı, düşünebilirlik/kiplik argümanı, açıklayıcılık gediği argümanı ve nitelik ikiciliği argümanı. Bu argümanların en temel ortak özelliği, hepsinin zihinsel olan ile fiziksel olan arasındaki epistemik/açıklayıcılık gediğine dayanmalarıdır. Varsayalım *P* evrendeki tüm fiziksel doğruların hepsini kapsayan ("ve" mantıksal bağıyla birbirine bağlayan) bir önerme olsun. Epistemik/açıklayıcılık gediği özetle şunu söylemektedir: Zihne dair herhangi bir önermeyi *P* önermesinden türetmek mümkün değildir. Eğer sorunun sezgisel, bilişsel vs. zeminlerde de durumunu göz önüne alırsak, söz konusu gediği "fenomenal-fiziksel gediği" ya da kısaca FF-gediği şeklinde isimlendirmek uygun olacaktır.

Aşağıdaki üç sınırlayıcı ilke FF-gediğinin ilkece kapatılamaz bir gedik olduğunu açıkça göstermektedir:

(*NF*) *Nomolojik-Fiziksel Sınırlama: Nomolojik olarak, belirli bir fenomenal durumu üçüncü şahıs perspektifinden gözlemlemek imkânsızdır.*

(*MF*) *Mantıksal-Fenomenal Sınırlama: Mantıksal olarak, belirli bir fenomenal durumun birden fazla fenomenal özne tarafından deneyimlenmesi imkânsızdır.*

Bu iki epistemolojik sınırlama aşağıdaki ispatlanabilirlik sınırlamasıyla birlikte ele alınınca apriori fizikalizmin savunulamaz bir versiyon olduğu sonucuna varıyoruz:

(*İS*) *İşaret Eden Terimler İçin İspatlanabilirlik Sınırlaması: İşaret eden iki terimin işaret ettiği nesnenin tek ve aynı olduğunu doğrulamanın potansiyel bir yolu olmadıkça bu iki terimin tek bir nesneye işaret ettiği ispatlanabilir değildir.*

O halde sonuç oldukça açık. Apriori fizikalizm savunulamaz. Çünkü FF-gediği ilkece kapatılabilir bir gedik değildir. Buradan yola çıkarak diyebiliriz ki fenomenal-fiziksel ilişkisi konusunda fizikalizmin ümit vadeden tek versiyonu aposteriori zorunlulukçu fizikalizmdir.

4. Tip-B Fizikalizm, Fenomenal Kavram Stratejisi Ve Fenomenal Öznelik Problemi

İkinci bölümde fizikalizmin her türlü versiyonunun zorunlulukçu olması gerektiğini ve dolayısı ile olumsal fizikalizmin savunulamaz olduğunu ileri sürmüştüm. Üçüncü bölümde ise FF-gediğinin ilkece kapatılamaz bir gedik olduğunu savunmuş ve bundan da apriori fizikalizmin savunulamaz olduğu sonucunu çıkarmıştık. Bu da demek oluyor ki fenomenal bilince dair fizikalist yaklaşım aposteriori zorunlulukçu bir tez olmak zorunda. Aposteriori zorunlulukçu fizikalizmin en ümit vadeden versiyonu Daniel Stoljar'ın (2005) adlandırdığı şekliyle "fenomenal kavram stratejisi"dir. Bu stratejiyi benimseyen pek çok teori öne sürüldü. Tek tek bu teorileri inceleme imkânımız bulunmamakta. Fakat bu teorilerin genel yapısını ve en temel akıl yürütmesini şöyle özetleyebiliriz: Fenomenal kavramlar—örneğin "kırmızı algısı" kavramı ya da "acı" kavramı—fiziksel kavramlardan—örneğin

“nörolojik olay x” kavramı ya da “c-fiberi yanması” kavramı—radikal bir biçimde farklıdır. İnsanın bilişsel doğasından kaynaklanan bu radikal farklılık, FF-gediğinin en temel sebebidir. FF-gediği kapatılamaz bir gediktir. Çünkü söz konusu kavramsal farklılık fenomenal bilinç ile ilgili düşünme süreçlerimizi direkt etkileyen kavramsal mekanizmalarımızdan kaynaklanıyor. Bu farklılığı apriori bilememizin nedeni de budur. Bilişsel doğamızı yöneten mekanizmalar hakkında ancak ampirik yöntemlerle bilgi edinebiliyoruz.

İlk bakışta bu strateji, hem anti-fizikalist argümanlara temel oluşturan FF-gediğini fiziksel terimlerle açıklamaya aday olarak gözükmekte, hem de her türlü versiyonunun zorunlu bir tez olması gereken fizikalist tezin neden ancak aposteriori olarak bilinebilmesini açıklıyor gözükmekte. Fakat bu strateji özünde kuale merkezli bir stratejidir. Yani fenomenal bilincin fenomenalitesini oluşturan niteliklilik ve öznellik yönlerinden sadece ilkine yoğunlaşıyor. Bu yüzden FF-gediğinin belki daha önemli etkeni öznellik özelliğini göz ardı ediyor.

Fenomenalitenin öznellik yönü benim “fenomenal öznelik problemi” şeklinde adlandırdığım problemle çok yakından bağlantılıdır. Problemi şu şekilde açıklayabiliriz: (a) “Deneyimleme” fenomeni tözsel olarak var olan üç unsurdan oluşmaktadır: Deneyimleyen özne (fenomenal özne), deneyimlenen şey (özne tarafından deneyimlenen kuale) ve bu ikisinin arasındaki deneyimleme bağı (fenomenal özne-yüklem-nesne bağı). Bu analiz doğrultusunda FF-gediğinin kaynağını sadece deneyimlenen ögenin yapısında aramak fenomenaliteyi eksik kavramanın bir sonucu olacaktır. FF-gediğinin kaynağının bütüncül bir şekilde deneyimleme olgusunun bütün unsurlarında aramak gerekir. (b) Ayrıca yukarıdaki üç unsurdan herhangi biri olmaksızın deneyimleme olgusunun gerçekleşmeyeceği açıktır. Bu da demek oluyor ki bu üç unsur birbirine, mantıksal ya da nomolojik bir bağla bağlıdır. FF-gediğinin tek kaynağı bunlardan sadece birisi, yani deneyimlenen öge olamaz. (c) FKS savunucularının merkeze aldığı kuale

fenomenal özne olmaksızın var olabilen bir unsur değildir. Bunun en açık delili hiç öznesi olmayan fakat fenomenal bilince sahip insanların sahip olduğu her türlü kualeyi deneyimleyebilen bir zombiyi düşünememizdir. Yani kualeler bir fenomenal özne olmaksızın yalnız başına var olabilen şeyler değildir.

Yukarıdaki nedenler ve onlarla bağlantılı başka diğer gerekçelerden yola çıkarak fenomenal öznelik probleminin, genel anlamda aposteriori fizikalizminin ve özelde FKS teorilerinin önünde ciddi bir engel oluşturduğunu söylememiz mümkündür. Yine de aposteriori fizikalistler ya da FKS savunucuları, fenomenal özneliğin ancak dilde var olduğunu, gerçekte fenomenal öznelerin tözsel bir varlığa sahip olmadıklarını ve bu yüzden fenomenal bilincin fiziksel terimlerle izahında herhangi bir engel oluşturmayacaklarını öne sürebilirler. Bunun hatalı bir yaklaşım olduğunu bir sonraki bölümde fenomenal öznelerin neden tözsel bir varlığa sahip olmaları gerektiğini göstererek izah etmeye çalışacağım.

5. Fenomenal Öznelerin Tözsel Varoluşu

Bir önceki bölümde FKS'nin karşı karşıya olduğu fenomenal öznelik problemine değinmiş ve bu probleme karşı olası cevaplardan bir tanesinin fenomenal öznelerin tözsel olarak var olmadıkları, ancak dilsel ya da formel olarak var oldukları olabileceğini söylemiştik. Bu bölümde bu itiraza karşı fenomenal öznelerin tözsel olarak var olduklarını ispatlamaya çalışacağım.

Bir canlının bilinçli olması ne demek anlayabiliyoruz. En kaba tanımla bir canlının bilinçliliği onun uyanık ve zeki olması şeklinde yorumlanabilir. Fakat bir zihinsel durumun bilinçli bir durum olmasından ne kastedildiği o kadar açık değildir. Akla gelen ilk teori öncesi tanım şöyle olabilir: Zihinsel durum x bilinçlidir ancak ve ancak bir özne S, x'in farkında ise. Aslında bu tanım, bilinçliliğin, zihinsel durumların özüne dair bir niteliği olmaması, aksine harici bir niteliği olması gerçeği ile birleşince bizi fenomenal öznelerin tözsel olarak var olduğu sonucuna

götürür. Eğer bilinçlilik bir zihinsel durumun harici niteliği (sonradan kazandığı ve bazen sahip olmayabildiği bir niteliği) ise ve bir zihinsel durumun bilinçli olmasının en basit teori öncesi tanımı bir öznenin o zihinsel durumun farkında olması ise, bu demektir ki bilinçlilik zihinsel durumlar ile fenomenal özneler arasında harici bir ilişkidir. Öte yandan biliyoruz ki zihinsel durumlar da tözsel bir varlığa sahiptir, ya da en azından tip-B fizikalizmi ve özelde FKS savunucuları, en başından benimsedikleri kabuller gereği bunu yadsımamaktadırlar. O halde soru şu: Tözsel bir ilişki (farkındalık ilişkisi) biri tözsel, diğeri tözsel olmayan bir varlığa sahip (dilsel, formel, kavramsal vs. bir varlığa sahip olan) iki şeyi birbirine bağlayabilir mi? Cevap açıkça "hayır" olmalıdır. Buradan da açıkça fenomenal öznelerin tözsel bir varlığa sahip olmaları gerektiği sonucunu çıkarabiliriz.

Fenomenal öznelerin tözsel varlığına dair yukarıdaki akıl yürütme ile tatmin olmayan bir fizikalist için daha başka gerekçeler de ileri sürülebilir. Örneğin fenomenal birlik ve fenomenal devamlılık gibi fenomenlerin gerçekliği ile fenomenal kendine özgülük, fenomenal faillik ve fenomenal ben-lik duygusu gibi olgular ancak tözsel bir varlığa sahip fenomenal öznelere başvurarak açıklanabilir.

6. Fenomenal Öznelik Problemini Çözmede Fizikalistlerin Karşı Karşıya Kaldığı Bir İkilem

İlk beş bölümü özetleyecek olursak, birinci bölümde fenomenal bilincin modern bilim ve felsefe için neden ve nasıl bir problem olarak ortaya çıktığını, ikinci bölümde materyalizm ya da fizikalizmin bu probleme nasıl yaklaştığını, üçüncü bölümde fizikalist teze karşı öne sürülen anti-fizikalist argümanları ve bu argümanların dayandığı fenomenal-fiziksel gediğini (FF-gediğini), dördüncü bölümde bu FF-gediğinin varlığını kabul eden fakat bu gediğin fizikalizm için bir problem oluşturmayacağını ileri süren tip-B fizikalizmi ile onun en ümit vadeden versiyonu fenomenal kavram stratejisini (FKS) ve bu ikisinin önündeki fenomenal öznelik problemini ve beşinci bölümde de fenomenal öznelerin tözsel olarak var

olduklarını ortaya koymaya çalıştık. Bu bölümde ise tip-B fizikalizmi ile FKS'nin fenomenal öznelik problemine ne çerçevede cevap verebileceğini ve cevap verme girişiminde yüzleşmek zorunda kalacakları bir takım sorunlara işaret edeceğim.

Dördüncü bölümde deneyimleme fenomeninin tözsel olarak var olan üç unsurdan oluştuğunu söylemiştik. Bunlar deneyimleyen özne (fenomenal özne), deneyimlenen şey (özne tarafından deneyimlenen kuale) ve bu ikisinin arasındaki deneyimleme bağı (fenomenal özne-yüklem-nesne bağı) idi. Eğer bu üç yapısal unsurun üçü de tözsel bir varlığa sahipse, zihnin haritasında fenomenal öznenin nasıl bir yer işgal ettiğini belirlemek için alternatiflerimiz sınırsız değil. Olası mantıksal dört model ancak şunlar olabilir:

(b) *Yapısal Olmayan Model*: Fenomenal özneler, deneyimlenen şeylerle (kualelerle) yapısal bir ilişki içinde değildir. Onlardan yapısal olarak ayrıktırlar.
(a) *Yapısal Model*: Fenomenal özneler deneyimlenen şeylerle (kualelerle) yapısal bir ilişki içindedir. Bu da üç türlü olabilir:

(i) *Kısmi Yapısal Model*: Fenomenal özneler kısmi olarak deneyimlenen şeylerin (kualelerin) sadece bir kısmından oluşmaktadır.

(ii) *Kapsayıcı Yapısal Model*: Fenomenal özneler tümüyle deneyimlenen şeylerin (kualelerin) bir kısmından oluşmaktadır.

(iii) *Örtüşen Yapısal Model*: Fenomenal özneler tümüyle deneyimlenen şeylerin (kualelerin) toplamından oluşmaktadır.

Bu modellerin tamamında hem tip-B fizikalizmi hem de FKS aşılması güç problemlerle karşı karşıya kalmaktadır. Sırasıyla ele alalım.

Eğer fenomenal özneler kualelerle yapısal bir ilişki içinde değilse tip-B fizikalizmi hem fenomenal öznelerin hem de fenomenal özne-yüklem-nesne bağının fiziksel karşılığını açıklamak zorunda kalacaktır: Aksi halde yeni bir FF-gediği ile karşı

karşıya kalır. Öte yandan eğer fenomenal özneler kualelerle yapısal bir ilişki içinde ise, o vakit de FF-gediği sanılanın aksine daha büyük ve mahiyeti daha farklı bir gedik olacaktır. Çünkü artık fenomenal bilincin fenomenal olma özelliği, sadece deneyimlenen şeyi değil, aynı zaman da deneyimleyen özneyi ve deneyimleme bağını da içereceğinden FF-gediğini fiziksel terimlerle açıklamak daha zor bir iş haline gelecektir.

FKS'nin durumu yukarıdakinden daha iç açıcı değildir. Yapısal olmayan modelde bir fenomenal öznenin kendi özneliğine ait fenomenal kavramı bilinçle ilgili düşünme süreçlerinde kullanmasında ciddi sorunlar çıkmaktadır. Yapısal üç modelde ise FKS savunucuları sadece fenomenal öznenin fenomenal öznelikle ilgili kavramı kullanmasında değil, diğer fenomenal kavramları kullanmasında da aşılması güç sorunlarla karşı karşıya kalmaktadırlar.

Tüm bu değerlendirmeler özetle bize şunu söylüyor: Genel olarak fizikalistler, özel olarak tip-B fizikalistler ya da FKS savunucuları, fenomenal öznelikle ilgili fenomenal terimlerle formüle edilmiş ciddi teoriler sunmaksızın FF-gediğini fiziksel terimlerle açıklama hedefini gerçekleştiremezler. FF-gediğini fiziksel terimlerle açıklayamazlar ise de fenomenal bilincin doğasını fizikalist tezle uyumlu bir şekilde açıklayamazlar. Bu da fizikalist tezin yanlışlanması anlamına gelecektir. Bu da demektir ki fenomenal öznelikle ilgili fenomenal terimlerle formüle edilmiş kapsamlı teoriler kaçınılmazdır.