PSEUDO-MYSTERIANISM OF MCGINN: A CONSTRUCTIVE AND NATURALISTIC SOLUTION OF THE HARD PROBLEM OF CONSCIOUSNESS IS POSSIBLE

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

BY

SENA IŞIKGİL

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

SEPTEMBER 2022

Approval of the thesis:

PSEUDO-MYSTERIANISM OF MCGINN: A CONSTRUCTIVE AND NATURALISTIC SOLUTION OF THE HARD PROBLEM OF CONSCIOUSNESS IS POSSIBLE

submitted by **SENA IŞIKGİL** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Philosophy, the Graduate School of Social Sciences of Middle East Technical University** by,

Prof. Dr. Sadettin KİRAZCI
Dean
Graduate School of Social Sciences _____

Prof. Dr. Halil TURAN
Head of Department
Department of Philosophy _____

Assoc. Prof. Dr. Aziz Fevzi ZAMBAK
Supervisor
Department of Philosophy _____

Examining Committee Members:

Prof. Dr. Mehmet Hilmi DEMİR (Head of the Examining Committee)
Ankara Social Sciences University
Department of Philosophy _____

Assoc. Prof. Dr. Aziz Fevzi ZAMBAK (Supervisor)
Middle East Technical University
Department of Philosophy _____

Prof. Dr. David GRÜNBERG
Middle East Technical University
Department of Philosophy _____

Prof. Dr. Murat ARICI
Selçuk University
Department of Philosophy _____

Assoc. Prof. Dr. Samet BAĞÇE
Middle East Technical University
Department of Philosophy _____

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: Sena IŞIKGİL

Signature:

ABSTRACT

PSEUDO-MYSTERIANIM OF MCGINN: A CONSTRUCTIVE AND NATURALISTIC SOLUTION OF THE HARD PROBLEM OF CONSCIOUSNESS IS POSSIBLE

IŞIKGİL, Sena
Ph.D., The Department of Philosophy
Supervisor: Assoc. Prof. Dr. Aziz Fevzi ZAMBAK

September 2022, 128 pages

Consciousness is the most important aspect of our minds, and there are several perplexing questions about consciousness. The hard problem of consciousness is the problem of finding an answer for the question of how and why non-physical mental states arise from physical brain states. Although there are several traditional attempts to solve the hard problem, none of them have provided successful explanation up to now. As a result of these failed attempts, in the last decade of the 20th century, Colin McGinn, a strong mysterian, suggested a dissolution to the consciousness-brain problem. McGinn's thesis, Non-constructive Naturalism, claims that although there is a naturalistic solution of the hard problem of consciousness, we are cognitively closed to this solution because of our limited cognitive faculties. However, McGinn's arguments for cognitive closure are not plausible enough to leave consciousness to dark side. The recent studies on consciousness in the science and philosophy shows that there can be promising approach on the table, e.g. The Integrated Information Theory (IIT), that strongly suggest a possible solution within the framework of constructive naturalism. Even if the IIT is not completely successful as of now, it shows us the possibility of solving the hard problem of

consciousness by taking the phenomenal nature of consciousness seriously. In the light of these, the main aim of this dissertation is to reveal that mysterianism of McGinn is pseudo-mysterianism and that a constructive and naturalistic solution of the hard problem of consciousness is possible.

Keywords: the hard problem of consciousness, non-constructive naturalism, pseudo-mysterianism, the Integrated Information Theory

ÖZ

MCGİNN'İN SÖZDE-GİZEMCİLİĞİ: BİLİNCİN ZOR PROBLEMİNİN YAPICI VE DOĞALCI BİR ÇÖZÜMÜ MÜMKÜNDÜR

IŞIKGİL, Sena
Doktora, Felsefe Bölümü
Tez Yöneticisi: Doç. Dr. Aziz Fevzi ZAMBAK

Eylül 2022, 128 sayfa

Bilinç, zihnimizin en önemli yönüdür ve bilinçle ilgili oldukça kafa karıştıran sorular vardır. Bilincin zor problemi, fiziksel beyin durumlarından fiziksel olmayan zihin durumlarının nasıl ve neden ortaya çıktığı sorusuna cevap bulma problemidir. Bu zor problemi çözmek için birçok geleneksel yaklaşım olmasına rağmen şimdiye kadar hiçbiri başarılı bir açıklama getirememiştir. Bu başarısız girişimlerin bir sonucu olarak, 20. yüzyılın sonlarında güçlü bir gizemci olan Colin McGinn bilinç-beyin problemini tümüyle ortadan kaldırmayı önerdi. McGinn'in Yapıcı Olmayan-Doğalcılık tezi, bilincin zor probleminin doğalcı bir çözümü olmasına rağmen sınırlı bilişsel yetilerimiz sebebiyle bizim bu çözüme bilişsel olarak kapalı olduğumuzu iddia eder. Fakat McGinn'in bilişsel kapalılığa yönelik argümanları bilinci karanlık tarafta bırakmak için yeterince makul değildir. Bilimde ve felsefede bilinç üzerine yapılan son çalışmalar bize Bütünleşik Bilgi Teorisi (BBT) gibi yapıcı doğalcılık çerçevesinde olası bir çözümü kuvvetle öneren umut vaat edici bir yaklaşımın var olabileceğini gösteriyor. BBT şu an için tam anlamıyla başarılı olmasa bile bilincin fenomenal doğasını ciddiye alarak bilincin zor problemini çözme olasılığını bize gösteriyor. Tüm bunların ışığında, bu tezin temel amacı, McGinn'in gizemciliğinin

sözde-gizemcilik olduğunu ve bilincin zor probleminin yapıcı ve doğalcı bir çözümünün mümkün olduğunu ortaya koymaktır.

Anahtar Kelimeler: bilincin zor problemi, yapıcı olmayan doğalcılık, sözde-gizemcilik, Bütünleşik Bilgi Teorisi

To the Memory of My Grandmother

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Mehmet Hilmi Demir for his effective advice, friendly guidance, and encouragement throughout my dissertation project. His intellectual and psychological support made everything easier to me.

Special thanks to my supervisor Assoc. Prof. Dr. Aziz Fevzi Zambak for his invaluable contributions with his helpful criticisms and effective suggestions from another perspective for this dissertation.

I would also like to express sincere thanks to Prof Dr. Murat Arıcı who has always kept his eyes on this dissertation from the beginning to the end. He made my work easier with his precious feedback and support.

Thanks to Prof. Dr. David Grünberg for his kind interest in my dissertation project. I am grateful to him for his valuable feedback and important comments for this dissertation.

I would also like to thank to Assoc. Prof. Dr. Samet Bağçe for his contributions and endless support. I am grateful to him; whenever I needed, he was always there both academically and friendly.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

TN: Transcendental Naturalism

FP: Folk Psychology

IIT: Integrated Information Theory

CHAPTER 1

INTRODUCTION

Consciousness has become the most perplexing and mysterious problem among philosophers, cognitive scientists and neuroscientists with an explosion of research on the brain for a few decades. There are advanced brain-scanning techniques like EEG, MRI, fMRI or PET scans to measure the physical activities of human brain, but human beings are still far away to understand the consciousness. All of us know what it is like to smell a camomile, see a red wall, or feel pain; however, none of the brain-scanning devices are effective to explain the subjective nature of what occurs within us when we smell a camomile.

Although consciousness is ordinary to us, for we directly and intimately know it, there is nothing more complicated to explain since consciousness cannot be easily fitted into the physical world. It is surely beyond doubt that conscious states are related with the physical brain states. We do not have a doubt that our redness experience is related with our physical brain states at the time of seeing red wall, but though we acquire the total knowledge of physical brain process, up to now we have not found an answer for the question of how physical brain states give rise to conscious states, which have subjective character. What is the relation between spatial brain properties and non-spatial conscious phenomena? How do non-material mental phenomena result from gray matter? These are manifestations of the hard problem of consciousness known as consciousness-brain (or mind-body) problem.

Even tough non-physical and subjective character of conscious state have been ignored by scientists for a long time, philosophers have been discussing it as a subjective phenomenon under the name of mind-body problem or the consciousness-brain problem thousands of years. There are some solution-oriented approaches for the problem of consciousness. While some of these approaches claim that the mental

and the physical are two distinct substances or two distinct properties of a physical substance, some of them identify the mental states with the neuro-physical properties of the brain. There are of course some approaches that dispute the hard problem of consciousness and totally reject the existence of conscious states as well. However, unfortunately none of them has been able to provide a satisfactory explanation for understanding of consciousness, and these failure stories drove some philosophers to despair about solving the problem of consciousness.

In the last decade of the 20th century, Colin McGinn, a pioneer of the new movement called Neo-Mysterianism, suggested a dissolution of the consciousness-brain problem, and he claims that although there is a natural and simple solution of the mind-body problem, we, as human beings, cannot attain this solution because of our limited cognitive capacity. That is, in his thesis called Transcendental Naturalism he defends the idea that because constructive solution of the consciousness problem is not possible, understanding of consciousness will remain inaccessible for humanity forever. I think that McGinn's claim is not digestible.

In my opinion, the solution of the consciousness-brain problem depends on revising either our understanding of consciousness or understanding of the physical world. So far philosophers and neuroscientists have developed various theories dealing with the concept of consciousness from different perspectives, but unfortunately none of them has succeeded in explaining consciousness on a physical basis. This is the sign for us to rethink the problem by revising our current knowledge about the physical world and its physical objects instead of ignoring the characteristics of consciousness. The main goal of this dissertation is to reveal that the mystery of consciousness does not stem from the consciousness' absurd character; on the contrary, insoluble character of consciousness is the result of humanity's partial knowledge of the physical world, and there is a promising theory such as IIT that naturally places consciousness in a physical world by starting from the sui generis character of consciousness. Given this overall goal, this dissertation will be organized as follows:

The second chapter will formulate the real problem of consciousness. For this purpose, firstly, I will introduce the readers to the concepts of consciousness. In the literature there are different kinds of concepts of consciousness, but in order to study the problem of consciousness and the brain, it is necessary to specify the concept of consciousness in the sense I use. For this reason, I will analyse the most common concepts of consciousness under two main categories: “creature consciousness” and “state consciousness”. I will first explain the concepts that highlight the conditions making a creature a conscious creature, then I will try to define what a state of consciousness is and to specify the types of states of consciousness by using Ned Block's distinction (P-consciousness and A-consciousness) and Thomas Nagel's "what it is like" locution. Secondly, I will highlight the features of conscious states such as subjective and qualitative character as a ground for the hard problem of consciousness. And then, I will formulate what the real consciousness problem is by comparing David Chalmers' formulation of the easy and hard problem of consciousness and Joseph Levine's explanatory gap. In a final step I will touch briefly on the question of why consciousness resists to the scientific explanation, and analyse the required conditions of the theory that opens the door to the scientific explanation of consciousness.

Chapter three will focus on the traditional solution-oriented approaches to the consciousness problem and McGinn’s DIME dance which is the result of the persistent failure of the traditional solutions. It is possible to divide the proposed solutions to the mind-body problem into two broad categories: the first category represents the approaches that consider the subjective and qualitative character of consciousness important, and the second category represents the approaches that ignore the subjective and qualitative character of consciousness. The persistent failures of these extreme categories trying to solve the problem have turned into an endless loop that oscillates between these two extremes. McGinn defines this vicious circle, which we can accept as the starting point of McGinn’s mysterianism, as the DIME dance. This chapter first analyses the dualistic views and the monist/materialist approaches with their problems and the objections, and then mentions the DIME circle which represents their persistent failure.

The dualistic approaches, which are based on the idea that the mind and the body are two different substances, or two different properties of one physical substance can be classified under the first category. The advanced version of substance dualism is based on Descartes' *Meditations.* Descartes' substance dualism supports the idea that although the mind and the brain are related to some degree, they are two distinct substances – the mind, thinking thing, and the body, extended thing –. It is not possible to be blind to the influence of substance dualism on the emergence of modern approaches to the mind-body problem, but substance dualism was generally rejected in the twentieth century. There is a contemporary and popular version of dualism known as property dualism. As opposed to substance dualism, which insists on the existence of two distinct substances, property dualism asserts that there is only one physical substance but two distinct kinds of properties. Property dualists claim that the mental properties cannot be reduced to the brain properties, but the brain can be bearer of both the physical and the mental properties. As kinds of property dualism, I also mention emergentism, which formulates consciousness as an emergent property of the brain, and in addition to this, I will also analyse panpsychism, the view that describes the mental as a fundamental property of physical things. There are also dualistic approaches that investigate the interaction between the mental and the physical. Among these views that examine the interaction between the mental and the physical, the most basic versions that will be summarised in this dissertation are interactionism, parallelism, occasionalism and epiphenomenalism.

As a reaction to the problematic aspects of dualistic views about the mental-physical relation, monist materialist approaches which can be classified under the second category of traditional solutions, have become popular among some philosophers. Materialism rejects the two distinct substances of dualism, and the materialists base their views on monism by claiming that if there is an interaction between the mind and the body, it is unreasonable to define the mind as a distinct thinking substance, it must be solely material (Feser, 2005, p: 46). This chapter will briefly discuss the materialistic approaches and their problems in terms of four main views:

Behaviourism, psycho-physical identity theories, functionalism, and eliminative materialism.

After I revealed that the solution-oriented approaches examined in the previous chapter could not provide a satisfactory solution to this problem, in chapter four I will analyse the mysterians, Thomas Nagel and Colin McGinn, who are pessimistic about achieving the scientific explanation of the problem of consciousness. While Nagel defends the weak version of mysterianism by leaving open to the possibility of solving the mystery of consciousness in the future, McGinn defends the strong version by claiming that the mystery is permanent due to the cognitive closedness of human being. And, although I address Nagel's thesis and compare it with McGinn's, in this chapter I will mainly discuss the strong version of mysterianism.

Especially in his article, "Can we solve the mind-body problem?", McGinn claims that there is a naturalistic solution of the mind-body problem, but it is not constructive (1989, p: 350). That is, according to him, there is a natural and simple explanation of consciousness-brain interaction; however, because of our limited cognitive capacity we can never specify the link between the mind and the brain. Scientific explanation of problem cannot be possible, so solution of the consciousness problem will remain an epistemic mystery for human being. He thinks that though consciousness will remain a mystery, facing with the insolubility of consciousness dissolve the mind-body problem for us. McGinn's theory is called as Transcendental Naturalism (TN), and he formulates this theory in three main steps given below:

(I) "There exists some property of the brain that accounts naturalistically for consciousness"
(II) "We are cognitively closed with respect to this property"
(III) "There is no philosophical mind-body problem" (1989, p: 352)

After I explained TN by respectively analysing these three main steps, I will show its problematic aspects in this chapter as well. Firstly, I will draw attention to McGinn's

first argument – I will call it “the argument for naturalism” and analyse his definition for the nature of property providing the mind-brain relation. McGinn asserts that property, *P*, must be non-spatial because of the fact that ordinary spatial properties of brain are not sufficient to mediate between the mind and the brain. I will reveal that although McGinn says that the *P* providing the link must have mediating characteristics, while defining the features of *P* he conflicts with this idea. In McGinn’s thesis the problem about consciousness-brain relation is replaced with the problems about “*P*-brain”, “*P*-consciousness” relations. Secondly, I will draw attention to the second argument which I will call “the argument for closure”, and I will show that McGinn’s cognitive closure idea on which epistemic mystery of consciousness is based is also problematic. After briefly explaining his last argument which I will call “the argument for dissolution”, at the end of this chapter I will have manifested that (I)-(III) are not satisfactory triad to leave consciousness to dark side, and McGinn’s mysterianism is pseudo-mysterianism.

The aim of McGinn is to impose that it is time to wave the white flag for philosophers who noodle over the problem of consciousness in hopes of finding its solution. In chapter five, however, I will defend the claim that a naturalistic and constructive solution to the consciousness-brain problem is possible. Until now, it has been believed that the non-physical or the non-spatial character of consciousness creates a hard problem because it cannot be fitted into the entirely physical world. However, there is an intuitional belief which is based on Einstein’s Relativity theory that time and space cannot be separated from each other. In other words, the entity which exists in time must be exist in space and vice versa. For this reason, it seems to be unplausible to assume that consciousness cannot be fitted into the spatial world. Nagel thinks that problem emerges from the dichotomy between the concepts of objective and subjective. The objective concepts that we have today and that form the basis of the scientific explanations ignore the subjective character of consciousness and it does not seem possible to solve the problem with the concepts we currently have. With the objective phenomenology thesis he proposed, he states that we need to new concepts that also include the subjective and qualitative aspects

of consciousness for the solution of the problem and expresses that he is not hopeless in this regard.

I think that for the new concepts we need to revolution in our knowledge and understanding of the physical world. Up to now we have generally tried to solve the hard problem of consciousness by revising the properties of consciousness in various way instead of thinking that the physical world we live in may have different fundamental entities than we currently know. In my view, the problem of consciousness emerges from the problem in the formulation of the problem. Formulating the problem by assuming that our knowledge of physical world is complete unsurprisingly leads to the problem with respect to the seemingly non-physical nature of the consciousness because we cannot understand how non-physical consciousness emerges from the physical brain. We cannot understand this because we cannot fit consciousness into the physical world. If our scientific knowledge of physical world can be adequately advanced and our understanding of the objects of the physical world can be revolutionized, then the seemingly non-physical consciousness can be unproblematically accepted as a denizen of the new world. I think that Giulio Tononi's Integrated Information Theory (IIT), which explains consciousness with the system's capacity of integrating information and identifies the subjective character of consciousness with the special way of integrating information, is considerably promotive example for the possibility of explanation of subjective consciousness in the scientific way. Unlike all theories that have tried to explain consciousness on a scientific basis, the IIT begins by strictly adhering to the characteristics of consciousness. It seems more promising approach to explain the seemingly non-physical consciousness, when it is compared with the theories based on the fixed framework of the classical physics because they exclude the subjective character of consciousness, but all theories of consciousness that ignore the real nature of consciousness are condemned to failure.

As a conclusion part, chapter six will summarize the whole dissertation. Consciousness problem, which is the problem of explaining the relation between the spatial brain and the non-spatial consciousness, is the hard nut of the philosophy of

mind. Although there is still no definitive solution, I think that the theory, which preserves the subjective side of consciousness and investigates consciousness with different physical types from material entities as we know, are quite promising in terms of scientific explanation of consciousness. With this claim at the end of this chapter, I will have concluded that strong mysterianism about consciousness is unwarranted; revolution with respect to the understanding of the physical world is like a light at the end of the tunnel for consciousness puzzlement.

CHAPTER II

AN INTRODUCTION TO THE PROBLEM OF CONSCIOUSNESS

Now I sit at the computer with the intention of writing the second chapter of this thesis. While I am writing these words, odour molecules enter my nose stimulating my olfactory receptor, and I get the tantalising smell of the coffee on my desk. My dog is lying on the carpet in my room gnawing on his big bone; I think he is so happy as he fills his belly with his favourite food. I am feeling a bit unwell today with a dull ache and a twinge in my neck. Just for a few minutes I stop to write, and my attention turns to the window to watch the children playing in the park. I remember times gone by, the hammock under the mulberry tree in our garden, and imagine my own happy childhood for a few moments. I take a sip of my coffee and while its taste sends me into euphoria, I get back to writing. While you read these lines, you may be tasting your coffee, or you may be feeling a little hungry. Maybe you have had enough of what is written here, or you are wondering why I am telling you an excerpt from my everyday life.

The reason of this tiny story is to remind you that out of dreamless coma our daily lives are alive with the states named as *conscious* mental states – intending, smelling and tasting, thinking, feeling an ache, feeling happy or hungry, imagining, wondering, etc. – and we, human beings, as owners of various kinds of conscious states, are at the centre of our own individual conscious experiences. You cannot escape your feelings or desires, your conscious experiences, and it does not seem possible for anyone else to be closer to your conscious experiences than you are. Your direct access to your own conscious state is the most important evident for the existence of it. However, although our own conscious states are directly achievable and so ordinary for us, consciousness is quite a perplexing subject about which we have no theoretical knowledge. As human beings we still do not reach a conclusion

about what consciousness is, what the nature of a conscious state is, how it arises and how it is related to the brain.

The purpose of this chapter is to formulate the problem of consciousness, which I will deal with the rest of the dissertation. To do this, however, I will first point out the concepts of consciousness in the literature, for although it is a term frequently used in everyday language, it does not yet have a commonly accepted meaning. My aim in analysing the concepts of consciousness is to highlight the terminological diversity of consciousness in the literature, because it is important to distinguish the concept of consciousness from its other definitions in the sense that I will use throughout the dissertation. Then I will look at the features of consciousness because these features play an important role in the emergence of the problem, and then I will set out what the real problem of consciousness is by taking help from Chalmers' hard problem thesis and Levine's explanatory gap argument.

2.1. What does the term consciousness mean?

To study on consciousness, it is necessary to understand what the term consciousness means. The term conscious refers to various kinds of phenomena in the literature, as there are different proposed answers by philosophers or scientists to questions about consciousness. The most common distinction in the field of consciousness that has been generally agreed upon is the distinction between “creature consciousness” and “state consciousness” (Rosenthal, 1993, 1997; Lycan, 1996). While the notion of creature consciousness is used to attribute consciousness property to a subject –a human or an animal– the notion of state consciousness is used to attribute consciousness to mental states.

The real controversy is under what conditions we can call a creature or a state conscious. Is there a single criterion for defining something as conscious, or are there distinct kinds of criteria? There is certainly no doubt that not all mental phenomena are conscious phenomena, but earlier philosophers such as Descartes and his successors do not have a need to use the term conscious to emphasize the difference

between conscious and unconscious mental states because they consider that we are conscious of all our mental states. Descartes makes a distinction between the mind and the body, while the mind represents the thinking substance, the brain represents the extended substance. In this view "thinking" has a broad sense that includes all kinds of mental states and being conscious is essential for the existence mental states (Gennaro, 2016, p: 10). There can be no mental state of which we are not conscious. However, with the increase in research on consciousness the view is widely accepted that there are both conscious mental states we are aware and unconscious states of which we are not aware.

There are many concepts that have been developed to explain what makes a state a conscious mental state, but before we analyse them, let us focus on the concepts of consciousness that underline the conditions for a creature's being conscious. In the most general and basic sense, consciousness is characterized by "wakefulness". In this sense, being conscious depends on a human's or an animal's being alert and receptive to sensory stimuli. It is generally assumed that an organism could be conscious only if it is awake, and it refers to the opposite condition in which an organism is unconscious as dreamless sleep or deep coma. For instance, Tononi supposes that "[consciousness] is what vanishes every night when we fall into dreamless sleep and reappears when we wake up or when we dream" (2008, p: 216). Papineau and Selina also begin their book, *Introducing Consciousness*, with the assertion that "consciousness is what we lose when we fall into a dreamless sleep or undergo a total anaesthetic" (2000, p: 5). Similarly, Searle defines consciousness as a state of awareness that begins "when we wake from a dreamless sleep and continue through the day until we fall asleep again, die, go into a coma or otherwise become 'unconscious'" (1998, p: 1936). However, as mentioned before there is no definition of consciousness agreed upon, and of course, there are philosophers who argue against the synchronisation between consciousness and wakefulness. To illustrate, Thompson defends the argument that consciousness continues in dreamless sleep

(2015, p: 6). According to him, we can remember the quality of our sleep because we have consciousness in our deep and dreamless sleep as well.[1]

In addition to wakefulness, some are willing to define the term consciousness in such a way that it is identical with the terms such as "awareness" or "attention". It can be acceptable to claim that being conscious of something is required to be able to aware of those things, but awareness and consciousness are semantically different from each other. As Gennaro asserts that there are cases for awareness without consciousness like subliminal occurrences, so defining consciousness as synonymous with the awareness is nonacceptable today. (2016, p: 7)

Similarly, it seems as if consciousness and attention are inextricably linked with each other. Some scholars argue that "attention is necessary for conscious perception" (Mack and Rock, 1998, p: 250).[2] The evidence for this strong argument comes from phenomena like "inattentional blindness"[3], which show that in the lack of attention there is no conscious perception.[4] The invisible gorilla experiment is the most famous example of inattentional blindness. It is a video in which two groups of students, one group in a black shirt and the other in a white shirt, pass the basketball to each other, and the viewers are in demand to count the number of passes among the students wearing white shirt. After a while, a person in a gorilla suit appears in the scene between the students, but many viewers do not notice the person in the gorilla suit (Simons and Chabris, 1999, pp: 1062-69). The conclusion of this

[1] According to Thompson, when a person wakes up, he/she has the idea and the memory of "I slept well", and this is the sign for the awareness in deep sleep. For more details you can see the Part 8 of the book, *Sleeping: Are We Conscious in Deep-Sleep?* (2015, pp: 231-272)

[2] There are many scholars defending the same argument about the relationship between consciousness and attention. For instance, you can see Posner (1994), Simons and Chabris (1999), Mole (2008), De Brigard and Prinz (2010), Prinz (2011, 2012).

[3] This expression belongs to Mack and Rock and is used as their 1998 book's title.

[4] Inattentional blindness is one of those two well-known phenomena, the other of which is the change blindness. These two phenomena have similar features in terms of their contributions to the strong argument about consciousness-attention relationship, so I will continue with the inattentional blindness and its illustration, but for more information about change blindness, you can see Rensink at all. (1997), Simons at all. (2000) and Simons (2000).

experiment seems to be consistent with the assertion that "attention is necessary for conscious perception", but it does not provide us with satisfactory evidence to assume that this assertion is exactly true, because even if there are viewers who do not notice, 42% percent of viewers notice the person in gorilla suit (Mole, 2008, p: 95). It seems reasonable to assert that consciousness cannot be restricted with attention because there are further case studies that show the circumstances in which consciousness exists without attention, and attention exists without consciousness.[5]

The concepts of consciousness described so far are generally associated with the properties of a creature. However, we make a distinction not only between conscious-unconscious creature, but also between conscious and unconscious states because a creature's being conscious does not mean that all its mental states will be conscious. We know something about the conditions that make a creature conscious being, but the main issue is what condition makes a state a conscious mental state.

In the field of state consciousness there are different concepts referring the property of conscious mental states. The most crucial and frequently used concept is characterized by Thomas Nagel: "What it is likeness". In his 1974 article "What is it like to be a bat?" he claims that "fundamentally an organism has conscious mental states if and only if there is something that it is like to be that organism – something it is like for the organism" (1974, p: 436). For explaining this claim Nagel uses the example of the bat and argues that even if we, human beings, know the processes of bats' perception and whole facts about them in detail, we cannot know what it is like to be a bat; we cannot understand what it is like to spend a day shimmying from top to bottom, or we cannot know what it is like to use echolocation to make prey. While we, human beings, can imagine these states from a third-person point of view, a bat can experience them from its own subjective point of view. When I am in a certain mental state, for example when I smell a cup of aromatic coffee, there is something special to me in that state; what it is like for me to smell that cup of aromatic coffee is my own subjective conscious experience. By using the phrase "what it is like"

[5] For further information about the case studies that reveal the dissociation between consciousness and attention you can see Boxtel at all. (2010).

Nagel emphasizes the subjective character of conscious states. In his view, the criterion of subjectivity or, in other words, first-person perspective is essential to define a state as a conscious mental state. That is, a pain in my neck can only be considered a conscious mental state if there is something that it is like to feel this pain from my perspective.

There are also different terms to characterize consciousness in Nagel's "what it is like" sense, e.g., "qualia". The term qualia (singular: quale) is used by some philosophers to refer to the subjective and qualitative nature – "what it is like" character – of conscious states (Lewis, 1929; Jackson, 1982). When you hit your head on the wall, feel a severe headache, or feel dizzy, each of these states expresses a particular raw feeling that is introspectively accessible to you. And some philosophers use the term qualia to express this aspect of consciousness. Another concept associated with the Nagelian sense of consciousness is phenomenal consciousness. In his works, Ned Block makes a critical distinction between two different types of consciousness, access (A) consciousness and phenomenal (P) consciousness, and he defines "P- consciousness" in the Nagelian sense by asserting that "a state is phenomenally conscious if there is something it is like for one to be in that state". While phenomenal consciousness refers to our conscious experience "when we see, hear, smell, taste and have pains" (Block, 2002, p: 206), access consciousness signifies the state's "availability for use in reasoning and rationality guiding speech and action" (Block, 1995, p: 227). In his view, A-consciousness and P-consciousness do not always coincide with each other; these two types are conceptually independent. Block's distinction emerges from the thought that the properties of P-consciousness are different from the functional or cognitive properties of consciousness (*ibid.,* p: 207). P-consciousness has experiential properties, that is, "what it is like" to have that experience, and this is the reason why phenomenal aspect of consciousness is more puzzling than A-consciousness. While access consciousness can be explained by cognitive science, explanation of phenomenal consciousness resists the methos of cognitive science because of its subjective and qualitative character.

As seen in the discussion above, although consciousness is a very common word in everyday language, it is used in different meanings. Even though there are lots of attempts to describe consciousness, it is still a quite ambiguous term for which there is no commonly held objective definition. However, for the remainder of the dissertation, I will use the term consciousness in Nagelian sense, which emphasizes the subjective and qualitative nature of consciousness, because the problem of consciousness essentially focuses on the explanation of phenomenal consciousness, i.e., the "what it is like" character of a conscious state.

2.2. What makes consciousness problematic: The features of consciousness

If we want to talk about what generates the hard problem of consciousness, it is required to address the features of it. Presenting the essential features of consciousness both contributes to understanding of what we mean when we talk over consciousness and helps us to clarify the reason why the hard problem of consciousness is so hard.

There are two essential features of states of consciousness that make the question of consciousness and the brain a difficult problem: The first is the subjective character and the other is the qualitative character of consciousness (Levine, 2001; Kriegel 2009). When we focus on a subject's phenomenal experience of a white rose – what it is like for the subject to see a white rose – while for-subject-ness of this state refers to the subjective character of consciousness, the component of what it is like – the whiteness of the white rose – refers to the qualitative character.

There is further problematic feature of consciousness related with its subjective and qualitative nature like immediacy. A subject has direct or immediate knowledge of her/his own conscious phenomena (Kim, 2011, p: 159). This means that your knowledge about your phenomenal consciousness is not based on evidence, inference or any other things. If we assume that you have a stomach pain, because of immediacy nature of consciousness you have direct and immediate knowledge of this pain without evidence and inference. This feature distinguishes your conscious

phenomena (stomach pain) from physical phenomena (physical condition of your stomach). While the knowledge of your pain is only private to you, you can have evidence like endoscopy results for knowing your stomach's physical conditions.

Because subject, who has conscious phenomena, can have immediate knowledge of this state from the first-person perspective, there are first person-third person distinction in accessing knowledge of conscious phenomena (ibid., p: 162). To illustrate, when you have a stomach pain as we said above, only you have information of this conscious state, your doctor or anyone else cannot have it, and this situation creates first person-third person distinction. While you are the first person because you have direct and immediate knowledge of your mental phenomenon, your doctor or anyone else is a third person because he/she can only access to the knowledge of your stomach's physical conditions by using some medical evidence or inference.

In the light of the explanation above, it can be said that although consciousness arises from the brain, a physical system, it has completely different properties from the brain. And this is the core of the puzzle of consciousness. It is obvious that scientific research, based on empirical methods like observation, experimentation and verification, solves the problems of the world in which we live. However, the problem of how something as complicated as consciousness with respect to its features arises from the grey matter in the head is radically different from all other scientific problems. While the brain is open to scientific investigation by scientific methods due to its observable character, consciousness seems to resist scientific research because of its features which are different from the entities in conformity with the scientific methods. For this reason, the relation between consciousness and the brain is considered as the hard problem of consciousness.

2.3. Stating the Hard Problem of Consciousness

Consciousness is the most distractive problem in the philosophy of mind. As Chalmers says, if we want to make progress on the consciousness problem, firstly we

should make a distinction between the problems of consciousness. Some problems of consciousness are easier than the other to solve; therefore, at first the truly hard part of consciousness problems should be separated from the relatively easy ones (2010, p:3). Chalmers asserts that some phenomena such as “the ability to discriminate”, “categorize or react the environmental stimuli”, “the reportability of mental states”, “the focus of attention”, “the deliberate control of behavior” or “the difference between wakefulness and sleep” are easy problems of consciousness (*ibid*., p: 4). The reason why he identifies such phenomena as easy is that even though we do not have complete knowledge about them yet, we get an idea on how we can give an explanation to these phenomena by using the methods of neuroscience and cognitive science. For example, to explain the difference between the states of wakefulness and sleep we need only observe the neuro-physical processes leading the contradictory behaviours of organism in these states. Or in a similar way, for explaining the reportability of mental states the only need is specifying a mechanism making achievable the information about the inner states for verbal report (*ibid*., p: 4-5).

However, easy problems of consciousness are not the subject of this dissertation. As Blackmore says that the real problem is associated with the questions about what the world is composed of (2013, p: 8). If the world is purely physical, how can non-physical consciousness be fitted into it? There is no doubt that the brain is the home of consciousness because even if we cannot observe consciousness in the brain, we can introspectively aware the correlation between states of consciousness and the brain activities. However, what is the relation between purely physical brain states and mental states?

Let us take an example, and suppose that *P* is your subjective visual experience of pine-tree in your garden, while you can answer some main questions about what the pine-tree in the world is made of, or its physical properties, spatial dimension and solidity, you cannot answer the same questions for *P*. There is no doubt that *P* is correlated with some set of neural activities in your brain, but while these neural activities have purely physical properties, and they have a place in the brain, we do not have enough knowledge about the nature of *P*. The relation between objective

brain and subjective *P* seems like a mystery for us, so the question of how the physical state of the brain leads to mental experience is the most brain-teaser problem of consciousness, and this problem is called by Chalmers the "hard problem" of consciousness.

Chalmers describes the hard part of consciousness as the problem of experience and explains it in the following words:

> The really hard problem of consciousness is the problem of experience. When we think and perceive, there is a whir of information processing, but there is also a subjective aspect. As Nagel (1974) has put it, there is something it is like to be a conscious organism. This subjective aspect is experience. When we see, for example, we experience visual sensations: the felt quality of redness, the experience of dark and light, the quality of depth in a visual field. Other experiences go along with perception in different modalities: the sound of a clarinet, the smell of mothballs. Then there are bodily sensations from pains to orgasms; mental images that are conjured up internally; the felt quality of emotion; and the experience of a stream of conscious thought. What unites all of these states is that there is something it is like to be in them. All of them are states of experience. (2010, p: 5)

As seen in the quotation above he emphasizes the essentially subjective character of consciousness because the real problem is how subjective character of consciousness emerges from the purely physical process of the brain. Unlike the easy problems, the methods of cognitive science and neuroscience are not sufficient to be able to explain this problem of consciousness. In the pine-tree example above, by doing empirical research we can achieve the information of neuro-physical process of the state – visual inputs coming from the pine-tree to organism's eyes, stimulation of the optic nerve back of the head of organism, neural firing in the brain etc.–, but objective research gives us nothing about the organism's subjective pine-tree experience – what its greenness is like or what it is like to have the smell of pinecone for the organism in that state – we need something new to be able to achieve the satisfactory explanation of conscious experience.

2.3.1. Levine's Explanatory Gap

You may be materialist, and you may claim that just one type of entity in the world, and consciousness is identical with this type.[6] That is, in the pine-tree example above you may assert that P is identical with the neuro-physical process in the brain. As a result of this claim the problem may go away for you. However, even if psycho-physical identity thesis is ontologically true, there is an epistemological gap in the explanation of psycho-physical identity, and Levine uses the phrase "explanatory gap" to describe the unexplained relation between the physical and the mental (1983, 1993, 2001).

For doing this, Levine uses Kripke's objection to materialism, and he transforms Kripke's metaphysical objection to the epistemological one. According to identity proponents, identifying a kind of mental state like pain with a kind of neural state like C-fiber stimulation is not problematic as identifying heat with molecular motion or water with H_2O. For identity theorists all these identities are contingent a posteriori statements as claimed in the orthodox view. The concepts of a priori or a posteriori are epistemological concepts related to the way a proposition is known. While a priori proposition can be known by reason, a posteriori proposition can be known by appealing to experience. And according to classical assumption, if a proposition is necessary, it is plausible to expect that this proposition must be known only by a priori, and if it is contingent, it is knowable by a posteriori. However, Kripke argues that the orthodox assumption that all a priori propositions are necessary truth, and all contingent propositions are also a posteriori is problematic because even if some necessary propositions are known a priori, there are also propositions which are neither known a priori nor formed in analytical sentences. These are a posteriori necessarily true propositions.

In order to describe the structure of such kind of statements Kripke uses the concept of "rigid designator". According to Kripke, "rigid designators" refer to entities which

[6] The materialist theories will be discussed at length in the next chapter.

have some essential properties. And because entities cannot be thought without their essential properties "rigid designators" designates concepts which refer to the same things in all possible worlds. The proper names and natural kind terms, for example, are rigid designators. In Kripke's view, identity statements including two rigid designators must be necessarily true. In other words, he argues that if A=B is true, then A and B have to designate the same thing in all possible worlds. A=B must be necessarily true, and this means that there cannot be any possible world in which A exists without B or vice versa (1980, pp: 146-150). For instance, "Hesperus (the evening star)" and "Phosphorus (the morning star)" refer to the same thing (Venus) in all possible worlds, and so they are necessarily identical. "Hesperus is Phosphorus" is an a posteriori necessary truth because its truth can only be proved by empirical data. Kripke enlarges on this argument by using two more different identity analogies:

(1) Heat = molecular motion
(2) Pain = C-fibers stimulation (ibid., p: 154)

Kripke argues that if (1) and (2) are true, then both must be necessarily true. That is, both sides of the identity in each statement must refer to the same thing in all possible worlds. If the necessity cannot be shown, it is not possible to talk about identity between "heat" and "molecular motion" or "pain" and "C-fibers stimulation". The identity statement (1) is necessarily true because there is not a possible world in which heat exists without molecular motion or the presence of molecular motion without heat. The identity thesis (2) is false because if it is true, it must be necessarily true; however, it is not necessarily true because pain without C-fibers stimulation, or vice versa is conceivable. Therefore, it cannot be assumed that pain is C-fibers stimulation.

The identity theorists think that (2) is contingent a posteriori truth like (1). However, Kripke argues that (1) is only seemingly contingent. That is, considering the identity between heat and molecular motion as contingent is an illusion that arises from our confusion between the sensation of heat and the heat. It is conceivable that molecular

motion may exist without the sensation of heat such as the feel of warm or cold. This can be plausible. However, heat is molecular motion in all possible world, even if the heat does not lead to the feeling of heat at every time. Therefore, even if it is seemingly contingent, in reality it is necessary truth for Kripke. Such an explanation is implausible for (2) because the sensation of pain, the pain feeling, is the pain itself (ibid., pp: 151-52). For this reason, if (2) is true, it must explicitly be necessary truth, but it is false because it is conceivable that there are possible worlds in which pain is not C-fibers stimulation.

As seen above Kripke's argument has aim to reveal that mental states cannot be identical with the brain states. However, Levine claims that Kripke's argument is not sufficient to say something about ontological truth of the psycho-physical identity thesis, because conceivability of something is not adequate to establish the metaphysical facts about that thing. According to Levine, for a physical reduction to be successful what is reduced should be fully explained; in other words, the physical description of mental states and properties is not sufficient alone, the reduction theory should give an explanation to these mental states and properties as well (1993, pp: 128-9). He justifies this claim by pointing out the kind of difference between (1) and (2). In statement (2) the identity between heat and molecular motion can be expressed totally with whole crucial points; on the other hand, in statement (1) there are something unexplained between pain and the C-fibers stimulation, in the words of Levine there is an "explanatory gap" in between (1983, p: 357). (2) is the theoretical identity thesis, and our theoretical knowledge on physics and chemistry makes the relation between heat and the molecular motion fully explainable. However, (1) is psycho-physical identity statement, even if C-fibers stimulation can be related with the causal role of the pain, the concept of pain has more than its causal role; it has qualitative character, and C-fibers stimulation is not sufficient to explain the qualitative character of the pain. "What it is left unexplained by the discovery of C-fiber firing is why pain should feel the way it does!" (*ibid.,* p: 357).

Although Levine does not directly argue against the identity thesis, by explanatory gap argument he argues for that the truth or the falsity of it cannot be

epistemologically accessible, and this outcome seems to be a vexing problem for materialists (*ibid.*, p: 360). There are many theories that try to close this gap, and the aim of the following chapter of this dissertation is to discuss the solution-oriented approaches giving an explanation to the hard problem of consciousness, from dualistic approaches to monist/materialistic approaches. However, before this, I will discuss the reason of why the hard problem of consciousness is the hard nut for the scientific explanation.

2.4. The Reason of Why Consciousness Resists to the Scientific Explanation: Possibility of Explaining Consciousness in The Scientific Manner

It is believed that phenomenal consciousness or the subjective/qualitative character of consciousness resists to the scientific explanation in metaphysical and epistemological ways (Arıcı, 2018, p. 6). If we look at it from a metaphysical point of view, all entities in the world are shaped according to the fundamental particles of physics and the physical laws they are attached to. There is an intuitional belief by many people that phenomenal consciousness cannot be explained in terms of fundamentals of physical world and the rules upon which they are bound.

It will not be different when we consider this situation from an epistemic point of view. There is a distinction known as “first-person perspective” versus “third-person perspective” in analytic philosophy and science. This distinction is expressed as the "subjective-objective" distinction, as Nagel puts it, or as the "non-scientific-scientific" distinction, since the third-person perspective is considered by many philosophers as a scientific-objective perspective. When questioned as to what makes research on consciousness so interesting and so difficult at the same time, it is thought that the problem arises from the distinction between first-person and third-person perspectives. There is a belief that a scientific explanation of consciousness cannot be made because it is thought that it is not possible to know what it means to be a bat or to know what another person feels and how he perceives the world from the third person view. In other words, while science tries to find solutions to

problems from a third-person perspective, consciousness is understandable from a first-person perspective.

I also agree with the idea that it does not seem possible to explain subjective consciousness metaphysically in terms of physical/material fundamental entities as we know them today. As will be discussed in the next chapter, psycho-physical theories that try to explain subjective consciousness by reducing it to the entities of classical physics have failed by ignoring the subjective and qualitative aspect of consciousness. And these failures have led some pessimistic philosophers, like Colin McGinn, to despair of a scientific explanation of consciousness and to regard it as a mystery that can never be explained. However, the purpose of this thesis is to reveal that McGinn's mysticism is an unfounded and excessive claim, and to suggest that there may be naturalist approaches that promise to explain consciousness on a physical and scientific basis by taking the subjective side of consciousness as a starting point. The answer to the question of what the basic features of the theory that can give us the scientific explanation of phenomenal consciousness should be, may be as follows:

i. Firstly, there must be a theory that cares about the phenomenological features of consciousness, such as qualitative and subjective nature, and should starts from there.

ii. It should be able to express the subjective character of consciousness in scientific and mathematical language and translate them into these fields.

iii. The entity and process in which the phenomenal consciousness is explained must be placed on a physical basis not in the sense that we currently know but in a way that expands our perception of space and physical world.

iv. What is modeled, translated into this scientific mathematical language, must also have scientifically testable predictions. These tests may be difficult, not very likely, but they should at least be logically testable.

v. The implications of the theory should not be counterintuitive. For example, if the implication of a consciousness theory denies the subjective experience of animals, as Descartes' theory did, it denies that they have no feeling like pain, and says that they have no consciousness, then this is a counterintuitive theory. The implications of theory should conform to all systems that provide the conditions for being conscious.

A theory that adapts to the conditions mentioned above may open new door for us that there may be a scientific explanation of subjective consciousness and even that the physical world may have foundations in a different sense than we currently know. We will discuss all these in the following chapters, but before this, in the next chapter we will discuss the traditional approaches to the hard problem of consciousness and their problems.

CHAPTER III

AVAILABLE THEORIES THAT PURPORT TO SOLVE THE HARD PROBLEM OF CONSCIOUSNESS

The hard problem of consciousness or, more traditionally, the mind-body problem continues to maintain its vitality and importance despite all the philosophical theories, developments in neuroscience, artificial intelligence, and technology. The problem of consciousness has been dealt with by both philosophers and scientists from ancient times to the present. The main goal is to understand how mental states take place in the natural, physical world. In other words, it must be understood how mental states such as pain and pleasure arise in a world composed of electrons, atoms, electromagnetic fields, and gravitational forces without contradicting the laws of physics. This chapter of the dissertation discusses the traditional solution-oriented approaches that have been developed for the hard problem of consciousness. Traditional solutions can be examined under two separate categories: The first of these categories represents approaches that preserve the subjective side of consciousness but fail to understand the consciousness-brain interaction since they cannot explain it on a physical basis. The other category represents approaches that ignore or do not give due importance to the subjective or qualitative aspect of consciousness while trying to explain consciousness on a physical basis. These two extreme ways of thinking are incapable of solving the hard problem of consciousness, and this inadequacy puts the problem of consciousness into an endless vicious circle that oscillates between two extremes like a pendulum. This vicious circle, which Colin McGinn describes as the “DIME circle”, forms the basis of McGinn’s mysterianism, which claims that the problem of consciousness is forever unsolvable for human being. For this reason, in this chapter, we will first discuss dualistic approaches that emphasise the phenomenal character of consciousness but fail to explain its connection with the physical, and then materialistic approaches that ignore the phenomenal character of consciousness and

try to domesticate it in the physical world with their existing objections and problems. Finally, we will make a preliminary preparation on the way to McGinn's mysterianism by expressing the circle caused by these two extreme ways of thinking with McGinn's DIME acronym.

3.1. The approaches that emphasise the subjective and the qualitative character of consciousness

There are approaches that prioritize the subjective and qualitative character of consciousness, distinguish it from the physical one and keep it superior. We will examine these approaches as dualist approaches that accept consciousness and the brain as two separate substances or properties. Dualism is a theory that generally emphasizes that the mental and the physical are two distinct realms. This theory divides into two main categories with respect to the ontological properties of the mental and the physical: the first is substance dualism, the second property dualism. While substance dualism claims that there are two distinct substances in the world, property dualism claims that there are two distinct properties but one substance. However, if there are two different substances or properties in the world, what kind of interaction exists between them? This question reveals some certain dualistic positions that try to understand the interaction between the brain and consciousness. These positions that examine the interaction between the mental and the physical include interactionism, epiphenomenalism, parallelism, and occasionalism.

3.1.1. Substance Dualism

Although dualism originated in a general sense in Plato's dialogues,[7] modern versions of dualism in philosophy of mind are based on arguments for or against Descartes' substance dualism, known as Cartesian dualism. Substance dualism holds that while the mind and the brain are related to some degree, they are two radically different substances. In general, substance is defined as something that has different

[7] In the Platonic dialogues, especially in *The Phaedo*, we find some expressions that define the soul as a reality that can exist separately from the body without owing its existence to the body.

properties but can exist independently of the properties it possesses. Descartes also states that "the notion of a substance is just this - that it can exist by itself, that is, without the help of any other substance. (1984, p. 159). This means that the mind as a substance has the ability to exist without the body. The properties that the mind has, e.g., thinking, believing, doubting, judging, etc., can exist independently of the existence of the body. for this reason, a distinguishing claim of substance dualism is that there is a mind that can exist independently of the body and vice versa.

In his important work, *Meditations on First Philosophy*, Rene Descartes formulates six meditations. Although these are not studies made specifically to defend substance dualism, we can see that Descartes' dualistic arguments are pronounced in these meditations, especially in the second and sixth meditations. Descartes begins his work with a methodological doubt about all things, including the entire external world and his own existence. Methodological doubt aims to reveal the undoubted and to save us from truly endless scepticism. He thinks it is logically possible that all of our knowledge based on the senses and sense experience could be completely false, due to sensory illusions and the lack of a criterion for distinguishing dream states from real states. While Descartes uses "dream argument" to point out that sense beliefs and knowledge based on sense experience are dubious, he uses his famous "evil genius hypothesis" to argue that mathematical knowledge is also dubious. According to him, even though mathematical knowledge is not based on our senses and dream states do not affect its accuracy, an evil genius could deceive us about its accuracy. However, after Descartes uncovers that knowledge of the external world based on sense experience and even the existence of one's own body as a sense object and the accuracy of mathematical information can be doubted, he finally arrives at a certain thing that is indubitable: his own existence as a thing that has a cognitive process is the only certain thing that he cannot doubt. And he explains this in the following words:

> [...] I have convinced myself that there is absolutely nothing in the world, no sky, no earth, no minds, no bodies. Does it now follow that I too do not exist? No: If I convinced myself of something then I certainly existed. But there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that

> case I too undoubtedly exist, if he is deceiving me; and let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. [...] I must finally conclude that this proposition, I am, I exist, is necessarily true whenever it is put forward by me or conceived in my mind. (1984, pp. 16-7)

As can be seen from the above passage, according to Descartes, the knowledge of the existence of the body, which is the result of his sense experiences, can be an illusion. Although Descartes doubts the existence of his body, he believes that even if a demon deceives him about his existence, deception and doubt are mental states, and the existence of mental states presupposes the existence of a mind. In other words, a person who doubts cannot doubt that he doubts, because in the state of doubting there is a doubting "I"; it is a way of thinking, a mental activity. This conclusion leads to his famous phrase "I think, therefore I am," which shows that he is convinced of the certainty of his own existence. Descartes refers to the nature of his own certain existence as “a thinking thing” because he concludes that he exists by saying that he thinks, and this is known as Descartes’ famous “cogito argument”.

Especially in the sixth meditation Descartes explains the exact differences between mind and body. The mind has no physical properties that the brain has, such as shape, size, solidity, volume, and so on. That is, the "I" as a thinking thing cannot be a material substance. While the immaterial mind represents the thinking substance, "res cogitans", the material brain represents the extended substance, "res extensa"; they have separate natures. He asserts that the material body is divisible, but the immaterial mind is indivisible. While the body, as a divisible physical entity, is destructible, the mind, because of its indivisible nature, is not; it is immortal (Jacquette, 2009, p. 15). If it were a material substance, when the body disappeared, it would not be possible for the "I" to continue to exist. However, even if the material substance disappears, the "I" can continue to exist as a thinking substance. This means that in Cartesian view the bearer of mental states is not the brain but the immaterial mental substance.

As a result of Descartes' argumentation, we can characterise the thesis of Cartesian dualism by three basic arguments: The first argument asserts that (1) the mind is a

distinct substance that can exist independently of the body, and the second emphasises that (2) the mind is non-physical unlike the body. However, characterising the position of the mind and the body according to (1) and (2) leads to some difficulties regarding the relationship between them. That is, the question arises as to what kind of relation exists between these two fundamentally different substances. Descartes does not deny their causal interaction, although he defines the mind and the body as two different kinds of substances. His interactionist argument explains that (3) the mind and the body causally interact in both directions. Mental states causally affect physical states and vice versa.

In his work *Passions of the Soul*, Descartes draws attention to a specific area of the brain concerned with the interaction of body and mind. He claims that the mind and the body interact with each other through the pineal gland, which he calls the "seat of the soul" (2002, p. 22). Signals between mind and body are transmitted through the pineal gland. However, there are serious objections to the idea that two essentially separate substances can causally interact. The first objection to Descartes' interaction thesis is raised by Princess Elisabeth of Bohemia with the following passage:

> […] how the mind of a human being can determine the bodily spirits [i.e., the fluids in the nerves, muscles, etc.] in producing voluntary actions, being only a thinking substance. For it appears that all determination of movement is produced by the pushing of the thing being moved, by the manner in which it is pushed by that which moves it. or else by the qualification and figure of the surface of the latter. Contact is required for the first two conditions, and extension for the third. [But] you entirely exclude the latter from the notion you have of the body. and the former seems incompatible with an immaterial thing. (as cited in Garber, 2000, p. 172)

As seen in the above passage, Elisabeth asks Descartes how there can be a causal interaction between material body and immaterial mind without any contact, and this is a plausible question. As Kim asserts that we can also integrate Elisabeth's question into the modern physics:

> For anything to cause a physical object to move, or cause any change in one, there must be a flow of energy, or transfer of momentum, from the cause to the physical object. But how could there be an energy flow from an immaterial mind to a material thing? What kind of energy could it be? How could anything "flow" from something

> outside space to something in space? If an object is going to impart momentum to another, it must have mass and velocity. But how could an unextended mind outside physical space have either mass or velocity? (Kim, 2010, p. 48)

Descartes' ideas about the control of the nonphysical mind over the physical body are generally rejected today for the reasons given by Kim in his remark. Moreover, in his work "Consciousness and Its Place in Nature", Chalmers mentions Descartes' argument as the most common version of interactionism under the category of "Type-D dualism", and he also claims that as the "most influential objection" to interactionism, "science tells us that the microphysical realm is causally closed" (2003, p. 126). That is, there is no place for a non-physical, thinking substance to have a controlling effect on the physical body in the causally closed network of the physical word; Descartes' interactionism is incompatible with the laws of physics (Chalmers, 2002, p. 2). Even though Descartes explained the mutual causal relation between mind and body through the pineal gland in the brain, he could not clearly explain the question of how the nonphysical mind and physical brain causally interact with each other, and he could not explicitly define the role of the pineal gland in this causal process. Descartes refers to the pineal gland as the seat of the mentality, but neuroscience now shows that mental states extend to the entire brain and that we cannot refer to a single brain region for mental activity. All this shows that the interactionist view of substance dualism does not provide a satisfactory solution to the problem of interaction between immaterial mind and material body.

In response to objections to interactionism, some dualists adopt epiphenomenalism, which rejects the causal role of the mind over the body. In medicine, the term epiphenomenon refers to the symptom that occurs during disease but do not causally affect the disease, and it is used in a way that preserves its meaning in the philosophy of mind in relation to the mind-body relation. That is, epiphenomenalism accepts the causal effect of physical activity in the brain on the existence of mental states, but rejects the causal efficacy of the mind on the brain. The nineteenth century biologist Thomas Huxley, one of the best-known proponents of this view in history, takes as his starting point Descartes' assertion that nonhuman animals are merely unconscious machines whose behaviour is the result of the causal role of their brains. Huxley

agrees with Descartes that animal behaviour is entirely caused by brain states, but disagrees with his assumption that they are merely unconscious machines. In his work, “On the Hypothesis That Animals Are Automata, and Its History”, he asserts that the conscious experience of animals depends only on the neurophysiological changes in their brains; however, their consciousness has no causal effect on the activation of their physical states, just as the steam whistle produced by the operation of the locomotive has no effect on the operation of the machine (2002, p. 29).

Huxley initially seemed to be concerned with animal consciousness, but at the end of his work he concludes that the same is true for humans as seen in the following passage:

> It will be said, that I mean that the conclusions deduced from the study of the brutes are applicable to man, [...] therefore, that all states of consciousness in us, as in them, are immediately caused by molecular changes of the brain-substance. It seems to me that in men, as in brutes, there is no proof that any state of consciousness is the cause of change in the motion of the matter of the organism. (2002, p. 30)

Since epiphenomenalism rejects the thesis that the mind causally contributes to the activation of physical states, it seems to have an advantage over the interactionist view in terms of adaptation to a causally closed network of the physical world. However, epiphenomenalism is also considered a problematic view because it is difficult to accept that mental states have a useless status that plays no impressive role in our physical states.

In addition to epiphenomenalism, there are also some dualistic views that reject the causal interaction between mind and body and justify the existence of this interaction with the intervention of God. Parallelism, generally associated with Leibniz and his theory of "pre-established harmony", for example, asserts that the correlation between the mind and the body is established by God; God initially established the universe in harmony between the physical and the mental. Similarly, according to Malebranche's view of occasionalism, whenever there appears to be a causal interaction between mental and physical states that lead to each other, what is

actually happening is the causal power of God over the apparent causal interaction that we observe (Kim, 2011, p. 96). That is, the relation between mind and body that appears to humans as a genuine causal relation is actually caused by divine power. Malebrane's occasionalism is also a kind of parallelism, but the difference between the two is as follows: according to Leibniz's theory, God initiates the movement of causality but does not intervene in it later. In the occasionalist view, God initiates the movement of the causal relation between the mind and the body and constantly controls it as the actual cause of every effect. Apart from this difference, both views argue that Descartes' interactionism is only an illusion; there can be no such causal relation between mind and body. Since these views prefer to explain the interaction between mind and body by the power of God rather than present an acceptable argument, they have only historical value for contemporary philosophy.

Not everyone agrees that body and mind are two different substances. In the twentieth century, many philosophers identify what we call mind as physical properties of the brain, behavioral activities, or functional states. According to the arguments we will examine in the next section under the heading of "monism," everything that exists is composed of a single material substance. However, before we move on to the versions of materialistic monism, it should be mentioned that there are also different types of monism. One of these types is idealism. Idealism holds that only spiritual events or spirits exist. In contrast to the dualistic view, which holds that both the mind and the body exist independently, idealism holds that there are no physical things or states that are not mental. Idealists do not deny the existence of objects in the physical world, but they do deny the existence of physical objects outside of a mind that perceives or thinks about them. The most important and well-known proponent of idealism is George Berkeley. According to Berkeley, the only source of our knowledge about physical things in the external world is our experience. Our experience does not tell us whether physical objects exist independently of our perception of them, so we cannot say on the basis of our experience that physical objects are material substance. By explaining the existence of physical things in terms of the existence of mind, idealism eliminates the problem of mind-body interaction, but it would not have explained it. Also, the view that

physical objects do not exist without a perceiving mind is not an accepted approach today. Another type that rejects the existence of body and mind as two distinct substances is the double aspect theory of Spinoza. Spinoza also adapts to the tradition of metaphysics based on substance. In his view, body and mind are two separate and interrelated aspects of one substance, which is neither physical nor mental in itself. However, Spinoza maintains that substance, as a being that needs nothing to exist, is a concept that can only be valid for God. Accordingly, the substance of his monism is only God, not body or mind. The body and the mind are two main aspects of God, who is a substance. Spinoza's view is thus opposed to both substance dualism and idealism.

3.1.2. Property Dualism

The most problematic aspect of substance dualism is the existence of mind without body. However, it is possible to be a dualist without considering the mind and the body as two different substances. This type of dualism, which is becoming increasingly popular today, is known as property dualism. Unlike substance dualism, which insists on the existence of two different substances, property dualism claims that there is only one physical substance but two different kinds of properties. Property dualists claim that the brain is the only physical substance that has both physical and mental properties. The main idea of property dualism is that mental properties are distinct from, and cannot be reduced to, the physical properties of the brain.

It can be assumed that property dualism is a kind of dualism between substance dualism and materialism. On the one hand, in contrast to substance dualism, it argues that only physical substance exists. On the other hand, in contrast to materialism, it claims that mental states cannot be explained by reducing them to physical states. In this way, property dualism attempts to avoid the problems that dualism and materialism run into. However, although it argues for the existence of a single physical substance, the interaction problem is also valid for property dualism, since it claims the existence of two different types of properties.

3.1.2.1. Emergentism and Panpsychism

Emergentism is an important variant of property dualism advocated by important philosophers such as J. S. Mill, S. Alexander, L. Morgan, and C. D. Broad. The most important factor distinguishing emergentism from the other types of property dualism is the answer of emergentism to the question of why some physical objects have mental properties, but others do not. According to this view, when lower-level microphysical systems or entities come together and reach a certain complexity, a new kind of higher-level property, namely consciousness, emerges from that physical system, and consciousness cannot be explained by reducing it to lower-level physical properties (Kim, 2011, p. 97). Emergentism assumes that the emergence of mental properties from the physical system that reaches a certain level of complexity is the fundamental fact of nature. The fundamental facts of nature are facts that cannot be derived from other laws of nature. That is, according to the statement of S. Alexander, we accept the existence of the fundamental fact about the emergence of mental properties with "natural piety" (1966, pp. 46-7).

The main problem of emergentism, which is a kind of property dualism, is the same as the main problem of property dualism in general. That problem, as we mentioned in the previous section, is to provide an explanation for mind-body interaction that is consistent with the principle of causal closure of the physical world. The alternatives that emergentism assumes with respect to mind-body interaction are generally the same as the alternatives that substance dualism assumes. Emergentism can adopt either interactionism or epiphenomenalism; the important difference between the two is that the former recognises "downward causation" i.e., the causal effect of higher-level emergent properties on lower-level microphysical properties, while the latter does not. Another problem facing emergentism is how mental features emerge from the adequate complexity of low-level brain features that do not have any mental characteristics. In other words, it has not been convincingly explained how mental properties such as thinking, pain, etc. emerge from the integration of the physical properties of the brain.

For some, the only satisfactory approach that overcomes these difficulties is panpsychism, which asserts that everything has some degree of mentality. While emergentism advocates the emergence of mentality from physical particles that reach a sufficient level of complexity, panpsychism holds that all physical things, including atoms, quarks, and other smaller particles that exist at a fundamental level, possess some form of mentality from the beginning. Panpsychists do not think it reasonable to believe that consciousness emerges from the combination of completely unconscious physical things. While the claim that a property like Y emerges from things that do not have that property is true for physical phenomena, it is not possible for consciousness. For example, H_2O molecules individually do not have the property of fluidity, but when a sufficient amount of H_2O molecules come together, fluidity emerges. It is not incomprehensible that fluidity emerges from individual non-fluid things; it is quite comprehensible to us that the property of fluidity arises from the fact that the molecules are connected together in a non-dense structure. However, if we think the same thing about consciousness, it is not plausible to assume that something completely non-concrete and non-spatial emerges from something concrete and spatial (Strawson, 2008, pp. 60-64).

Panpsychism also faces the objection referred to by William Seager (1995) as the "combination problem." This objection, first raised by William James (1890), states that the emergence of macro-experiences from the combination of micro-experiences is not comprehensible. James expresses his concern about this problem as follows:

> Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. (p. 160)

If we transfer the subject to recent period, William Lycan similarly states the following in his 2006 work:

> Suppose I am looking out of my kitchen window, and simultaneously seeing a rabbit in my back yard, hearing my wife's cat yowling that he wants to behead the rabbit, feeling the touch of my fingertips on a bottle of salad dressing, smelling the

> spaghetti sauce in the pot, suffering an ache in my right shoulder, and imagining in anticipation a very tall frosty beer. In what way could such a mental aggregate consist of or be determined by or otherwise 'arise from' a swarm of smaller mentations? (p. 69)

As can be seen from the above passages, the idea of the emergence of a macro-mind or consciousness from the combination of micro-minds seems as incomprehensible as the idea of the emergence of consciousness from the combination of unconscious things. And this means that the hard problem of consciousness returns to where it began.

3.2. The approaches that ignore the subjective and the qualitative character of consciousness

The influence of substance dualism on the development of modern views of the mind-body problem cannot be overlooked. However, this kind of dualism was generally rejected in the twentieth century. As a reaction to the problematic aspects of substance dualism in relation to the mental-physical interaction, the materialist approach has become popular among philosophers. Materialism rejects the two distinct substances of dualism, and materialists base their views on monism, claiming that if there is an interaction between the mind and the body, it is unreasonable to define the mind as an independent thinking substance, it must be exclusively material (Feser, 2005, p. 46). However, while monistic/materialistic approaches tried to make room for consciousness on a physical basis, they generally ignored the sui-generis phenomenal character of consciousness. Some of them reduced mental states to neurophysiological brain states and remained silent about the subjectivity and the qualitative character of consciousness, and some completely ignored the mental concepts that define consciousness as will be discussed in detail below.

3.2.1. Varieties of Materialism

I will briefly discuss the development of the materialist approach that try to domesticate phenomenal consciousness in the physical world in terms of four main

views: Behaviourism, Psycho-Physical Identity Theory, Functionalism and Eliminative Materialism.

3.2.1.1. Behaviorism

Behaviorism is the first materialistic view, which rises in the mid-twentieth century. Gilbert Ryle, with his 1949 work *The Concept of Mind,* is seen as one of the pioneers of philosophical behaviorism, also known as logical and analytical behaviorism. Especially in the section titled “Descartes’ Myth”, Ryle argues against Descartes’ substance dualism by identifying cartesian mind-body relation as “the ghost in the machine”. He draws an analogy between the phrase "mind in the machine" and the thought that someone who has no idea what a machine is has when he encounters a machine. When such a person observes that the machine can move by itself, he begins to think that there is a ghost or spirit inside the machine that is controlling it. Like this man, Descartes identified "the body including the brain and nervous system as a mechanical system" and identified the mind "as a ghost harnessed to the bodily machine" (Place, 1998, p. 370). Descartes' thoughts are compatible with the animism belief that defines death as the parting of the soul from the body and claims that the disembodied soul continues to exist (Place, 1998, p. 371), and for Ryle, these thoughts of Descartes are so dogmatic and primitive. Ryle's critique of Descartes shows that his primary goal is not to develop a theory of mind, but to reveal the problematic aspects of Cartesianism and to refute it. That is, he tries to reveal that there is “not Two-Worlds, but One-World; not a Ghost, but a Body; people are not Occult but Obvious” (Hampshire, 1950, p. 238). It is hard to summarize Ryle's views that he develops in the mentioned book and the rest of the thesis does not seem necessary, but it is accepted that the attitude of the book generally involves a kind of behaviorism. According to Ryle’s general view, the mind cannot be identified as an immaterial substance that independently exists from the body as Descartes believes; it corresponds only to the actual and possible behaviors of the organism.

As seen in Ryle’s attitude, behaviorism is an approach which is developed as a reaction against the unreasonable consequences of the Cartesian definition of mind.

It aims to eliminate the unreasonable consequences of the traditional dualist view by reducing mental concepts to accessible and observable behaviors, rather than explaining them in subjective inner states. In other words, according to behaviorist approach, a mental state like “pain” is explained by some certain publicly observable pain behaviors such as “wincing” (Chalmers, 2002, p. 3).

Although behaviorists generally agree that mental states can be explained in terms of behavior, there are different views of what is meant by behavior. For behaviorists, the most important criterion for an activity to be considered a behavior is that it is publicly observable/accessible. Kim says, “in much behaviorist literature there is an assumption that only psychological responses and bodily motions that are in a broad sense ‘overt’ and ‘external’ are to count as behavior” (2011, p. 67). That is, although activities or processes that take place in internal organs, such as activities in the brain, are observable, they are not considered behavior by some philosophers. However, for some, besides bodily motions, physiological activities in the body are counted among the behaviors that construct the meaning of mental terms. This view can be illustrated by Hempel's argument. In his 1949 work, Hempel proposes to explain mental terms by behavioral terms, and he uses the statement "Paul has a toothache" as an example. To understand the meaning of this sentence, he argues, we must consider the conditions that confirm it. Thus, we can construct a set of sentences about our behaviors that define the truth conditions of the statement "Paul has a toothache." The person with a toothache may weep and make gestures indicating a toothache. When asked if he has a toothache, he sincerely affirms that he does. Upon closer inspection, we also find signs of tooth decay, as well as certain changes in blood pressure and central nervous system. It should be noted here that all these behavioral conditions, including physical changes, constitute for Hempel the meaning of what we call a toothache, and are not merely a sign of a mental condition. That is, the mental terms can be translated without loss of meaning into behavioral terms, including physical and verbal behaviors and physiological processes. Paul's toothache is nothing more than a behavioral process.

If behaviorism were an approach that merely advocated a closer relation between mental states and behaviors than we think, there would be no strong objections to this argument. However, philosophical behaviorism is not limited to this, but is an extreme approach that claims that what we call mind is nothing but behavior and behavioral dispositions. For this reason, there are also strong objections to behaviorism. It seems very plausible to argue that given mental states can be different from given behavioral processes or dispositions (Chalmers, 2002, p. 3). To illustrate, in his work “Brains and Behavior”, Putnam imagines some special beings he calls “Super Spartans” He claims that Super-Spartans can "suppress all involuntarily pain behaviour" or behave differently from people who are in pain (1968, p. 9). As seen in Putnam's example, although these beings have real pain, they do not exhibit pain behavior. Conversely, a perfect actor without pain can behave as if it were in pain (Gennaro, 2016, p. 43). These objections show that it is not a plausible approach to explain the mind by reducing it to behavior. Furthermore, Kim claims that some mental states such as pain can be associated with meaningful behaviors such as groaning, screaming etc., but there are mental states that cannot be readily associated with specific behaviors, which makes the behaviorists' idea very problematic (2011, p. 63).

3.2.1.2. Psycho-physical Identity Theory

Psychophysical identity theory, whose best-known advocators are U. T. Place (1956) and J. J. C. Smart (1959), asserts that mental states are the physical brain states. It is mostly accepted view that there is a causal relation or correlation between the mental states and the neuro-physical brain states; however, identity theory goes one step further and claims that mentality is totally identical with the neurobiological process in the brain; they are the same and the one. In other words, each of mental states such as “the desire for a cookie” or “the belief that it is raining” is supposed to be one and identical with the specific brain process in the brain (Feser, 2005, p. 53). It can be more explicitly explained by commonly used example, “pain” and the “C-fiber activation”. According to identity theory, “pain” and “C-fiber activation” are two different concepts, but they are totally same things as in the example of “water is

H_2O". "Water" and "H_2O" are different concepts linguistically, but they are the same things (Kim, 2011, pp. 99-100).

In his work, "Is Consciousness a Brain Process?", Place asserts that

> That is, I am not claiming that statements about sensations and mental images are reducible to or analysable into statements about brain processes, in the way in which 'cognition statements' are analysable into statements about behaviour. To say that statements about consciousness are statements about brain processes is manifestly false. [...] What I do want to assert, however, is that the statement 'consciousness is a process in the brain', although not necessarily true, is not necessarily false. 'Consciousness is a process in the brain', on my view is neither self-contradictory nor self-evident; it is a reasonable scientific hypothesis, in the way that the statement' lightning is a motion of electric charges' is a reasonable scientific hypothesis. (1956, pp. 44-5)

As seen in the above passage, for Place not a conceptual or linguistic reduction, but an ontological reduction of mental states to brain states is a strong possibility. While behaviorism explains mentality through conceptual analysis and verification of our behavioral concepts, identity theory is a claim based on empirical verification of the ontological identity between mental activities and bodily activities. According to this view, mentality cannot be explained by conceptual analysis, that is, unlike the conceptual identity between "bachelor" and "unmarried man", the identity between the mental states and the bodily states is a scientific identity like the identity between "water" and "H_2O", or as Place exemplified in the above passage, between "lightning" and "motion of electric charges".

However, there are objections to the identity of the mental states and the brain states as well. Identity theory implies that if there is not C-fiber activation, then there is not pain as well, but there may be some organisms which do not have C-fiber, but have pain experience. For example, some animals' nervous systems are different from human beings, and although they do not have C-fiber activation, they can have pain experience (Kim, 2011, p. 112). This objection, which is called "multiple realizability" argument by Putnam, will be explained in more detail in the next section, but in brief we can say that Putnam claims that since some creatures like

Martians may have the pain even if they do not have the corresponding brain state, identifying a mental state with a type of physical state is not plausible. The argument of multiple realizability of mental activities makes it necessary to distinguish between two versions of identity theories: Type identity theory and token identity theory. Type and token distinction tells the difference between general types and particular instances. For instance, planet is a type, but Mars or Venus are instances of this type. According to the type identity theory, each type of mental activity is identical to a type of brain activity. The identity theory mentioned above is the explanation of type identity theory in general. However, according to the token identity theory, a particular mental activity is identical with a particular brain activity. That is, for type identity theory, in all circumstances "pain" is a specific physical type, "C-fiber activity", however, for the token identity theory, while an organism's pain state is identical to the E-fiber, another organism's may be identical to a physical particle such as the G-fiber activity. For this reason, while the multiple realizability argument is a strong objection for type identity theory, it seems easier to deal with this objection for the token identity theory.

3.2.1.3. Functionalism

We define many things by their job or function. For example, the heart is defined as the organ that pumps blood to other organs. We define the mousetrap as the tool that catches the mouse, and the knife is identified with its cutting function. When we define something by its function, we are generally concerned with what it does instead of what it is made of. For instance, if the function of the heart is to pump blood, the valves of the heart, the vascular system, the circulatory system are not important for this functional definition. Or for the knife, whose function is to cut, it does not matter what it is made of, as long as it performs the cutting function. There is not a specific kind of material the knife must be made up of; various kinds of material can be suitable for the formation of knife, such as plastic or steel, as long as it fulfils the cutting function (Feser, 2005, p. 55). This means that functional systems are multiply realized by different physical substrates (Chalmers, 1996, p. 247).

In the same way, according to functionalism, we should think of mentality in functional terms. That is, we can explain what we call a mental state by what it does – what its functional role is – rather than what it is made of. Functionalism is an approach that is a descendant of behaviorism and identity theories, but developed in response to them by important philosophers, such as Hilary Putnam, David Armstrong, and David Lewis, in the 1960s (Chalmers, 2002, p. 5). Functionalist theories, contrary to behaviorism, accept the mentality as internal states, and contrary to identity theories, which explains mental states by reducing them to certain physical brain states, functionalist view claims that mental state like pain is not identified with a special physical process in the brain like C-fiber activation, rather it is identified with its function, its causal role in the cognitive system. In his famous article, "Psychological Predicates", Putnam reveals that a function does not match one-to-one with a single type of internal structure, and this is named as the "multiple realizability" argument. The claim of functionalism that mental activities are multiple realizable is one of the important objections to type identity theory. Identification of a mental state with a particular system in the brain is problematic because, as mentioned in the previous section, it may vary depending on the organism. Putnam asserts that

> the physical-chemical state in question must be a possible state of a mammalian brain, a reptilian brain, a mollusc's brain (octopuses are mollusca, and certainly feel pain), etc. At the same time, it must not be a possible (physically possible) state of the brain of any physically possible creature that cannot feel pain. Even if such a state can be found, it must be nomologically certain that it will also be a state of the brain of any extraterrestrial life that may be found that will be capable of feeling pain before we can even entertain the supposition that it may be pain. (1967, p. 164)

Mental activities are functional types, and there is not just one type of internal structure that an organism must have for this function to be performed. The same function can be realized by organisms with different kinds of internal structures. As seen in the above passage, for human being "pain" is identified with C-fiber activation, but for some animals it may correspond to different physical process. Therefore, for functionalist view physical states of the brain is not important to explain the mentality; a state of pain, for instance, can be perfectly explained by its

function. When we have tissue damage, we have pain sense, and pain which is activated tissue damage activates our behavioural responses such as “escape” or “avoidance”. That is, “job description” of pain is “tissue-damage detector” (Kim, 2011, p. 131). In the pain example, tissue damage corresponds to “typical pain input”, and behavioural responses correspond to the “typical pain output”. For a functionalist the concept of pain is identified with the function of pain, and the function of pain is to causally intermediate between the pain input and the pain output (Kim, 2011, p. 133).

The advent of computers and the rise of computer science have been the most important factors in the effectiveness of functionalism. Putnam’s functionalist approach, called machine functionalism, makes an analogy between computing machines and mentality, and it explains the mind by appealing to the operation system of Turing machines which is introduced by the British mathematician and logician Alan M. Turing. The Turing machine is a hypothetical computing device with “a tape divided into square”, a finite alphabet including certain symbols that solves problems and performs calculations in accordance with an instruction involving some general rules that tell what the machines does when a symbol is encountered.[8] It has two separate tapes – the first for the problems it takes as input, and the second for the final information, called the output, produced as a result of calculating the problems – and two separate heads, printer and scanner. Machine functionalism considers the mentality as the Turing machine. That is the Turing machine calculates the symbols on the input tape (“sensory stimulus”) which it receives with its scanner (sense organ), and it emits the final symbols on the output tape (a certain behavior as a response) (Kim, 2011, p. 143). According to machine functionalism, like Turing machine, human mentality is also identified with its causally intermediate function between sensory input and behavioral output.

However, there is a “qualia” obstacle for functionalism. Functionalist view ignores the qualitative characteristics of mental states while identifying them. Since

[8] Kim, in his 2011 work, explains the components and operating system of the Turing machine in detail. For more information about this issue, you can see pp. 139-144.

functionalism identifies the mentality with the functional states, two systems which have the same functional state must have the same mental state. However, according to the one of the major objections to functionalism with respect to the qualia, which is called "Absent Qualia Argument" by Ned Block, although the system in which the qualia is completely missing can have exactly the same functional state with a organism with qualia. In his work, "Troubles with Functionalism", Block illustrates this claim with a robot named with a "homunculi-headed robot" whose outer appearance is the same with the human body but internally different (1978, p. 275). For instance, if we call a particular qualia "Q" and a particular functional state "S_q", even if, "the homunculi headed system" is functionally equal to a conscious individual, i.e., both are in "S_q", a conscious individual can be in "Q" while "homunculi-headed system" cannot be in (1978, p. 278). The other objection which is known as "inverted qualia" or "inverted spectrum" hypothesis claims that "two systems could have the same relevant functional states as a conscious system while having different qualitative states" (Chalmers, 2002, p. 6). This claim is illustrated by Martine Nida-Rümelin with the "psedunormal people" who see green what we see red and see red what we see green. Even though pseudonormal people and the normal people use the same colour names and share the identical functional state, they are exactly in different qualitative states (1996, p. 145-50). And if these objections are true, functionalism is not satisfactory in explaining qualitative mental states by reducing them to functional states.

3.2.1.4. Eliminative Materialism

It would not be wrong to say that eliminative materialism, which defends the falsity of "our ordinary or common-sense understanding of the mind" (Gennaro, 2017, p. 45), is the most radical response to the hard problem of consciousness. Earlier suggestions of eliminative materialism had been made by Paul Feyerabend (1963a; 1963b) and Richard Rorty (1965); however, more recent, prominent proponents of this position are Paul M. Churcland (1981, 1985) and Patricia S. Churchland (1986).
P. M. Churchland asserts that the concepts of consciousness constituted by "folk psychology" will ultimately be replaced with physicalistic concepts "constituted by a

matured and successful neuroscience" (1985, p. 8). He describes "folk psychology" (FP) as the following:

> "Folk psychology" denotes the prescientific, commonsense conceptual framework that all normally socialized humans deploy in order to comprehend, predict, explain, and manipulate the behavior of humans and the higher animals. This framework includes concepts such as belief, desire, pain, pleasure, love, hate, joy, fear, suspicion, memory, recognition, anger, sympathy, intention, and so forth. It embodies our baseline understanding of the cognitive, affective, and purposive nature of people. Considered as a whole, it constitutes our conception of what a person is. (1999, p. 3)

As explained in the above passage, FP is a theory constructs the commonsense conceptual framework to explain the human (or animal) behaviour. It could be said that the one-sentence summary of eliminative materialism is that folk psychology is an inadequate and "a radically false theory" (1981, p. 67).

P. M. Churchland thinks that none of the previous major positions doubts the falsity of FP. The identity theory optimistically argues that FP can be unproblematically reduced to the completed neuroscience. The dualist view, on the other hand, argues that FP cannot be reduced to neuroscience because it corresponds to the non-physical domain of natural phenomena. Functionalism also agrees with its irreducibility, but argues that FP can be explained by functional organizations. However, eliminative materialism, unlike them, argues that FP and "[i]ts principles are radically false", "and that its ontology is an illusion". (1981, p. 72).

The three main reasons P. M. Churchland explained for rejecting FP are: First, FP seems inadequate to explain many kinds of psychological phenomena such as "mental illness", "sleep", "creativity", "intelligence difference", and "memory". Secondly, in its 2500 years history, FP has remained virtually unchanged. The FP used by the ancient Greeks is essentially the same as the one we use. Third and the last, FP cannot be integrated into any of the scientific theories such as physics, chemistry, biology, psychology, and neuroscience, which attempt to explain human behaviour (1999, p. 8). For these reasons, according to eliminative materialism, the irreparably wrong and unrealistic FP is eventually replaced with a better one.

However, there are also some objections for the eliminativist position of materialism. For instance, some argue that the claim of eliminative materialism is self-refuting. That is, the eliminativist believe that mental concepts such as thought, desire, belief or the other concepts of FP do not correspond to the reality. However, if the eliminativist has a belief about the non-existence of “belief” as a mental state, then his claim must be contradictory and eliminativism must be false. As another common objection to eliminativist position is that it is very difficult and utopian to imagine a future in which mental concepts are not used at all (Gennaro, 2017, p. 46).

3.3. “DIME” Dance

As detailed in the above discussion, the available approaches to solving the so-called mind-body problem or the hard problem of consciousness encounter some difficulties. While dualistic approaches could not successfully explain the interaction between mind and body, physicalistic approaches could not fill the explanatory gap that reveals the impossibility of explaining the subjectivity of mental states, which is also emphasised by Nagel, in terms of physical states.

In his 1993 work, *Problems in Philosophy: The Limits of Inquiry*, Colin McGinn explains the inadequacy of the available theories with respect to the problem of consciousness by labelling them in the "DIME" shape. The "DIME" shape corresponds to four types of positions that attempt to explain the origin of consciousness:

“D” corresponds to the domesticating position. According to McGinn, domesticating position tries “to convince us that consciousness is really nothing more than such and such” (1993, p. 32). In other words, this position explains consciousness by identifying it with the unproblematically physical facts. Materialistic approaches, which is examined on the above sections, such as identity theories and functionalism are the most explicit examples of this position. However, such theories ignore the essence of consciousness for McGinn.

"I" corresponds to the position of irreducibility which has respect for the "*sui generis* character of consciousness states" domesticating position ignores (1993, p. 33). According to "I-philosophers", consciousness cannot be explained or defined by reducing it to the physical facts because of its brute, irreducible nature. Although this position defines consciousness as non-miraculous and natural thing, it defends the duality between mental states and physical states. However, the members of I-position cannot satisfactorily explain the tie of consciousness to the physical world.

"M" corresponds to "magical", "miraculous", "mystical" position (1993, p.16). M-philosophers invoke to the supernatural entities like God to explain consciousness since they believe that the world cannot be explained only by the physical laws and the natural forces, there are also supernatural things independently exist from the physical world. The approaches such as Descartes' substance dualism or Leibnizian pre-established harmony are the most explicit examples of M-position since they invoke to the supernatural or divine entities.

"E" corresponds to eliminativism. According to McGinn, when D-position has been shown to be unable to provide a successful explanation for consciousness, and arguments of I-position have started to sound hollow, and M-position have been delusory, the most plausible option left seems to be to deny the existence of the entity which generates the problem (1993, p. 34). E-members rejects the existence of conscious states, for them there is no room for consciousness in the scientific world.

McGinn claims that none of the positions, which is explained above, is convincing; each of them has some deficiencies and problems which lead to the "DIME dance". He asserts that "D yields reluctantly to I", "I encourages a flirtation with M", "M propels one to E", and "outraged common sense then demands a reexamination of D. (1993, p. 17; Inwagen, 1996, p. 254). McGinn calls this endless circulation between D and E "DIME dance".

McGinn introduces the acronym, "DIME", to propose his thesis, Transcendental Naturalism, as an alternative to the available unsatisfactory theories. His thesis

claims that consciousness is actually a natural phenomenon, but our cognitive capacities are inadequate to explain it. According to McGinn, because of this inadequacy, we are doomed to be trapped inside the DIME dance, even if we do our best.

McGinn thinks that “[l]ongstanding historical failure is suggestive but scarcely conclusive” (1989, p. 354). However, I think that the inadequacy of the available traditional approaches to the problem of consciousness does not imply that there will be no explanation of this problem for human being. With his extreme thesis, McGinn seems unaware of the recent studies on consciousness in the science and philosophy. There are some approaches on the table that strongly suggest a possible solution within the framework of naturalism, and none of these is completely successful as of now, but they show us the possibility of solving this problem by taking the nature of consciousness seriously. We will examine these approaches in detail later, but now we will discuss the problematic aspects of McGinn's pessimistic thesis in detail in the next chapter.

CHAPTER IV

NON-CONSTRUCTIVE NATURALISM OF MCGINN

In his work, *The Science of the Mind,* Owen Flanagan makes a distinction between two types of mysterianism: the old mysterianism and the new mysterianism. The old mysterians, e.g., Descartes and Leibniz, claim that consciousness cannot be understood by empirical methods because it has supernatural properties and is controlled by supernatural principles (1991, p: 313). There are also some thinkers who believe that although consciousness is a natural phenomenon operating according to natural principles, human beings will never explain it. Flanagan calls this second type "the new mysterianism" and in his earlier works refers to Thomas Nagel and Colin McGinn as the leaders of the new mysterians (1985, p: 375; 1991, p: 313). However, it is not a reasonable approach to evaluate Nagel and McGinn under the same category because while McGinn is a strong mysterian in terms of achieving the explanation of consciousness, Nagel has only an agnostic attitude towards physical explanations of consciousness due to its subjective nature that resists physical reduction. For this reason, I propose to distinguish between two types of mysterianism: the weak version and the strong version. The main purpose of this chapter is to argue against the strong version of mysterianism. In accordance with this purpose, I will first briefly mention Nagel's attitude towards the consciousness explanation as a weak version of mysterianism. I will then explain and discuss the strong version on the basis of McGinn's theory, Transcendental Naturalism. Finally, I will show that McGinn's theory of consciousness is ill-structured and illusionary, which is why I will call his mysterianism 'pseudo-mysterianism'.

4.1. The Weak Version of Mysterianism

Flanagan describes the new mysterianism as a movement of "a postmodern group" that tries "to drive a railroad spike through the heart of scientism" by defending that

“consciousness will *never* be understood” (1991, p. 313). In his earlier works he equates the position of Nagel with the position of McGinn in terms of mysterianism, but I think they tell quite different stories.[9] While McGinn claims that consciousness can *never* be understood, Nagel’s position is not as strong as McGinn’s; he argues only for the *current,* but *non-permanent* ignorance in explaining consciousness. I call this kind of mysterianism, to which I count Nagel as an advocate, ‘weak mysterianism’. Since the weak version of mysterianism is not the result of an inadequacy in cognitive capacity, it can easily be overcome with the right discoveries and developments in the related fields.

Flanagan calls Nagel “the founder” of the new mysterians and claims that:

> Thomas Nagel was the founder of this group. In his famous paper "What Is It like to Be a Bat?" Nagel argued that there can be no remotely plausible naturalistic account of consciousness, that something essential will always be left out of even our very best theory. Nagel writes, "Consciousness is what makes the mind-body problem really intractable [...] Without consciousness the mind-body problem would be much less interesting. With consciousness it seems hopeless." (1991, p. 313)

It is true that Nagel emphasises the perplexing nature of consciousness in his article “What is it like to be a bat?”, and claims that the current theories, which exclude the essential feature of consciousness, are far from being able to explain it. However, this does not show that Nagel has a pessimistic stance on the future explainability of consciousness.

The main purpose for Nagel is not to close the door on the investigation of consciousness, but to make visible the real reason that makes the explanation of consciousness intractable. In the above mentioned article, by using the example of

[9] The aim of this part is not to argue against the general view of Flanagan regarding the position of Nagel because in his later works Flanagan also changed his view of Nagel’s position, and he identified Nagel as a principled agnostic instead of a new mysterian. I will specifically follow Flanagan’s earlier problematic attitude towards Nagel to emphasise the difference between the weak and the strong version of mysterianism. I do not give a place to Flanagan’s latter term, the principled agnostic because my primary concern is the strong mysterian McGinn; but for the details of his claim you can see his books, *Consciousness Reconsidered* (1993, pp. 1-10, 89-90) and *The Really Hard Problem: Meaning in a Material World* (2007, p. 228).

the bat[10], Nagel tries to reveal that the fugitive character, the subjectivity, of conscious states is the reason the problem of consciousness cannot be explained with the analogies of the physical reductionists. He asserts that

> It is impossible to exclude the phenomenological features of experience from a reduction in the same way that one excludes the phenomenal features of an ordinary substance from a physical or chemical reduction of it […] If physicalism is to be defended, the phenomenological features must themselves be given a physical account. But when we examine their subjective character it seems that such a result is impossible. The reason is that every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view. (1974, p. 437)

As can be deduced from the quotation above the problem emerges from the clash between objective and subjective points of view.[11] In this respect, it would not be a mistake to assume that Nagel transforms the consciousness-brain problem into the subjectivity-objectivity problem. The problem exists because in our present circumstances, we try to explain consciousness by reducing it to our current partial objective concepts that exclude the subjectivity of consciousness, but "[c]onsciousness should be recognized as conceptually irreducible aspect of reality" (Nagel, 1998, p. 337). Nagel suggests that in order to bridge the gap between subjectivity and objectivity, we need to develop new concepts and methods that describe rather than exclude the subjective character of experiences, and he calls this proposal "objective phenomenology" (1974, p. 449).

Flanagan justifies Nagel's mysterianism with Nagel's claim that we do not have objective concepts to express the subjectivity of consciousness:

> In his book *The View from Nowhere*, a work that Nagel describes as "deliberately reactionary," Nagel develops his argument. He insists that "we have *at present* no conception of how a single event or thing could have both physical and phenomenological aspects, or how if it did they might be related. " (1991, p. 313, italics mine)

[10] The details of the example of the "bat" were explained under the title "What does the term consciousness mean?" in the second chapter of this dissertation.

[11] For the details of Nagel's claim, see "Subjective and Objective" in *Mortal Questions* (1979) and "The Limits of Objectivity" in *Tanner Lectures of Human Values* (1980).

To assume someone to be a mysterian in the sense that Flanagan identifies it, there must be evidence to show that he has pushed what he thought to be mysterious outside of our cognitive faculties. However, according to Nagel, we do not at present have suitable objective concepts to explain the subjectivity of consciousness, and this does not mean that the explanation of consciousness will be forever inaccessible to us. On the contrary, he believes that "the mental-physical relations will eventually be expressed in a theory whose fundamental terms cannot be placed clearly in either category" (1974, p. 450, footnote 15). He is not a pessimist regarding the solution of the problem of consciousness; we can easily see that from his following remarks:

> Yet I believe it is not irrational to hope that some day, long after we are all dead, people will be able to observe the operation of the brain and say, with true understanding, "That's what the experience of tasting chocolate looks like from the outside." (Nagel, 1998, p. 337)

As seen from this expression, Nagel is optimistic about the explanation of the states of consciousness; he insists that new concepts are required, but he never gives an argument to justify the inaccessibility of these concepts for human beings. Therefore, it is unreasonable to define Nagel as a mysterian in the McGinn's sense. McGinn himself is also aware of Nagel's optimistic position:

> [A] careful reading of Nagel reveals an optimistic strain in his thought … : see, in particular, the closing remarks of 'What is it Like to be a Bat?', in Mortal Questions. Nagel speculates that we might be able to devise an 'objective phenomenology' that made conscious states more amenable to physical analysis. Unlike me, he does not regard the problem as inherently beyond us. (McGinn, 1989, p. 354, footnote 9)

In the light of the information above, we can clearly see that Nagel does not push the consciousness problem to the dark side; on the contrary, he leaves the door open to the possibility of explaining consciousness with a conceptual revolution in the future. He is a weak mysterian who insists on the current but non-permanent difficulties in explaining consciousness. The really problematic position, which will be discussed in the rest of this chapter, is the position of McGinn who closes the door to all possibilities for the explanation of consciousness.

4.2. The Strong Version of Mysterianism

Colin McGinn is a pioneer of an approach which claims that consciousness will *never* be explained by human beings, and I will call this approach 'strong mysterianism'. Unlike the weak version of mysterianism, strong mysterianism is *permanent*, which is a result of an inadequacy in the cognitive capacity of humankind. In his 1989 article, McGinn describes his extreme approach to the problem of consciousness by saying that "the approach I favour is naturalistic but not constructive". His thesis, which he describes as non-constructive naturalism, is based on three main arguments (1989, p. 352):[12]

i. There exists some property of the brain that accounts naturalistically for consciousness (which I will call 'The Argument for Naturalism')
ii. We are cognitively closed with respect to that property (which I will call 'The Argument for Closure')
iii. There is no philosophical mind-body problem (which I will call 'The Argument for Dissolution')

I will explain his hypothesis by respectively analysing each step.

4.2.1. The Argument for Naturalism

Although the world we live in is completely spatial, consciousness has a non-spatial nature. It can be clearly seen that the brain is the causal basis of our mental lives; it is "the seat of consciousness", or rather it is like "a womb" for consciousness (1999,

[12] McGinn applies the approach that he defined for the problem of consciousness in his 1989 article to other philosophical problems, e.g., free will, the self, meaning and so on, in his 1993 book, and calls this thesis "Transcendental Naturalism" (TN) in which he argues that the solutions to such problems, including the problem of consciousness, are beyond our cognitive capacity. McGinn takes Chomsky's distinction between "problems" and "mysteries" as a starting point for TN. According to Chomsky, while problems are answerable questions within our cognitive faculties, mysteries are questions whose solutions are beyond our cognitive capacities (1976, p. 281; 1998, p. 137; 2013, p. 663). And according to McGinn, philosophical questions are mysteries whose (scientific) explanations are inaccessible by our cognitive faculties. For the details of this argument, see his 1993 and 1994 works, for the rest of the dissertation I will continue with his argument for the problem of consciousness which he especially discussed in his 1989 article.

pp. 4-5). However, there is a fundamental difference between consciousness and its causal basis. While we attribute some spatial properties, such as shape, volume, solidity, extension, and dimension, to the brain, these properties are not compatible with consciousness (1995, p. 149), and the non-spatial character of consciousness creates "a problem of emergence" and "a problem of interaction" (1999, p. 116).

We know, by means of introspection, that there is a strong correlation between the brain and consciousness; the changes in brain activities cause an alteration of conscious states that we are introspectively aware of (1997, p. 45). Even if we know that there is a causal nexus between the brain and consciousness, we do not understand how a piece of meat causes non-spatial consciousness. We feel as if "the water of the physical brain is turned to the wine of consciousness" (1989, p. 349). It is obvious that conscious states do not emerge from brain states in a miraculous way, like a djinn rising from a lamp (1989, p. 352). If we do not have a supernatural position, believing "the magic touch of God's finger" or an eliminativist position that excludes consciousness, we must accept that there must be "some natural property of the brain" through which consciousness is linked to the brain and also a theory which explains this psychophysical nexus (1989, p. 353). McGinn constructs this claim with the following words:

> [T]here exists some property *P*, instantiated by the brain, in virtue of which the brain is the basis of consciousness. Equivalently, there exists some theory *T*, referring to *P*, which fully explains the dependence of conscious states on brain states. (1989, p. 353)

According to McGinn, consciousness itself is a biological phenomenon which naturally arises from the brain, so the relation between the brain and consciousness is as natural as the relation between the liver and bile (p. 362). There is a natural property, *P*, in the brain which is responsible for the link between the mental and the physical.

However, McGinn insists that even though *P* is one of the properties of the brain, "the brain has this property, as it has the property of consciousness" (1989, p. 359),

therefore it must have different features from the physical properties of the brain. It must be non-spatial because so far nothing spatial in the brain convince us that we have found the consciousness-brain link we seek (1989, p. 357). He underlines that

> spatially defined properties […] that seem inherently incapable of resolving the mind-body problem: we cannot link consciousness to the brain in virtue of spatial properties of the brain. […] Consciousness does not seem made up out of smaller spatial processes. (p. 357)

Consciousness is non-spatial property of the brain, and non-spatial property itself cannot be composed of the spatial properties of the brain. For this reason, there must be any other property which is in conformity with the nature of consciousness, and McGinn's *P* is this property in his view.

4.2.2. The Argument for Closure

McGinn claims that even if there is a naturalistic theory of the nexus between the brain and consciousness, his approach to the solution of the consciousness problem is non-constructive (p. 350). That is, he believes that we can never explain the natural property of the brain which makes possible the link between the brain and consciousness. He justifies this claim with his idea of "cognitive closure". According to the cognitive closure idea,

> A type of mind *M* is cognitively closed with respect to a property *P* (or theory *T*) if and only if the concept-forming procedures at *M*'s disposal cannot extend to a grasp of *P* (or an understanding of *T*). (1989, p. 350)

That is, if an organism's concept forming capacity is not sufficient for understanding a property or a theory, the organism is cognitively closed to the property or the theory. Each species has its own cognitive limitations, so a theory or a property may be achievable for some species but cognitively closed to others. For instance, "what is closed to the mind of a rat may be open to the mind of a monkey, and what is open to us may be closed to the monkey" (1989, p. 350). If we consider the property of an electron, while human minds have sufficient cognitive capacity to understand the

property of an electron, monkey minds can never understand what it is; they are cognitively closed to this property.

McGinn claims that different species can achieve different things in the world, and there is no species which is capable of understanding all properties of the world. However, being closed to some minds does not make a property less real than the realisable properties of the world; so human minds being cognitively closed to *P*, does not mean that *P* is not real (p. 351). McGinn supports the possibility of the cognitive closure thesis by following realism, which is an approach assuming an object's existence independently of a mind conceiving it. This is a reasonable starting point. If we do not adopt a dogmatic assumption like idealism, which claims that everything in the world can be understood by humankind, it is reasonable to assume that there are qualities in the world that are independent of our minds. And it is possible that the human mind cannot grasp all the qualities of the world. From this point of view, it is also possible that there are some problems which are unsolvable by human minds due to the qualities that human minds cannot comprehend. However, something is *possible* does not mean that it will be *actual*.

The strong mysterian position not only claims that the problem *may be* unsolvable, but also claims that it *is* unsolvable. McGinn, as a strong mysterian regarding consciousness, claims that the property of the brain making possible the link between the brain and consciousness will never be understood by human beings because of the limitations in their concept forming faculties. The solution to the consciousness-brain problem will forever remain a mystery for humanity. He turns the *possibility* of the cognitive closure idea into an *actuality* and justifies the actuality of cognitive closure by showing the failure of our two distinct cognitive faculties in terms of understanding consciousness.

According to McGinn, there are only "two possible avenues" to understand *P* that will lead us to the solution: introspection and perception (1989, p. 354). That is, we understand the entire world through either "introspection-based" or "perception-based" faculties (2003, p. 158); we do not have any faculty to understand *P* which

would make the consciousness-brain link possible. We can try to understand *P* either by introspectively having direct knowledge of our mental states, or by conducting observations on the brain, but none of them would allow us to understand *P*.

Let us begin with introspection. Through introspection we become aware of the properties of consciousness; we can immediately access our mental states, and thus we can form concepts from our inner lives. However, this is not sufficient cognitive faculty to achieve *P*, because *P* has a mediating function between the brain and consciousness, but "introspection does not present conscious states *as* depending upon the brain in some intelligible way"; it gives us only one side of the consciousness-brain nexus (1989, p. 354). A further point is that we cannot obtain the concept *P* by analysing on the concepts of consciousness because we can only form the concept for the state of consciousness we can access by our own introspection. For instance, we cannot form the concept for the echolocational state of the bat; it is the bat's own subjective experience. Similarly, a blind person cannot comprehend the concept for an experience of 'red' she has never seen (1989, p. 355). In the same way, the conceptualisation of *P* as a solution of consciousness problem is beyond the realm of introspection; the concept of *P* cannot be structured. In the light of this explanation, McGinn claims that "the faculty of introspection, as a concept-forming capacity, is cognitively closed with respect to *P*" (1989, p. 355).

Another avenue through which we can try to reach *P* is the faculty of perception. However, according to McGinn, we cannot get a better result than by introspection because the function of the perception is to form our comprehension of the brain. For the investigation of the brain we must use our senses, but senses only present to us objects of space with their spatial properties. McGinn claims that

> You can stare into a living conscious brain, your own or someone else's, and see there a wide variety of unstantiated properties-its shape, colour, texture, etc.-but you will not thereby see what the subject is experiencing, the conscious state itself. Conscious states are simply not potential objects of perception: they depend upon the brain but they cannot be observed by directing the senses onto the brain. In other words, consciousness is noumenal with respect to perception of the brain. (1989, p. 357)

Senses are effective to respond to spatial properties, but because nothing spatial can solve the consciousness problem, *P* must be a non-spatial property as consciousness itself. For this reason, according to McGinn, we cannot understand how *P* provides the consciousness-brain nexus and explain consciousness by the investigation of spatial properties.

Being perceptually closed to something does not entail being cognitively closed to it, since the method of concept formation may work towards introducing theoretical concepts for *P* (1989, p. 358). However, McGinn claims that this method is also useless to give us *P*. We cannot obtain a theoretical concept of *P* by making an inference from the observational data, because we use the "principle of homogeneity" while we generate theoretical concepts based on our observations. The principle of homogeneity declares that the theoretical concept we formulate by means of the inference from physical data must be homogenous with the entity of our observation. He asserts that "if our data, arrived at by perception of the brain, do not include anything that brings conscious states, then the theoretical properties we need to explain these data will not include conscious states either" (1989, p. 358). According to McGinn, this method works for concept formation of unobservable material entities. Although they are not the objects of observation, we introduce theoretical concepts about them by analogical extension of the physical data we do observe. However, *P* cannot be given by an analogical extension of the observable entities of the brain; neither consciousness itself nor *P* are entities of the same kind as the spatial properties of the brain, so their theoretical formation by perceptual brain data is prohibited by the homogeneity principle (1989, p. 358-9). For this reason, *P* is not only perceptually but also cognitively closed to us.

Consequently, according to McGinn, these two distinct concept-forming faculties, introspection and perception, present us with a partial picture of the relation between the brain and consciousness; they thus fail to reveal the underlying property that unifies the brain and consciousness. For McGinn, "it is a bit like having to view an elephant either from the tail end or the trunk end and never being allowed to take in the whole elephant" (1999, p. 48). Our partial knowledge about the brain-

consciousness nexus is the result of the inherent limitations in our concept-forming faculties and unfortunately there is no way to change or extend them. Therefore, the brain-consciousness problem is condemned to remain a mystery for us.

4.2.3. The Argument for Dissolution

McGinn has a pessimistic position with respect to achieving a constructive solution to the consciousness-brain problem. However, he believes that we do not need to understand the nexus between the brain and consciousness to remove the philosophical perplexity. He claims that just knowing that there is a natural solution to the problem of consciousness in the science to which we will never have access to because of our inherent limitations, also removes the philosophical consciousness problem:

> The philosophical problem about consciousness and the brain arises from a sense that we are compelled to accept that nature contains miracles-as if the merely metallic lamp of the brain could really spirit into existence the Djin of consciousness. But we do not need to accept this: we can rest secure in the knowledge that some (unknowable) property of the brain makes everything fall into place. What creates the philosophical puzzlement is the assumption that the problem must somehow be scientific but that any science we can come up with will represent things as utterly miraculous. And the solution is to recognize that the sense of miracle comes from us and not from the world. There is, in reality, nothing mysterious about how the brain generates consciousness. There is no metaphysical problem. (1989, pp. 362-3)

Consciousness is not more complex than processes like “digestion” or “sexual reproduction”, and the brain’s release of consciousness is as natural as the liver’s secretion of bile. The philosophical perplexity about the relation between the brain and consciousness is based on our sense that this nexus arose in a miraculous way. However, this sense, according to McGinn, is the result of the inadequacy of our cognitive capacity, and not the result of a miraculous world. That is, the mystery is not ontological but epistemological in the sense that the science which explains consciousness is beyond our cognitive limitations, and being aware of these limitations is enough to remove the philosophical consciousness-brain problem.

4.3. Pseudo-Mysterianism

McGinn attempts to prove the non-constructive but naturalistic solution of the consciousness-brain problem by his hypothesis. Even if I find McGinn's insistence that there is a naturalist solution to the consciousness problem rational, I think his thesis that this solution is noumenal for us is unacceptable. He implies that it is time to wave the white flag for philosophers who ponder consciousness in hopes of contributing to its solution. The nexus between the brain and consciousness is something mysterious that will forever remain on the dark side of the world for us because of our limited cognitive capacity. This is not a digestible claim. McGinn constructs his mysterianism with his cognitive closure thesis. However, I will reveal that the main argument he uses to justify his cognitive closure thesis is inconclusive and his mysterianism is illusionary. For this purpose, I will first argue for the problematic aspect of *P* he identified as providing the nexus between the brain and consciousness. I will then argue against his cognitive closure thesis which is at the core of his mysterianism.

4.3.1. Objections to the Argument for Naturalism

As explained above McGinn identifies a property of the brain, *P*, making consciousness a natural product of the brain. *P* "does for consciousness and the brain what gravity does for the planets and their orbits, or what kinetic energy does for molecules and the behaviour of the gases they compose" (1991, pp. 58-9); that is, it removes the (ontological) mystery of the emergence of consciousness from the brain and provides a naturalistic solution. According to him, *P* is the property not only of the brain but also of consciousness because it must have sufficiently similar features with both to connect these two distinct characters to one another.

McGinn insists that even though *P* is a property of the brain, it must also be non-spatial like consciousness because spatial properties of the brain are not effective to mediate between the brain and consciousness. However, some commentators argue that the non-spatial character of *P* has similar characteristics with the supernatural

position of Cartesian dualism, and thus contradicts the naturalist solution of the consciousness-brain problem. For instance, Brueckner and Beroukhim claim that

> [W]hat becomes of McGinn's claim to have given a naturalistic solution to the mind-body problem, a solution that is preferable to Cartesianism? […] To say that P is inaccessible to our best possible physicists' minds is one thing, but to say that P (along with consciousness) is non-spatial is another. If P is non-spatial in character, then it is hard to see what its being a *natural* property comes to, if not just being a real property of things. According to the Cartesian, properties of non-physical mental states and substances are natural in *that* sense. (2003, pp. 403-4)

As seen in the above quotation, Brueckner and Beroukhim defend their thoughts on the non-naturalist position of McGinn's thesis by comparing McGinn's non-spatial property with the supernatural substance of Cartesian dualism. Cartesian dualism constructs the mind-body problem by claiming that mental states are the states of the non-spatial substance which is distinct from the body. From this point of view, according to Brueckner and Beroukhim, if McGinn's position is naturalistic, then we must assume that Cartesian dualists are naturalist in this sense, otherwise we have to accept that McGinn's thesis is also non-naturalistic like Cartesian dualism.[13]

However, I think that interpreting McGinn's position in this way is a result of missing some crucial points in his thesis. McGinn's *P* and the Cartesian substance are quite different from each other. Let us first explore the reason why we assume that the Cartesian position is supernatural. There are two main claims of Cartesian dualism: (1) the mind and the brain are two separate substances, and (2) the mind is a non-spatial substance and therefore different from the brain. These two claims are logically independent from each other. The existence of (1) does not necessitate (2). Descartes, the most famous defender of Cartesian dualism, could have identified two different spatial substances, but he assumed that the brain is an extended substance, and the mind is a non-extended thinking substance. The reason we assume Cartesian dualism is supernatural is not the existence of different substances but the non-spatial character of one of them. McGinn's position also seems non-natural because of the

[13] In his article, "Against McGinn's Mysterianism", Demircioglu also argues against McGinn's naturalistic position, and I also replied to Demircioglu in a similar way. For the detail see Işıkgil (2017).

non-spatial *P*, but he protects his thesis from such an objection with some additional explanations. On the one hand, McGinn insists on the non-spatial character of *P* for the reason that the spatial property of the brain does not have the mediating feature that links the brain and consciousness. On the other hand, he tries to reveal that his concept of non-spatiality corresponds to the *seeming* non-spatiality which is the result of our cognitive limitations. That is, according to him, because of our cognitive limitations we cannot understand the objective space itself because the objective space also naturally involves non-spatiality. He asserts that

> It is not that consciousness is non-spatial, after all; rather, space is quite other than we think, and consciousness fits comfortably into the nature of space as it really is. So, when I repeated that the mind has no spatial properties, I must be taken to have meant that it does not have the spatial properties we attribute to space, which is consistent with saying that it has the properties that space objectively has. (1999, p. 123)

McGinn defends his idea regarding the objective space with the possibility of a pre-Big Bang universe. Following the claims of cosmologists, he explains that before the Big Bang there was no spatial universe; the universe had the dimension of non-spatiality, and space came along with matter after the Big Bang. This means that the origin of spatiality was not spatial itself but non-spatial or pre-spatial; a transformation from the non-spatiality to spatiality during the Bing Bang. However, McGinn claims that during the transformation of the current universe during the Big Bang, the earlier state of the universe did not entirely disappear from the universe; the dimension of pre-Big Bang, non-spatiality, is conserved in some form like consciousness in the animal or human brain (1999, pp. 120-1). Even though objective space also naturally includes the dimension of non-spatiality, due to our cognitive limitations we conceive space as if it excludes the feature of consciousness. The non-spatial character of consciousness is the result of our ignorance of what real space is. That is, the clash between the non-spatial consciousness and the spatial brain is not an ontological fact but only an epistemological one (2004, p. 108). The *seemingly* non-spatial consciousness is included in space due to space's "hidden dimensions" which originated from the earlier state of the universe.

Whether McGinn's explanations are reasonable or not is a matter for another debate, but in the light of these explanations, it does not seem plausible to define his position as supernatural in the Cartesian sense. McGinn implies that the space we perceive does not correspond to real space, and this idea makes McGinn's position totally different from the Cartesian view in terms of the naturalistic explanation of consciousness. McGinn himself also defines the Cartesian attitude as supernatural with the following words:

> The other form [of purported solutions to the problem], which has been historically dominant, frankly admits that nothing merely natural could do the job, and suggests instead that we invoke supernatural entities or divine interventions. Thus we have Cartesian dualism and Leibnizian pre-established harmony. These 'solutions' at least recognize that something pretty remarkable is needed if the mind-body relation is to be made sense of; they are as extreme as the problem. (1989, p. 350)

As seen in the quotation, because the Cartesian view admits that "nothing merely natural could do the job" McGinn identifies the Cartesian position as non-naturalistic. In this sentence "nothing natural" does not correspond merely to "nothing spatial" because McGinn also believed that nothing currently conceived as spatial could explain consciousness. However, the Cartesian approach asserts that "no brain property" could explain consciousness, and this is the reason why the position of Cartesian dualism cannot be assumed natural in the same sense as McGinn's position. According to the Cartesian view, the mind comes into being not through the properties of the brain, but "through some quite different agency, possibly God's" (1999, p. 118), and such assumptions push dualism out of the reach of naturalism. McGinn's position is also dualistic with respect to properties of the brain, but this is quite different from the classical version of dualism. Assuming that consciousness is a part of natural space means that space is not simply composed of the physical things in the sense that we understand, and this is a naturalised version of dualism. Consequently, while the Cartesian approach insists on the exclusion of consciousness from space, McGinn insists on the hidden inclusion of consciousness in space, and this crucial distinction keeps McGinn's position from being non-natural like the Cartesian approach.

I think that the non-spatiality of McGinn's *P* does not create a contradiction with respect to McGinn's naturalistic position, but this does not mean that his naturalistic thesis is satisfactory for the actual solution of the brain-consciousness problem. I also believe that there are some problems in the characterisation of *P*. To be able to mediate between two different character there must be a property which is sufficiently homogeneous with each, and McGinn insists that spatial properties of the brain cannot be adequate to provide the link between the brain and consciousness because they are not homogeneous with non-spatial consciousness. However, I think that McGinn's *P* is also unsatisfactory with respect to providing a mediating role. Even though McGinn says that "*P* must be a property both of the brain and of consciousness, since its role is to link the one to the other" (1991, p. 60), he characterises *P* the same as consciousness. He identifies *P* as non-spatial property of the brain as consciousness is, and because of its non-spatial character *P* is unperceivable as consciousness itself is. In this case, the question of how a property which has the same features with consciousness can relate to the brain is as perplexing as the question of how consciousness arises from the brain. Regarding this issue, Whiteley also claims that

> I am not convinced that McGinn's property P could solve [the problem]. If variations in P can occur independently of the spatial properties of the brain, it cannot play its intended part as a mediator between brain and consciousness in a theory accounting for the physical determination of conscious states. If, on the other hand, its variations are wholly determined by these spatial properties, shall we not have the same sort of misfit between cause and effect which disturbed us in the causal relation between brain and consciousness? (1990, p. 394)

As seen in the above quotation, Whiteley shares the view that a property which provides the link between two heterogenous things must be homogeneous with each thing, but he also expresses the failure of McGinn's *P* in terms of a mediating role. A necessary condition for a mediating *P* is at least one common feature which is shared between both consciousness and *P*, and *P* and the brain, but *P* seems to have nothing in common with properties of the brain. So then, why do we need a property that has the same features as consciousness? What can such a property do for us other than increase the number of problems we face?

I think that McGinn's *P* not only fails in its mediating role, but it also raises two other problems besides the consciousness-brain problem: the problem of *P*-brain and the problem of *P*-consciousness. Until now, we were looking for what kind of connection there is between consciousness and the brain, but now we do not know what kind of connection there is between *P*-brain and between *P*-consciousness. About this issue Hanson also argues that:

> McGinn has in effect merely replaced one unintelligible connection with two: first, the unintelligible connection between the spatial properties of the brain and P, and second, the unintelligible connection between the mysterious P and consciousness. (2010, p. 583)

I totally agree with Hanson about *P*'s unintelligible connections, and he also makes an important point by saying that this situation may lead to a regress (2010, p. 583). If we continue with McGinn's idea, to explain these new connections, new properties must be proposed. As a natural explanation for the brain-consciousness relation, McGinn proposes an epistemically mysterious *P*, and maybe he could propose P^1 for the *P*-brain relation, and P^2 for the *P*-consciousness relation.

Consequently, it can be said that the non-spatiality of McGinn's *P* is not in contradiction with his naturalistic position regarding the solution of the brain-consciousness problem, but its one-sided harmony undermines its mediating role and renders it unnecessary.[14]

4.3.2. Objections to the Argument for Closure

McGinn claims to defend a naturalistic approach to the problem of consciousness, but according to him, this natural and simple solution of consciousness is noumenal for human beings with respect to the cognitive faculties, and this is the reason why McGinn calls his thesis transcendental naturalism. McGinn establishes his cognitive

[14] There are limited direct objections to the nature of McGinn's *P* in the literature. Apart from the objections I have discussed in this section you can see Garvey (1997). Garvey insists that McGinn's *P* is one of the ordinary properties of the brain, but while establishing his thesis McGinn clearly explains why the property providing the brain-consciousness link cannot be ordinary property of the brain. As an objection to Garvey's argument, see Worley (2000).

closure thesis, which is the core of his mysterianism, with a double-stage method. As detailed above, he firstly tries to show the possibility of cognitive closure, and then he claims the actuality of human cognitive closure by appealing to two distinct concept forming faculties: introspection and perception. However, his cognitive closure thesis extending from possibility to actuality has some problems. In this section, I will first try to reveal that the arguments he mounts to justify the cognitive closure thesis are inconclusive, and second, I will show that his mysterianism regarding the consciousness problem is pseudo-mysterianism by replacing the cognitive closure with the psychological closure.

McGinn tries to justify the possibility of human cognitive closure with respect to the brain-consciousness link by making an analogy between animals and human beings. He claims that "minds are biological products like bodies, and like bodies they come in different shapes and sizes, more or less capacious, more or less suited to certain cognitive tasks" (1989, p. 350). Each biological species has its own cognitive limitations, and a problem that is cognitively closed in one species may be open to another. For instance, a monkey mind cannot gain access to the knowledge of what an electron is. Similarly, a rat mind cannot understand solutions to trigonometry problems, while a human mind is open to them. However, McGinn claims that there is no guarantee that the human mind is powerful enough to understand the solution of every problem in the world:

> [W]e are natural beings, descended from apes, living in a natural world, and our capacities are as finite as can be. We can, it is true, do more with our mind than apes can, but that does not mean that we somehow magically escape the constraints of biology. We are animals all the way down, or up, not angels. (1999, p. 42)

This means that in the same way that monkeys cannot understand the concept of electrons or rats cannot understand the solutions of trigonometry, human beings also cannot understand some properties of the world like the property which makes possible the solution of the brain-consciousness problem.

I also do not rule out the possibility that the human mind, as a biological system, is cognitively closed to some things in the universe. I think that there may be very interesting and important properties of the universe that we are not aware of, do not know what they are about, and have any conceptual knowledge to ask questions about them. However, a consciousness that we can think and formulate questions about, and do research on some aspects of, does not seem to be among these properties. It is not reasonable to argue for the possibility of human cognitive closure with respect to the solution of the problem of consciousness by making an analogy between animals and humans because there is a radical difference between animals' closedness to some domains of the world and humans' closedness to the solution of consciousness. While human beings can formulate and understand the problems of consciousness, animals are unable to formulate and understand the problems about the domains they are cognitively closed to. That is, in contrast to animals, human beings are aware of what they are ignorant of. It is not surprising that monkeys are closed to the concept of electrons, and that rats cannot solve trigonometry problems, or dogs ponder on political problems, since none of these cases exist as problems in their world. Dennett (1991, 1995) and Kriegel (2003, 2009) also draw attention to this distinction between animals and human beings. According to them, McGinn fails to justify the possibility of human cognitive closure with the analogy argument because he ignores our linguistic ability which is the most important difference between the human mind and the animal mind. The passages I have quoted below clearly show us their attitude towards McGinn's cognitive closure claim:

> Monkeys, for instance, can't grasp the concept of an electron, McGinn reminds us, but I think we should be unimpressed by the example, for not only can the monkey not understand the answers about electrons, it can't understand the questions (Dennett 1991d). The monkey isn't baffled, not even a little bit. We definitely understand the questions about free will and consciousness well enough to know what we're baffled by (if we are), so until […] McGinn can provide us with clear cases of animals (or people) who can be baffled by questions whose true answers could not unbaffle them, they have given us no evidence of the reality or even likelihood of "cognitive closure" in human beings. (1995, pp. 382-3)

Like Dennett, Kriegel also insists that considering other species cannot be evidence for the possibility of human cognitive closure with respect to consciousness. He says that

> Rats' minds do not understand trigonometry. Nobody would want to deny that. But trigonometric problems do not pose themselves to rats. Indeed, that rats' minds do not understand trigonometry is precisely why trigonometric problems do not pose themselves to rats. For trigonometric problems pose themselves to rats, rats' minds would have to understand a great deal of trigonometry. (2003, p. 183)

As mentioned in the above quotations, monkeys are cognitively closed in terms of the concept of electrons because there is not a question like "what is an electron" in a monkey's mind; they cannot comprehend such a question. The property of an electron has never existed in the life of monkeys as a problem that needs to be solved. However, unlike monkeys (or any other species), "we understand the unanswered question about consciousness" (Dennett, 1991). Kriegel emphasises that "there are conceptual connections between understanding a question and understanding its possible answers" (2003, p. 184). That is, if we do not have innate limitations to understand a problem itself, this situation is necessarily coupled with being cognitively open to its possible solutions as well. This does not mean that understanding a problem also requires knowledge of the correct one among its possible solutions. It only requires understanding what its solution would be, and with the right development and discoveries we can ultimately achieve the correct solution. It does not seem to be reasonable to assume that with our cognitive capacity, which is powerful enough to comprehend the problem, the solution of the problem can never be understandable for us. Considering all of these, it seems rather weak to base the claim that humans are cognitively closed in terms of the problem of consciousness on the analogy between humans and animals' closure.

In order to prevent this weak position, McGinn tries to turn the possibility of cognitive closure into actuality by claiming that humankind does not have the required cognitive faculty to understand the solution to the consciousness problem. In McGinn 1989, his closure argument goes like this: We can achieve *P* providing the brain-consciousness link neither by introspection and introspection-based

concepts nor by perception and perception-based concepts, so the solution of the consciousness problem will remain for us forever on the dark side of the world. However, I think that this claim is too strong and without foundation. As explained in detail in previous sections, the main argument he proposes in order to show that we cannot reach *P* and introduce its theoretical concepts with the observation of the brain is the principle of homogeneity.

McGinn uses the homogeneity principle as a protection against "a form of magical emergentism with respect to concept formation" (1989, p. 358). However, this principle unfortunately "works by restricting inferential space to only the simplest, most straightforward sorts of connections" (Flanagan, 1993, p. 113). That is, by observing the brain we can only find the links between spatial brain states and introduce theoretical concepts about these data, and by introspection we can only comment on our own conscious states and introduce theoretical concepts about them. This seems undoubtably true, but while introducing theoretical concepts about our observational data, we invoke third avenue such as "investigating consciousness indirectly via third person observation of behaviour and associated reasoning" (Hanson, 2010, p. 583). Even if introspection has a prominent role in shaping our concept of consciousness, it is clear that consciousness does not present itself only through introspection; we can examine the presence of consciousness in other creatures through its manifestation in their behaviours. And although McGinn claims that "to explain the observed physical data we need only such theoretical properties as bear upon those data, not the property that explains consciousness, which does not occur in data" (1989, p. 359), in the explanation of some of the observable data we invoke both the concepts of consciousness and the concepts of the brain properties. Flanagan illustrates this with the case of blindsight. Blindsight patients have lesions in their primary visual cortex (V1), also known as the striate cortex, but they can give statistically significant responses to visual stimuli that they cannot consciously see. In other words, when these patients are asked questions about the objects in the direction they look, they correctly answer the questions about the objects. They could also perform some actions correctly and identify the features of the objects in the direction they were looking, even though they say that they see nothing. In this case,

as opposed to McGinn's claim, to explain the observed data we need to appeal both to the properties of consciousness (or lack of them), and to the properties of the brain to reveal the lesion in the visual cortex (1993, p. 114). The differences between fully-sighted people and blindsight patients with respect to visual consciousness, and the reason why visual consciousness is lacking in the blindsight cases while fully-sighted people have it, cannot be explained by purely physical data; theoretical explanations for such cases warrants a simultaneous appeal to both neural properties and conscious properties.

Besides this, humans are creatures who can think about, and introduce concepts for, the entities they may never have perceived. McGinn also admits that in our theories we refer to properties that are not the objects of our perception. However, McGinn insists that we introduce theoretical concepts about unobservable entities which comply with the procedure of the homogeneity principle. He claims that we form theoretical concepts about such entities by making an analogical extension from the physical data we observe, but because *P* is not a spatial property we can observe, we cannot form any theoretical concept about *P*. However, McGinn's homogeneity principle is excessively restrictive. We form concepts about unobservable objects by making an inference from our observations, but to say that our method of inference and concept formation is based solely on the analogy with the physical objects we observe would be to take science back to the dark ages. McGinn's homogeneity principle may be successful for the formation of theoretical concepts of unobservable objects at a relatively macro level, but as Perez says, there are also theoretical concepts from the field of the quantum physics at the microstructural level (e.g., "energy", "spin", "valance" etc.) that we form without the basis of an analogy with the spatial entities we observe (2005, p. 39). If we were producing concepts only through the analogical extension as McGinn claimed, we would be closed to many scientific concepts that we have today. Therefore, the homogeneity principle, which is the basis of McGinn's cognitive closure thesis, seems rather unwarranted.[15]

[15] There are also some other objections in the literature that directly target McGinn's cognitive closure thesis. For these see Kirk (1991), Sacks (1994), Kukla (1995), Megill (2005) and Velerick and Boudry (2017).

As can be seen in the above discussions, McGinn's closure thesis is merely an inconclusive assumption; he gives us no evidence or plausible reason why we should accept that the problem of consciousness is outside of our cognitive power. As Dennett (1991) says in the review of McGinn' book: "In order to defend a thesis about the outer limits of our powers, one should at least take a peek at the concepts made available to those who have armed themselves with the new technology". Unfortunately, however, McGinn does not address the scientific concepts related to the working system of the brain during his defence of cognitive closure thesis. In order to defend a claim that the solution of any problem is cognitively closed to us, it needs to be proved that the problem is surveyable neither by our current scientific tools, nor by any possible future science. McGinn himself also describes *P* providing the link between the brain and consciousness as a natural denizen of the objective space. Then if our scientific knowledge of space can be adequately advanced, the seemingly non-spatial consciousness can be unproblematically surveyable as a physical subject of future science. Even our current scientific studies, both in quantum physics and developments in information theories, are considerably promising in revealing that consciousness can be explainable on a physical basis.

There must be another reason underlying McGinn's persistent defence of the cognitive closure thesis, and why he ignores all the developments in research of consciousness and the changing structure of science that progresses with paradigm shifts in the historical process. I think that McGinn's closure thesis, which he constructs without any acceptable data or scientific evidence, is not a cognitive closure which is the result of the human mind's inadequacy, but a 'psychological closure' which is the result of McGinn's desire to drop the subject. When a person encounters a problem, and if it does not lead to a conclusion within his cognitive system, he may feel uncomfortable and have a desire to stop and conclude the search for knowledge on a particular subject. That is, what I mean by psychological closure

is an individual's need and desire to close and finalise the subject that gives uncertainty.[16]

I think that McGinn attributes his own need for closure with respect to the problem of consciousness to the cognitive inadequacy of all humanity. In his 1989 and 1993 works, McGinn frames the consciousness-brain problem with the following words:

> [The mind-brain link] strikes us as miraculous, eerie, even faintly comic. Somehow, we feel, the water of the physical brain is turned into the wine of consciousness, but we draw a total blank on the nature of this conversion. (1989, p. 149)
>
> Consciousness puzzles us in a special way. […] The head spins in theoretical disarray; no explanatory model suggests itself; bizarre ontologies loom. There is a feeling of intense confusion, but no clear idea about where the confusion lies. (1993, p. 28)

As can be seen in the above quotations, while talking about the problem of consciousness, McGinn characterises some states that he calls "philosophical bewilderment", such as "eeriness of the mind-brain link", "head spins in theoretical disarray", "feeling of intense confusion" etc., and his main aim is to remove the intense confusion created by the problem of consciousness and to immediately reach a final conclusion about this problem. For this purpose, instead of contributing to the possible solutions of the problem he prefers to stop thinking about consciousness and puts forward the thesis of cognitive closure, which is the short-cut dissolution.

In the light of the above discussion, I think that McGinn's cognitive closure thesis, which is not based on any evidence, is an illusionary thesis which is the result of his need to reach a conclusion quickly, and the epistemic mysterianism built on his closure thesis is pseudo-mysterianism.

[16] For this argument, I was inspired by the psychological and sociological analysis which belongs to Kruglanski et al. Kruglanski uses the statement of "Need for Cognitive Closure (NFCC)" to describe the individual's need to reach for a certainty instead of confusion and uncertainty regarding a particular subject (Kruglanski 1990, Webster and Kruglanski 1994, 1998). In other words, the need for cognitive closure is the individual's need to simplify complex information and to avoid uncertainty when faced with a problem.

CHAPTER V

A CONSTRUCTIVE AND NATURALISTIC SOLUTION IS POSSIBLE

As has been seen so far, the subjective side of consciousness, qualia, or phenomenal consciousness is embraced by some who classify it as something that cannot be physically identified. In this case, there is silence about the interaction between the physical brain and non-physical consciousness. Others describe consciousness by reducing it to the physical properties of the brain, in which case the subjective part of consciousness is completely ignored. While this unsuccessful dichotomy pushes some to eliminate the existence of phenomenal consciousness, McGinn accepts that consciousness has a physical explanation that preserves its subjective side; however, as a strong mysterian, he believes that no matter how hard we study the neural correlates of consciousness, the human mind still cannot understand how physical processes of the brain generate consciousness. He pessimistically argues that the physical explanation of phenomenal consciousness is forever beyond our understanding.

As has been claimed from the beginning, the purpose of this dissertation is to reveal that McGinn's epistemological mysterianism is excessive and that there is a promising approach that can explain phenomenal consciousness on a physical basis. This approach, called Integrated Information Theory (IIT), not only preserves the intuitive data that consciousness is subjective, direct, and immediate, but also aims to show that consciousness can be explained on a physical basis by identifying consciousness with the capacity of a system to integrate information. This chapter of the dissertation will firstly explain the main argument of IIT in detail and then discuss how the IIT model explains the subjective character of consciousness in the physical system. However, I will begin by briefly analysing the place of consciousness in the space-time unity of the world.

5.1. The Place of Consciousness in Space-Time Unity

Einstein's Special Theory of Relativity tells us that time and space cannot be separated from each other. In other words, in relativity

> there is not a sufficiently clear distinction between space and time […]. You can't be just in time, and you can't be just in space, since on this view there's only one thing: space-time. Space-time is only artificially divisible into separate dimensions, and so you're either completely in it, and therefore in space and time, or you're not in it at all. (Lee, 2007, p. 341)

Although classical Cartesian theory accepts that mentality exists in time, it denies its existence in space. However, time and space are like an interwoven continuum with respect to each other, and in this respect, conscious states must be spatio-temporal. There is no clear distinction between time and space, and therefore it is not plausible that something exists only in time and not in space, or vice versa.

Although relativity tells us that consciousness must be spatio-temporal, it does not tell us how it can fit into space-time. Thinking that consciousness exists in space does not mean accepting that physicalism in the traditional sense is true. Invoking traditional physical identity theories to show that the existence of consciousness is compatible with space-time coexistence would mean that what was at stake regarding the hard problem of consciousness (the subjective/phenomenal side of consciousness) remains at stake. There are, however, other paths which are compatible with space-time coexistence; ones which neither try to explain the subjectivity of consciousness in the reductive way of classical physics, nor completely ignore the subjective character of consciousness.

As discussed in the third chapter, there are many conflicting theories about the consciousness-brain relationship, from substance dualism to property dualism, and from identity theories to eliminative materialism, but none of these theories have come up with a promising and satisfactory explanation of how consciousness, which is thought to be mental, takes place in the physical world. The subjective side of consciousness seems to stubbornly resist scientific explanation. However, it is much

more difficult for some because the essential reason of this failure is the belief that our current mental (subjective) and physical (objective) concepts are insufficient to accurately explain the mind-body relationship. Nagel is the oldest and best-known of those who have expressed the concern that our current physical concepts are objective concepts, and objectivity does not represent the subjective aspect of consciousness. Nagel firmly accepts the subjectivity of consciousness and the existence of a first-person perspective. Rather than arguing that physicalism is wrong, he argues that physicalism must become compatible with the existence of subjectivity and that the commonly accepted theory of consciousness must explain the subjective character of consciousness. Even if Nagel thinks that in today's sense, physicalism—which we explain with objective concepts—ignores the subjective side of consciousness and is insufficient to explain it, he still asserts that the new concepts can be developed which include the subjectivity of consciousness. In *The View From Nowhere*, Nagel refers to this thesis as the "objective understanding of the mind" (1986, p. 17). It is the distinction between subjective concepts and objective concepts that leads to the seeming impossibility of an objective explanation of the mind. For Nagel, giving a scientific and naturalistic explanation of the subjective character of experience seems possible with the expansion and the revision of our conceptual framework. This is the essence of Nagel's theory of objective phenomenology.

According to Nagel, we can bridge the gap between the mental and the physical and find a place for consciousness in a physical world with the formation of new concepts which include the subjective aspect of consciousness as well. However, it would be more accurate to take it a step further and say it in this way: in the light of developing technology, research, and advances in the sciences, and if we revise what we understand by "being physical", we will be able to produce an unproblematic physical and scientific explanation for the phenomenal consciousness that refers to the subjective aspect of consciousness. In fact, Nagel's suggestion for the formation of new concepts to include the subjective character of consciousness also depends on the change in what we understand when we say the physical world.

Instead of thinking (like McGinn) that the problem of consciousness will forever remain unsolved (due to the inability of traditional approaches to explain consciousness on a spatial basis by preserving its subjective aspect), it would be more reasonable to think about what the concept of "spatial" means and to pursue the question of whether the knowledge we have about space is sufficient. Our current knowledge of the limits of the spatial and the limits of real space— which exist beyond what we know—are not the subject of this thesis, and should be left as a research topic for physicists and scientists. However, the important thing for us is that as our knowledge of space develops and changes (through new investigations, developing technologies, and advances in the scientific world), the spatiality of consciousness can not only be explained scientifically but can also be compatible with the view of space-time coexistence. Therefore, even if it is accepted that there is only one kind of fundamental category of existence, and if it is revealed that this category is not spatial as we currently know it, but also spatial in reality, the explanation of consciousness— whose fundamental parts are not ordinary physical entities, but exist in time and space like ordinary physical entities—will not be mysterious to us at all.

Until now, classical physicalist theories have tried to fit consciousness into rigid physical patterns in order to solve the puzzle of consciousness. Firstly, they defined the characteristics of the physical world and believed that the non-physical or the non-spatial character of consciousness was problematic because it was not compatible with the entirely physical world as formulated below:

(1) The world we live in is physical world.
(2) If the world is physical, then the objects in the world and their properties are also physical and must comply with the laws of physics.
(3) Consciousness is the property of the brain which is a denizen of the physical world.
(4) Consciousness is non-physical and does not comply with the laws of physics.

In this formulation, the inference we must take, by following (1), (2), and (3), is contradicted by (4). That is, if we assume the truth of each statement in this formulation, then there is a contradiction. We call this contradiction "the hard problem of consciousness", and until now we have generally tried to solve this problem by trying to tame the *sui generis* nature of consciousness to conform to physical rules. However, the problem of consciousness emerges from its formulation. Formulating the problem by assuming that our knowledge of the first three premises is complete, unsurprisingly leads to the problem of the seemingly non-physical nature of consciousness because we cannot understand how non-physical consciousness emerges from the physical brain. We cannot understand this because we cannot fit consciousness into the physical world.

The real starting point should not be the physical world but rather the nature of consciousness. A successful theory of the explanation of consciousness initially requires an analysis of the nature of consciousness to be put forward, followed by an investigation into the physical foundations that protect this nature rather than ignoring it. The Integrated Information Theory, developed by neurologist Giulio Tononi in 2004 to explain consciousness, is a promotive example of such a theory. This theory tries to develop an approach based on what we know about consciousness, instead of starting from the physical basis. Tononi accepts that consciousness intrinsically exists; everyone can experience their own consciousness since it is absolutely subjective, and consciousness is an indivisible whole. Tononi starts from these facts and thinks that human consciousness is more than mere neural firing. The essential thing is the information passing through these networks, not the neuron itself; consciousness derives from the integrated information in the related mechanism of the brain. Let us now examine the details of this theory more closely.

5.2. The Integrated Information Theory of Consciousness

The Integrated Information Theory, which was initially proposed by neuroscientist and psychiatrist Giulio Tononi and his colleagues, aims to reveal what consciousness

is and how it can be physically explained[17]. According to the IIT model, to explain consciousness we need to solve the two main problems. The first problem concerns "the conditions that determine to what extend a system has consciousness" (2004, p. 2). Indeed, the solution to this problem will also give us answers to questions such as: What does consciousness mean and under what conditions can we talk about consciousness? This was referred to in the second chapter. With the solution to this problem, we will be able to know the amount of consciousness a system can produce; we will have measured "the quantity or level of consciousness". We can call this problem "the quantity problem of consciousness". The second problem is related to the conditions that determine the quality of the consciousness produced by a system, i.e., how to understand the conditions that construct the subjective character of conscious experience, which is the main subject of this dissertation. We can call this problem "the quality (or qualia) problem of consciousness" and the solution to this problem will provide the explanation for how the system generates the subjective character or quality of consciousness.

The main idea of the IIT is constructed on the claim that "consciousness corresponds to the capacity of a system to integrate information" (2004, p. 1). According to this theory, both the quantity problem and the quality problem of consciousness can be explained by the amount and specific way the information integrates with the system. If the system's way of integrating information sheds light on the explanation of the subjective aspect of consciousness, even if there are objectionable or missing aspects of the theory, it can be considered as an answer to McGinn's mysterainist thesis which obscures the understanding of consciousness, since it at least makes constructive and naturalistic solutions of consciousness possible. I will now explain how a constructive and naturalistic solution of consciousness is possible in the light of the IIT, and for this I will begin with the quantity problem of consciousness.

[17] Giulio Tononi is the founder of the IIT and pioneered it with his work with Gerald Edelman, but Christof Koch is also one of the major advocates of IIT, especially with his long-term work with Francis Crick. For the details, see (Koch&Tononi, 2013). However, in this chapter, for the analysis of IIT, which is a promising approach to explain phenomenal consciousness, I will continue with the arguments and works of Tononi, the founder of this thesis.

5.2.1. The Quantity Problem of Consciousness

The two essential properties on which the IIT is based are information and integration. Before analysing the construction of the IIT on the basis of these properties, let us begin with the description of information**.**

5.2.1.1. Information

Information theory, or the mathematical theory of information, was first formulated in the 1940s (under the leadership of Claude Elwood Shannon and Warren Weaver) to examine the rules on obtaining, processing, and transferring information. Information is classically described as the "reduction of uncertainty". The greater the number of alternative outcomes, the greater the reduction of uncertainty and thus the greater the amount of information. In information theory, the results of a random experiment (the information that any outcome carries in the experiment) is calculated as the inverse logarithm of the probability of that outcome occurring as formulated below:

$$H= - \sum pi \, . \log_2^{pi}$$

This function is called the "entropy function". In this function "Φ" refers to the number of alternative outcomes of an experience, and "pi" represents the probability of each of the outcomes of this experiment. For example, according to the entropy function, the state of tossing a coin corresponds to:

$$H= - (½ \, . \log_2^{1/2} + ½ \, . \log_2^{1/2})$$
$$H= - (\log_2^{1/2})$$
$$H= - (\log_2 2^{-1})$$
$$H= \log_2^{2}$$
$$H= 1 \text{ bit of information}$$

because there are only two alternative outcomes of this experiment, and each outcome has a probability of 1/2 in this experiment. However, when throwing a dice there are six probabilities and this state corresponds to "$\log_2^6$ = 2.59 bits of information" (2008, p. 217). As the number of probabilities increases, the amount of information which is produced by the experiment also increases.

Tononi explains the importance of the information produced by an experiment with "the photodiode thought experiment" (2004, p. 2-3; 2008, p. 217-18). In this experiment, a blank screen is shown to both a photodiode and a person, and they are requested to say "light" when the screen is on and "dark" when it is off. While the person can consciously distinguish when the screen is on and off, the photodiode makes this distinction with a photosensitive sensor and a detector that tells light from dark depending on this sensor; while the person has the subjective experience of seeing light or dark, the photodiode cannot have this experience, and according to the IIT, this fundamental difference between the person and the photodiode has to do with the amount of information produced when making the distinction. In other words, the essential question is: How much information is produced when a photodiode or a person detects light or dark?

If we use the concepts of classical information theory, and if the results are the same in both cases, the information produced by the photodiode and the subject will be the same, i.e., one bit of information. However, according to the IIT model, a different path is followed for this entropy calculation. The photodiode selects one of two options: there is light or dark. When the light and dark states are equally probable, the mean information entropy of the photodiode, with respect to logarithm base, corresponds to one bit of information. However, when a person looks at the screen, she not only detects the light as a photodiode does, but also has the experience of the lightness. The reason a person, unlike a photodiode, has an experience is because the lightness of the screen is just one of the numerous alternatives that a person can experience. For example, the screen is not just light or dark, but can also be green or red; there can be this or that movie or many different frames on the screen, such as a tree, a mountain, or a landscape. In fact, there are an extraordinarily large number of

probabilities corresponding to the frames that can be seen on the screen. The person is not only choosing between light and dark, but is choosing from a much wider range of alternatives. Therefore, the person's statement of “the screen is light” or “the screen is dark” carries higher information than one bit. In the IIT model—as opposed to the classical description mentioned above—information is actually “the ability to discriminate among a large number of alternatives” (2008, p. 218); the difference between the person’s conscious response to light and the response of the photodiode is to be able discriminate between many alternatives. While the person can have “the *meaning* of the discrimination” between dark and light, “the photodiode has no mechanism to discriminate colored from achromatic light, even less to tell which particular color the light might be” (2008, p. 218). Therefore, unlike the photodiode, the person can specify that whatever she discriminates “is not colored (in any particular color), does not have a shape (any particular one), is visual as opposed to auditory or olfactory, sensory as opposed to thought-like and so on” (2008, p. 218).

According to the thought experiment mentioned above, the key difference between the photodiode and the person in terms of consciously seeing is to be able to differentiate one state among many (2004, p. 3; 2008, p. 218). However, even if it is necessary for the presence of conscious experience to be able to specify one alternative as distinct from many alternatives— the information level— it is not sufficient. Additionally, Tononi uses another thought experiment to explain this idea (2004, p. 3; 2008, p. 218). This time, the person is compared to an ideal digital camera whose sensors consist of one million photodiodes. The detectors of such a camera have the potential to discriminate among $2^{1.000.000}$ states, and the entropy of this corresponds to one million bits of information. This means that the camera could respond differently to every movie frame shown on the screen. However, this does not mean that the digital camera is a conscious being.

According to the IIT, the essential condition for the presence of conscious experience and the difference between the person and the camera is “information integration”. The camera could be thought of as a unity or single system which is capable of distinguishing $2^{1.000.000}$ states; but it is not. Each of the camera’s one million

photodiodes makes its own independent distinction. There is no interaction between the photodiodes of the camera; therefore, each state which is differentiated within the photodiode of the camera is causally independent from the others. This means that there is no information which is integrated among the photodiodes of the camera: millions of photodiodes in the camera respond not causally but independently of each other. The failure of a certain number of detectors in the camera corresponds to the same loss of information regardless of which detectors or which pixels are corrupted. The performance of the camera does not change with such a failure. However, for the person, every piece of information in the scene is part of a unity: the information is integrated. When the person consciously sees a frame on the screen, that image is “an integrated whole” and cannot be separated and experienced individually (2004, p. 3; 2008, p. 219). For a person, each part of the scene can carry information about another part of the scene. This is because there are causal interactions between the elements of the brain of the person, and due to this causal interaction, the state which is differentiated by the element of the brain is dependent upon the states of other elements. For this reason, the person’s brain, unlike the camera, can integrate the information, and again, unlike the camera, when the connection between the elements in the person's brain becomes inoperative, this has a disruptive effect on the conscious experience. Tononi claims that

> The integration of information in conscious experience is evident phenomenologically: when you consciously "see" a certain image, that image is experienced as an integrated whole and cannot be subdivided into component images that are experienced independently. For example, no matter how hard you try, for example, you cannot experience colors independent of shapes, or the left half of the visual field of view independently of the right half. (2004, p. 3)

As Tononi argues that consciousness is a result of an integrated whole, and we also know that in the “red square” experience we do not experience the “redness” without the experience of the shape of the square, or the shape of the square independently of its “redness”, the red square experience is integrated information of its “redness” and its “square shape”. The system of our brain informs us about the nature of our “red square” experience—its redness and its shape—by discriminating it from numerous possible alternatives; a green cube for example.

In summary, from all the above we can say that even if the amount of information is important, according to the IIT it is not adequate. To be able to measure "the repertoire of states" the system can achieve, we use the entropy function as explained above. This kind of measurement cannot give us anything concerning the information integration, and what is essential for the presence of consciousness is the ability to integrate information: "Consciousness has to do with the ability to integrate a large amount of information" (2004, p. 3). The conscious experience is generated by the way of neurons' physical communications in the systems of the brain which are connected with each other by cause-effect power. However, what is really important is how such a system is identified by the IIT model. Let us look at this.

5.2.1.2. The Physical System Integrating Information

The IIT insists that the presence of consciousness cannot be explained by starting from the physical matter because there is not, as classical identity theories claim, an identity between mental states and neural states (2017, p. 250). The IIT takes an opposite attitude by starting from the experience itself, and instead of beginning with identifying neural correlations of consciousness in the brain, Tononi, at first, identifies the essential characteristics of conscious experience.

5.2.1.2.1. Essential Characteristics of Phenomenal Existence: Axioms[18]

Before Tononi investigates the necessary and sufficient conditions for a system to be conscious, in his 2017 article, he gives five axioms to briefly explain the essential characteristics of phenomenal existence (or experience):

[18] The difference between axiom and postulate is that although these are common notions as general statements and general truths, postulates correspond to general truths specific to geometry. That distinction, however, in modern mathematics and modern geometry somehow disappeared. In the contemporary literature axiom is the only concept in use, so at first glance it seem that Tononi appeals to an outdated distinction. Yet, the way Tononi uses axiom-postulate concept pair is the following: axioms are properties of the phenomenal experience, postulates are their physical implementations or physical requirements that may implement those properties.

i. Intrinsic Existence: The first axiom is related with the existence of consciousness. Consciousness has intrinsic, undeniable existence and we can immediately and directly know it.

ii. Composition: The second axiom emphasises that every experience has "internal structure", consists of more than one phenomenal component and that these are interconnected in various ways. When we consider a visual experience, the same experience includes various experience components such as different colors and dimensional experience. Each experience has its own constitutive elements and internal structure.

iii. Information: The third axiom is related with the specific form of each experience. The composition parts of each experience are interconnected in a specific way, and this way distinguishes an experience from others.

iv. Integration: The fourth axiom emphasises the unity and irreducibility of consciousness. Consciousness is a result of an integration of its many components. The structure of an experience, that is, the constitutive elements of a conscious state are experienced as a unity. And a conscious state is not an experience that can be reduced to its components.

v. Exclusion: According to fifth axiom, conscious experience is definite in content. That is, it has only the content it has; there can be no more or less than this content. For example, my visual experience is a composition of different constitutive elements such as different colors, and dimension experience. These elements cannot be lacking in my visual experience, and more cannot be added to these components like blood pressure during this experience.

5.2.1.2.2. Postulates

According to the IIT, what is certain is the existence of consciousness and its essential properties, and this theory takes an approach from the phenomenal to the physical.

It should be recalled that according to the first axiom, consciousness certainly exists. The first postulate claims that if there is consciousness, there must also be a physical substrate that is in a cause-effect relationship with its environment. It cannot be assumed that there is something in the physical world that cannot make a difference on nothing, or anything can make a difference on it (2017, p. 245). Therefore, the existence of a physical substrate that affects and is affected by its environment and produces cause-effect relationships is a presupposition regarding this axiom.

According to the second axiom, there is an internal structure of consciousness, and experience is the unity of its compositional parts. And according to the second postulate, the physical processes related to consciousness are also structured, and constitutive parts within this structure should have a cause-effect relationship with their environment. For instance, let us suppose that a system ABC is composed of its constitutive parts A, B, C, AB, AC, BC, and this means that in the unity of ABC, each of these compositional parts also has a cause-effect power within the whole (2017, p. 245).

As explained in the previous section, the third axiom concerns the specific form of each conscious experience. Previously, we described how the IIT model identifies information and its measure: entropy. In the experiment in which the photodiode and the person determine the lightness of the screen, we explained that the person selects the light and dark options among many possibilities such as green screen, red screen, a frame of a mountain, a frame of a landscape, and much more. The components of the experience are distinguished from a multitude of possibilities; that is, the experience is specified by the system within the numerous alternatives. The third

postulate says that in the physical system, the structure of the cause-effect power between the subunits or the compositional parts also has a specific form.

As mentioned above, according to the fourth axiom, consciousness emerges as a result of the integration of many components and cannot be reduced to its parts. As explained in the example of the red square, the experiences of redness and the shape of square, which are experienced separately, do not lead to the red square experience. In other words, the red square experience is not the sum of the redness experience and the square experience. One inference we can make from this is that the physical substratum of consciousness must also be indivisible and irreducible to its components. The IIT model expresses this indivisibility and irreducibility of the physical system to its components with a measurement of "integrated information" and denoted by the symbol "Φ" (phi). This symbol indicates that "the information (the vertical bar 'I') is integrated within a single entity (the circle 'O')" (2004. p. 4). The main idea here is that when the system is completely separated into its components, there will be the greatest reduction in integrated information. If we look at the same case in reverse, the undivided cause-effect repertoire is where integrated information is at its highest.

The fifth axiom continues the idea that consciousness is the integration of information. As mentioned in the previous section, Tononi calls this the "exclusion principle". Exclusion affects the "maximality principle for a cause-effect power" of the physical system. The system excludes the competing structures and only one cause-effect structure exists over the other structures. That is, the structure "wins the competition for intrinsic existence" and has a maximal amount of irreducible integrated information (Φ^{max}) over the other structures (2017, p. 247).

5.2.1.2.3. The Conceptual Structure and The Complex

In the light of the axioms that describe the characteristics of consciousness and postulates that analyse the properties of the physical substrate of consciousness, IIT constructs an identity statement and claims that "every experience is identical with a

conceptual structure". A conceptual structure corresponds to the concepts and their communications, and an irreducible cause-effect space is composed of the conceptual structure. That is, it can be thought that "a conceptual structure" is "a form in cause-effect space", and this means that the conscious experience is "a form in cause-effect space" (2017, p. 248). Tononi explains and illustrates the conceptual structure with the following figure[19]:

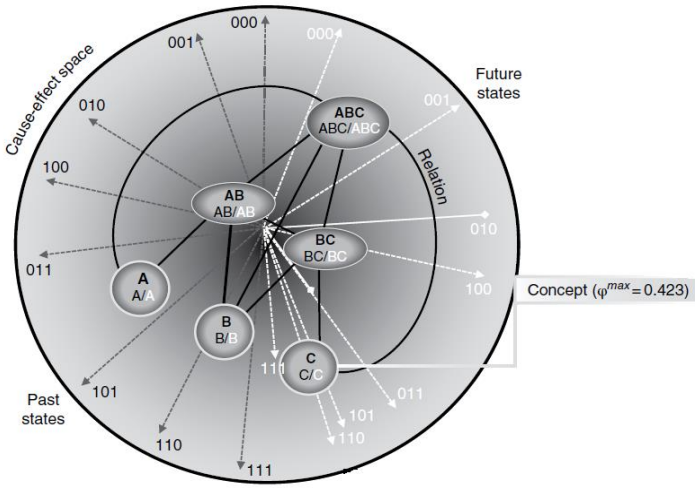

Figure 1. The model corresponding to the conceptual structure of the physical system

In the above figure, the elements A, B, C, AB, BC etc. correspond to the concepts, the connections of these concepts with each other, and also their self connections which constitute the cause-effect space. Tononi likens each of the concepts to a star, and a conceptual structure is a "constellation of stars", the brightness of it depending on "its irreducible Φ^{max}". According to the IIT, "the system of mechanism that specifies a conceptual structure is called a complex" (2017, p.248). That is, the physical system, which is called complex, distiguishes elements of conceptual structure from other possible alternatives, and they are individual physical entities that are capable of integrating information. A physical mechanism like a brain is ideal for having more than one complex. It contains multiple, small or large complexes with a lower Φ, but at any given time the main complex, which has a

[19] This figure is a part of the Figure 17.1, which Tononi used in his 2017 article to explain and illustrate the integrated information system (p. 249).

comperatively higher Φ, generates the dominant experience for the mechanism (2008, p. 221).

5.2.2. The Qualia Problem of Consciousness

Let us imagine that the pianist and composer, Fazıl Say, and I listen to Beethoven's 9th symphony. The tunes of Beethoven—the input—are the same, but as an ordinary person listening to the music, the information I receive while listening to Beethoven is significantly less than the information that Fazıl Say will receive, because his brain's way of integrating information while listening to Beethoven will not be the same as mine. This example illustrates the subjectivity of conscious experience in the IIT.

After the IIT clarifies what conditions a system must have in order to be conscious, it moves on to what Chalmers defines as "the hard problem of consciousness": the problem of how and why subjective consciousness (or phenomenal experience) arises from the physical brain.

> What determines that colors look the way they do, and different from the way music sounds, or pain feels? And why can we not even imagine what a "sixth" sense would feel like? Or consider the conscious experience of others. Does a gifted musician experience the sound of an orchestra the same way you do, or is his experience richer? And what about bats? Assuming that they are conscious, how do they experience the world they sense through echolocation? Is their experience of the world vision-like, audition-like, or completely alien to us? (2004, p. 6)

As previously mentioned, like Nagel, Chalmers, and McGinn, Tononi also starts with the subjective character of consciousness, and he emphasises the private, specific, and irreducible qualities of conscious experience. He agrees with Nagel's famous expression that consciousness is "what it is like to be", and he claims that "since information can only be integrated within a complex and not outside its boundaries, consciousness as information integration is necessarily subjective, private, and related to a single point of view or perspective" (2004, p. 6).

The IIT explains the quality of consciousness using informational relations between elements as shown in the following figure[20]:

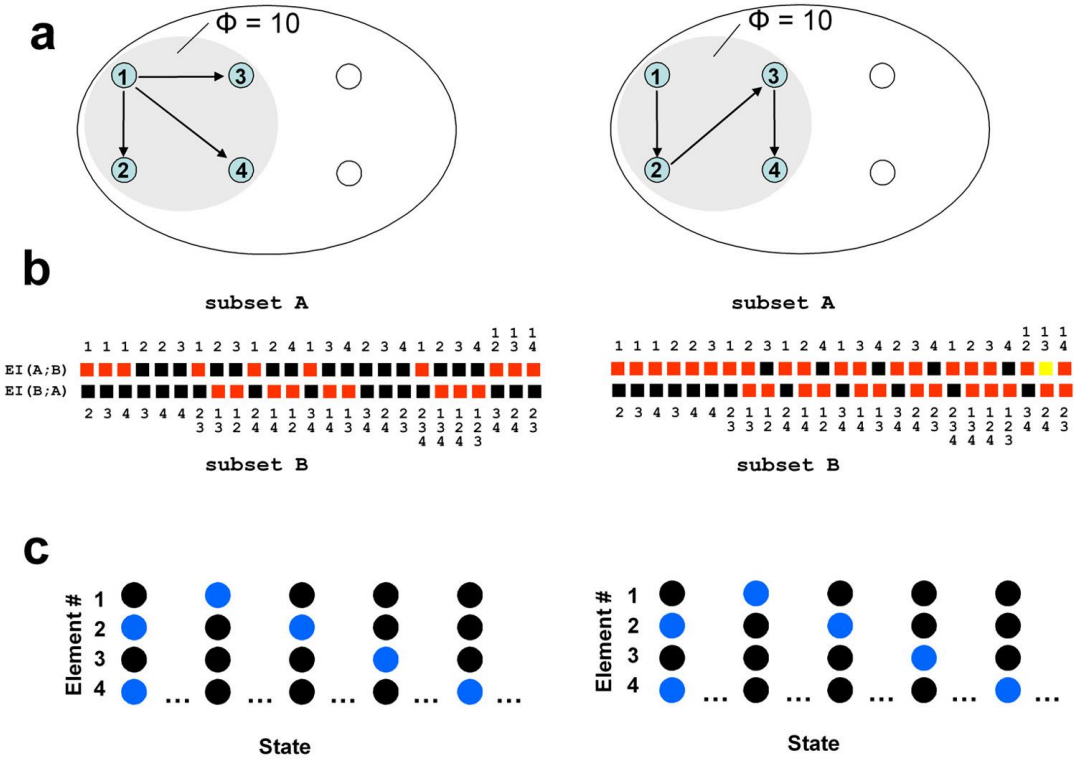

Figure 2. Qualitative character of consciousness: Two complexes having the same activity states and Φ value but different way of integrating information

Figure *a* exemplifies two different systems and the causal interactions of the elements in these systems. On the one hand, the system on the left shows that there is a divergent digraph between elements: element 1 is connected to the other elements, 2, 3, and 4, with an equal power; this system constructs a single complex, and its Φ value is 10 bits. On the other hand, the system on the right represents the "chain link" between the elements. That is, element "1 is connected to 2, which is connected to 3, which is connected to 4". Similar to the system on the right, this system also constructs a single complex with Φ=10 bits. Figure *b* represents the matrix of "effective information" values of both systems. That is, it represents the value of the informational relation between all the elements. The first line, "from A to B",

[20] This figure is taken from Tononi's 2004 article (p. 8, figure 2).

represents one direction of the information relation between elements, and the second line, “from B to A”, represents the reciprocity of the relation. Figure *c* represents the states— “the intrinsic dynamics of the system” or “inputs from the environment”—at a given time, and these five states are identical for each physical system. The given matrix is coded with three different colours. While black reflects the “zero” with respect to the value of the effective information, red corresponds to the “intermediate value” and yellow corresponds to the “high value”. As seen in the above, the color-coded matrix of both systems is different from each other. For instance, there are more zero entries for the divergent complex than the chain complex. When we examine the diagram, we see that even though these two systems have the identical input and the same Φ value, both have 10 bits of information integration value; the informational relations of both systems, which correspond to their qualitative characters, is different from one to the other (2004, pp. 7-9). Tononi claims that

> These two examples are purely meant to illustrate how the space of informational relationships within a complex can be captured by the effective information matrix, and how that space can differ for two complexes having similar amounts of Φ and the same number of dimensions. (2004, p. 8)

Undoubtedly, having a high phi value has a significant effect on having a very large qualia space, but the main claim of the IIT is that phenomenality is directly related to the structure of the informational relationship within the complex. This means that due to the structural difference in the qualia space of the two systems, the identical state of both complexes corresponds to different experiences for each of them at a given time. Just as the amount of information integration specifies the quantity of consciousness, the structure of all informational relations specifies the quality of consciousness. In other words, “how integrated information is generated within a complex determines not only the amount of consciousness it has, but also what kind of consciousness” (2008, p. 224).

Let us recall that the photodiode example given in the previous section distinguishing the state of light on the screen, does not mean only choosing a state of light that is the opposite of dark among the alternatives, but that the mechanism in the person's

brain also specifies what it is like to see the light for that person at any given time. This is the mechanism's specific contribution to the generation of conscious experience.

As has been emphasised in this chapter, IIT begins with phenomenology. It attempts to make its physical explanation by starting from the characteristic features of the phenomenal. To summarise, the first of the two basic claims of this theory is that consciousness corresponds to the information integration capacity of the brain's neural system, which we call complex, and that the amount of integrated information can also be measured and represented by the symbol Φ, called phi. We are organisms that can have a state of consciousness of many things: while watching a theatre scene, we can see the flower represented by the painting on the wall, as well as being able to discern any sound coming from outside. When we think of all the movie frames we have watched or will watch, just like the example we gave in the photodiode experiment, each frame has a specific state of consciousness for us. This means that the mechanism of our brain differentiates the inner state into numerable possible alternatives and this information is highly integrated within the system of the brain to generate conscious experience. Because of the cause-effect interaction between the related parts of our brain, we experience a state as a unity. For example, no matter how hard we try, we cannot consciously perceive the right side of the area we are in and fail to ignore the left side; or when we look at the face of a crying friend, it is unlikely that we do not notice her crying. The information we are conscious of cannot be divided into subparts; it is given as a complete whole. Let us think about the syndrome, which is called a split brain. It is a picture that occurs as a result of surgical cutting of the corpus callosum (connecting the two hemispheres) for the treatment of uncontrollable epilepsy. After this surgery, it is seen that the consciousness of the patients is divided into two; their experiences are simultaneous, but they are not integrated as a whole. Patients perceive the stimuli in their right and left visual fields differently and cannot form a unity; the patient's right hemisphere is aware of the left visual field, and the left hemisphere is aware of the right visual field. However, the patient is unaware of the stimulus presented to him as a whole. In this case, split-brain patients do not have phenomenal unity since "the 200 million

wires linking the two cortical hemispheres" have been cut. (Koch, 2009). This example also reveals that to be conscious there should be communication between elements in the neural complex of our brain; we need an integrated entity.

Furthermore, the second and main claim that we seek to answer in this thesis is on the question of whether the subjective side of consciousness can be explained scientifically. With his IIT theory, Tononi argues that this can be explained scientifically, and explains it by identifying qualia with the structure of the informational relations of elements within a complex. According to Tononi, even if the states and information integration values for two different complexes (or at different times for the same complex) are the same, the way of integrating information for both complexes is different, and this situation reflects the private and subjective aspect of consciousness.

The IIT may be problematic because it might be considered too complicated to measure the quantity of consciousness. However, when we look at the advantages of the theory rather than its incompleteness, IIT can be considered as a study that approaches the problem of consciousness holistically since it tries to explain consciousness in the physical system without ignoring the subjectivity (qualia) of experiences. The second advantage is that it argues that consciousness should be handled in an inclusive way not only for humans, but also for non-human creatures and even non-living things. For the IIT, the condition, that consciousness depends on the integrated information capacity of a system, is true "whether or not it has a strong sense of self, language, emotion, a body, or is immersed in an environment, contrary to some common intuitions" (2004, p. 20). Since it is not related to language, it becomes possible to talk about the existence of consciousness in cases where there is no verbal report, and with this claim it explains the existence of consciousness in animals and infants, as well as some neurological phenomena such as comas or sleepwalking. In addition, this theory claims that even though the input and the output signals are weakened in REM sleep, consciousness continues to be preserved (2004, p. 20). These examples can be defined as conditions where the phi value is low. In other words, according to the IIT, the experience takes place in the case of

phi max. As the integration capacity of the system increases, so does synergy and consciousness. People do not report their experiences below the phi max, but this does not mean that the phi value reaches zero.

5.2.3. Information is Physical

Obviously, to show that consciousness is objectively debatable by scientists and philosophers, it must be shown that consciousness is a physical phenomenon. As we have previously analysed, many theories tried to explain consciousness using classical physics by reducing it to matter, but this created an explanatory gap with respect to the subjective consciousness. Therefore, the main issue is: what kind of physical entity corresponds to the state of consciousness and its subjective aspect? If this entity is defined as a physical entity in the sense that we currently understand it, then reducing the subjective aspect of consciousness to a physical entity will not take us any further than traditional identity theories. However, there are opinions that claim that consciousness is physical in a different sense than our current understanding.

Although we are not yet aware of it, these studies and theories show us that a silent revolution has taken place in theoretical physics. Physicists have long avoided arguing about consciousness, because unfounded thoughts about the hard problem of subjective consciousness may be enough to end a physicist's career. However, this view began to change due to a new and fundamentally different idea that was spreading rapidly in the physics community. This idea is the claim put forward by the neuroscientist Tononi, where consciousness is an integrated information as detailed in this chapter. One of the two salient features of a system that exhibits consciousness is that the system must be able to differentiate and integrate large amounts of information. The second is that this information must be a single whole in such a way that it cannot be separated into independent parts. The fact that these two properties of a theory can be expressed mathematically made it possible for physicists to work on them. Although the problem of consciousness has yet to be

fully resolved, due to the IIT it has finally come close to being formulated as a set of mathematical problems that researchers can understand, explore, and discuss.

The physicality of information gives us a chance to explain phenomenal consciousness without any room for superstition or despair, and there are important physicists trying to prove that information—which is a mathematical concept—is a measurable physical entity. For instance, in his 1961 work, Landauer argues that, by emphasising the relation between thermodynamics and information, measurable information is not just an abstract mathematical concept but a physical representation.[21] And later, the German physicist, Melvin M. Vopson, developed the theory where, based on Landauer's work, he defines information as the fifth state of matter, claiming that information is physical. He asserts that

> In fact, one could argue that information is a distinct form of matter, or the 5th state, along the other four observable solid, liquid, gas, and plasma states of matter. It is expected that this work will stimulate further theoretical and experimental research, bringing the scientific community one-step closer to understanding the abstract nature of matter, energy, and information in the Universe. (2019)

Vopson explains and constructs his theory in an article published in 2019 (based on Einstein's general theory of relativity), by formulating "a new principle of mass-energy-information equivalence proposing that a bit of information is not just physical, as already demonstrated, but it has a finite and quantifiable mass while it stores information". The theory proposes that every piece of information has a finite and measurable mass (2020).[22] Vopson's theory means that information is the fifth state of matter, and this idea completes the missing part of physics in a new and exciting way. The physicality of information, in the sense that we cannot currently describe changes our understanding of space and spatiality.

[21] For the details see Landauer's 1961 and 1996 works.

[22] For the details of the theories of Vopson see his 2019, 2020, 2021 works.

Alongside Vopson, Max Tegmark, the famous physicist of the Massachusetts Institute of Technology, also shares the view that consciousness can be expressed within the framework of integrated information theory. He thinks that by combining quantum mechanics and information theory, some questions about the nature of reality in consciousness can be explored by scientific experimental methods. Tegmark, like Landauer and Vapson, explains information as a measurable physical entity.[23] He shows how the basic properties of consciousness can arise from the laws that govern the universe, and states that due to these properties, physicists can form an idea about the conditions that give rise to consciousness.

If we leave the explanation and proof of information as a physical entity to physicists, and return to the contribution of this knowledge and the existence of such a study, the claim that information is a physical entity in a different sense than we currently understand it, tells us that it is possible to expand our spatial perception without contradicting physical theories, that there are fundamentally more than four states of matter, and that consciousness can be explained by such a physical being.

[23] In his theory on consciousness, Tegmark deals with the concept he calls "perceptronium", which he defines as the most general component that is subjectively self-aware. This component not only stores and processes information, but also provides an indivisible integrity, and with "perceptronium" Tegmark describes consciousness as a state of matter. I will not go into the details of his thesis, but for the details, see his 2015 work, 'Consciousness as a State of Matter'.

CHAPTER VI

CONCLUSION

The relationship between consciousness and brain, in other words, the duality between the mental and the physical, has been one of the most fundamental metaphysical and epistemological problems of philosophy. However, the interest and the orientation of science to this field began much later than philosophy. With the development of advanced scanning and examination techniques related to brain functions in the last thirty years, studies in the field of neuroscience on consciousness have increased. With increasing studies in the field of neuroscience, although there is a general belief that the classical dualism (Cartesian dualism) based on the mind-body distinction is not valid, that is, mind and body, brain and consciousness cannot be two different substances, phenomenal consciousness (subjective and qualitative aspect of consciousness) is still a "hard problem" for the scientific investigations. According to the basic premise of natural sciences, quantifiable, measurable and mathematical properties of nature provide a single, precise and objective knowledge of the world. However, our subjective experiences of consciousness such as pain, fear, or pleasure from a work of art are accepted as qualities that cannot be measured quantitatively and are seen outside the study of scientific field because they cannot form a reliable basis in scientific studies.

There are many approaches that have attempted to explain the problem of consciousness. Some of these traditional theories accept the subjective character of consciousness but failed to explain the relationship between the subjective consciousness and the physical brain, while others try to explain consciousness in a physical sense by reducing it to neurophysiological states, but in this case, they ignore the qualitative-subjective aspect of consciousness.

The persistent failure of traditional theories has led some philosophers, like Colin McGinn, to believe that there is no humanly accessible solution to the hard problem of consciousness. According to McGinn, even if there is an explanation of the relation between the brain and consciousness, we as human beings cannot achieve this because of limited cognitive faculties. He believes that the link between the brain and consciousness is natural and simple like the link between the liver and bile. However, we have just two distinct cognitive faculties, introspection and perception, and they cannot be successful to explain property of the brain mediating between consciousness and brain. In other words, according to McGinn's view, we are cognitively closed to the scientific explanation of consciousness.

However, the aim of this dissertation is to reveal that McGinn's epistemic mysterianism with respect to the explanation of consciousness-brain relation is pseudo-mysterianism since McGinn cognitive closure idea that constructs his mysterianism is groundless and excessive argument. With a groundless claim, McGinn closes the books on the problem of consciousness and drives people to unnecessary despair. However, there is a promising theory that on the one hand, it explains consciousness with a measurable, quantifiable property of the physical world, on the other hand, it does not reduce it to a classical physical property in the sense that we currently know. For this purpose, if we summarize what we have studied so far, in the first chapter we introduced the problem that will be discussed in this dissertation. In the second chapter I formulated the hard problem of consciousness, analysed the sui-generis character of consciousness resisting to the scientific explanation, and examined the required conditions that the theory must have to open the door to the scientific explanation of consciousness. In the third chapter I discussed the traditional solution-oriented approaches to the hard problem of consciousness under the two main categories: dualism and monism. After I analysed the dualistic and monistic view, their problems, and the objections to them, in the chapter four, I will discuss the mysterians, Thomas Nagel and Colin McGinn, who are pessimistic about understanding the subjective and qualitative character of consciousness. In this chapter I divided mysterianism into two main categories: weak version of mysterianism which is advocated by Nagel emphasises the possibility of

solving the mystery of consciousness in the future, the strong version which is defended by McGinn claims that the mystery is permanent due to the cognitive closedness of human being. Even though I mentioned Nagel's thesis and compared it with the thesis of McGinn, the essential purpose of the fourth chapter was discussion of the strong version of mysterianism and its problems. With this discussion, I aimed to show that McGinn's arguments are not satisfactory enough to leave consciousness to the dark side and mysterianism which advocated by him is pseudo-mysterianism. Not only did I show that McGinn's mysterianism is refutable, I went a step further and argued in chapter five that there might be a constructive and naturalistic solution to the hard problem of consciousness, and I supported this claim with Tononi's Integrated Information Theory (IIT), which is the most promising approach to the explanation of consciousness. The IIT provides us the most reasonable data and working method on the explanation of consciousness. Here in the remaining part of this chapter, I will support this idea by investigating its compliance with the required conditions given in the second chapter of this dissertation.

6.1. The required conditions that a theory must have to explain consciousness in the scientific manner: The IIT Theory as an Example

Let's remember the criteria that we have listed in the second chapter of this dissertation as the necessary conditions that should be in the thesis that will explain consciousness

Firstly, we claimed that "there must be a theory that cares about the phenomenological features of consciousness, such as qualitative and subjective nature, and should starts from there." As explained in the fifth chapter of this dissertation, the IIT does not start from the question of how the brain lead to consciousness; by identifying the characteristics of consciousness with the axioms the IIT starts from the essential features of phenomenal consciousness. And then it investigates the physical substratum which correspond to the special character of consciousness.

Secondly, we emphasised that such a theory "should be able to express the subjective character of consciousness in a scientific and mathematical language and translate them into these fields". As explained in detail in the previous chapter the IIT defines the existence of consciousness as the system's capacity to integrate information. This theory uses a mathematical framework to explain the quantity and quality of consciousness. It measures the existence of consciousness in a system with the system's level of integrated information that corresponds to Φmax and formulates the qualitative character of consciousness with the system's special way of integrating information.

Our third criterion assumes that "the entity and process in which the phenomenal consciousness is explained must be placed on a physical basis not in the sense that we currently know but in a way that expands our perception of space and physical world". As we mentioned in the fifth chapter Tononi uses not the neural network itself, but the information passing through this network to explain consciousness. Even though classical identity theories reduce consciousness to neuro-physical properties of the brain, the IIT does not apply to entities that are physical in the sense that we currently know. Information, which is the basis of the IIT, appears as a physically definable entity. The idea that information is physical, firstly is defended by Landauer in 1961, has been improved by recent studies. Important physicist Vopson advocates the idea that information can be the fifth state of matter that can be measured massively, and he develops this idea based on the relationship between mass energy and information. These studies give us hope that consciousness can be explained scientifically by a physical entity in a different sense from the one we know, and that unknown aspects of consciousness can be clarified as our knowledge of the physical world improves.

Another criterion asserts that "what is modelled and translated into the scientific mathematical language, must also have scientifically testable predictions. These tests may be difficult, not very likely, but they should at least be logically testable". The IIT predicts that for the existence of consciousness the complex of the brain must

achieve the maximum cause-effect power (Φmax). Even though it is thought that measuring this prediction is difficult, it is not impossible.

The last condition we determined for the theory explaining consciousness is related with the consistent implications of the theory. That is, “the implications of the theory should not be counterintuitive”. As explained in chapter five, the information integrating system has to be consistently conscious, no matter what it is. The system can be a carbon-based organic material or a synthetic material. This shows that the implications of the IIT is valid both for animals and infants, and it also opens the door to artificial intelligence as long as the conditions are met.

Integrated Information Theory may have some problems; for instance, finding a computable and certain quantity corresponding to consciousness can be considered very difficult and challenging. However, it is not among the purposes of this dissertation to claim that the IIT is the ultimate theory of the explanation of consciousness. However, it is clear that the IIT is an effort to explain the problem of consciousness holistically; it does not pass over the subjectivity (qualia) of experiences. In addition, it argues that consciousness should be considered in an inclusive way not only for humans, but also for non-human creatures and even non-living beings. In the final examination, it is clear that even if the IIT theory is proven wrong, it refutes the mysterians like McGinn who believe that consciousness cannot be explained and will contribute to the right path for the next generation of researchers to follow.

REFERENCES

Alexander, S. (1966). *Space Time and Deity*. London: Macmillan

Arici, M. (2018). The Problem of Phenomenal Consciousness. *MetaMind*, 1(1): 1-19

Block, N. (1978). Troubles with Functionalism, *Minnesota Studies in the Philosophy of Science*, 9: 261-325

Brueckner, A., Beroukhim E. A. (2003). McGinn on Consciousness and the Mind Body Problem. In Q. Smith, A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*. New York: Oxford University Press

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

Chalmers, D. J. (2002). *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press

Chalmers, D. J. (2003). Consciousness and Its Place in Nature. In Stephen P. Stich & Ted A. Warfield (eds.), *Blackwell Guide to the Philosophy of Mind*. Blackwell. pp. 102-142

Chomsky, N. (1976). Problems and Mysteries in the Study of Human Language. In Asa Kasher (ed.), *Language in Focus: Foundations, Methods and Systems*. Boston Studies in the Philosophy of Science, 43, pp: 281-357

Chomsky, N. (1998). *On Language*. London, New York: The New Press

Chomsky, N. (2013). What can we understand?. *The Journal of Philosophy*, 110: 663-684

Churchland, P. M. (1981). Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy*, 78 (2), 67-90

Churchland, P. M. (1985). Reduction, Qualia and the Direct Introspection of Brain States. *The Journal of Philosophy*, 82(1): 8-28

Churchland, P. S. (1986). *Neurophilosophy: toward a unified science of the mind-brain*. Cambridge, MA: MIT Press

Churchland, P. M. (1999). Folk Psychology. In *On the Contrary: Critical Essays*, (pp. 3–15). MIT Press.

Demircioglu, E. (2016). Against McGinn's Mysterianism. *Cilicia Journal of Philosophy*, 1: 1-10

Dennett, D. (1991), Review of McGinn's The Problem of Consciousness. The Times Literary Supplement

Dennett, D. (1995). *Darwin's dangerous idea*. London: Penguin Books

Descartes, R. (1984). Meditations on First Philosophy, In J. Cottingham, R. Stoothoff, D. Murdoch (eds.), *The Philosophical Writings of Descartes*, Cambridge University Press

Edelman, G., and Tononi G., (2000). *A Universe of Consciousness: How Matter Becomes Imagination.* New York: Basic Books.

Feser, E. (2005). *The Philosophy of Mind: A Short Introduction*. UK: Oneworld Publication.

Flanagan, O. J. (1991). *The Science of the Mind.* Cambridge: MIT Press

Flanagan, O. J. (1993). *Consciousness Reconsidered*. Cambridge: MIT Press

Flanagan, O. J. (2007). *The Really Hard Problem: Meaning in a Material World.* Cambridge: MIT Press

Garber, D. (2000). *Descartes Embodied: Reading Cartesian Philosophy through Cartesian Science*. Cambridge University Press

Garvey, J. (1997). What Does McGinn Think We Cannot Know?. *Analysis*, 57: 196-201

Hampshire, S. (1950). The Concept of Mind by Gilbert Ryle. *Mind*, 59: 237-255

Hanson, P. (1993). McGinn's cognitive closure. *Dialogue*, 32 (3), 579-585

Inwagen, P. V. (1996). [Review of the book: Problems in Philosophy: The Limits of Inquiry by Colin McGinn], The Philosophical Review, 105/2: 253-256

Işıkgil, S. (2017). An Analysis on McGinn's Mysterianism: Reply to Erhan Demircioglu. *Cilicia Journal of Philosophy*, 1: 11-18

Jacquette, D. (2009). *The Philosophy of Mind: The Metaphysics of Consciousness*. UK: Continuum Publishing

James, W. (1890). *The Principles of Psychology*. New York: Dover

Kim, J. (2011). *Philosophy of Mind*. Boulder, CO: Westview Press

Kirk, R. (1991). Why shouldn't we be able to solve the mind-body problem? *Analysis*, 51 (1), 17-23.

Koch, C. (2009). A 'complex' theory of consciousness. *Scientific American*, https://www.scientificamerican.com/article/a-theory-of-consciousness/

Koch, C., and Tononi, G., (2013). *Can a Photodiode Be Conscious?* New York Review of Books.

Kriegel, U. (2003). The New Mysterianism and The Thesis of Cognitive Closure. *Acta Analytica*, 18: 177-191

Kriegel, U. (2009). Mysterianism. In T. Bayne, A. Cleermans, & P. Wilken (eds.), *The Oxford Companion to Consciousness*. Oxford: Oxford University Press

Kruglanski, A. W. (1990). Motivations for judging and knowing: Implications for causal attribution. In The Handbook of Motivation and Cognition: Foundation of Social Behavior, 2nd ed. (Eds ET Higgins, RM Sorrentino), New York, Guilford Press.

Kukla, A. (1995). Mystery, Mind and Materialism. *Philosophical Psychology*, 8 (3), pp: 255-64.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* 5(3), 183–191

Landauer, R. (1996). The physical nature of information. *Phys. Lett. A.* 217(4-5), 188–193

Lee, G. (2007). Consciousness in a Space-Time World. *Philosophical Perspectives*, pp: 341-374

Lycan, W. (2006). Resisting ?ism. In A. Freeman (ed.), *Consciousness and Its Place in Nature*. Imprint Academic, pp. 65-71

McGinn, C. (1989). Can we solve the mind-body problem? *Mind*, 98 (391), 349-366.

McGinn, C. (1991). *The Problem of Consciousness: Essays toward a Resolution.* Oxford: Blackwell.

McGinn, C. (1993). *Problems in Philosophy: The Limits of Inquiry*. Oxford: Blackwell.

McGinn, C. (1994). The Problem of Philosophy. *Philosophical Studies*, 76 (2-3), pp: 133-156.

McGinn, C. (1995). Consciousness and Space. In T. Metzinger (ed.), *Consciouss Experience*, Paderborn: Ferdinand Schöningh.

McGinn, C. (1996). *The Character of Mind: An Introduction to the Philosophy of Mind*. Second edt. New York: Oxford University Press

McGinn, C. (1999). *The Mysterious Flame: Conscious Minds in a Material World.* New York: Basic Books

McGinn, C. (2003). What Constitutes the Mind-Body Problem? *Philosophical Issues*, 13, pp:148-162

McGinn, C. (2004). *Consciousness and its objects*. New York: Oxford University Press

Megill, J. (2005). Locke's Mysterianism: On the unsolvability of the mind-body problem. England: Centre for Philosophy, Lancaster University, pp: 119-147

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83 (4): 435-450

Nagel, T., (1979). Subjective and Objective. In *Mortal Questions*, Cambridge: Cambridge University Press. 196–213.

Nagel, T., (1980). The Limits of Objectivity. In *Tanner Lectures on Human Values*, 77–139.

Nagel, Thomas. 1986. *The View From NoWhere*. Oxford and New York: Oxford University Press.

Nagel, T., (1998). Conceiving the Impossible and the Mind-Body Problem. *Philosophy*, 73 (285): 337–52

Nida-Rümelin, M. (1996). Pseudonormal Vision: An Actual Case of Qualia Inversion?, *Philosophical Studies*, 82: 145-157

Perez, D. I. (2005). Mysteries and Scandals. Transcendental Naturalism and the Future of Philosophy. *Critica*, 37(110), 35-52

Place, U. T. (1956), Is Consciousness a Brain Process?, *British Journal of Psychology*, 47: 44-50

Place, U. T. (1998). Ryle's Behaviorism. In W. O'Donohue, R. Kitchener (eds.). *Handbook of Behaviorism*, pp. 361-398, Academic Press.

Putnam, H. (1967). Psychological Predicates. In W. H. Capitan, D. D. Merrill (eds.), *Art Mind and Religion*. Pittsburgh: University of Pittsburgh Press

Sacks, M. (1994). Cognitive closure and the limits of understanding. *Ratio*, 7 (1), pp: 26-42.

Seager, W. (1995). Consciousness, information, and panpsychism. *Journal of Consciousness Studies*, 2(3): 272–88.

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

Strawson, G. (2008). *Real Materialism and Other Essays*, Oxford: Clarendon Press

Tegmark, M. (2015). Consciousness as a State of Matter. Chaos, Solitons & Fractals.

Tononi, G., and Edelman, G. (1998). Consciousness and Complexity. *Science*, 282 (5395).

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5(42), pp. 1-22

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biology Bulletin*, 215: 216–242

Tononi, G. (2012). *The Integrated Information Theory of Consciousness: An Updated Account.* Arch Ital Biol.

Tononi, G., Boly, M., Massimini, M., Koch, C. (2016). Integrated Information Theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7), pp. 450-461

Tononi, G. (2017). The information integration theory of consciousness. In Max Velmans and Susan Schneider (eds.), *The Blackwell Companion to Consciousness.* Blackwell Publishing

Webster, D., Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. Journal of Personality and Social Psychology, 67: 1049-1062.

Webster D. M., Kruglanski A. W. (1998). Cognitive and social consequences of the motivation for closure. *European Review of Social Psychology*, 8:133-146.

Whiteley, C. H. (1990). McGinn on the Mind-Body Problem. *Mind*, 99: 289.

Worley, S. (2000). What Is Property P, Anyway? *Analysis*, 60, pp: 58-62.

Vlerick, M. Boundary, M. (2017). Psychological Closure Does Not Entail Cognitive Closure, *Dialectica*, 71: 101-11

Vopson, M. M. (2019). The mass-energy-information equivalence principle. *AIP Advances* 9, https://doi.org/10.1063/1.5123794

Vopson, M. M. (2020). The information catastrophe. *AIP Advances* 10, https://doi.org/10.1063/5.0019941

Vopson, M. M. (2021). Estimation of the information contained in the visible matter of the universe. *AIP Advances* 11, https://doi.org/10.1063/5.0064475

APPENDICES

A. CURRICULUM VITAE

Institutional affiliation: Middle East Technical University, Department of Philosophy

Email: @gmail.com

AREAS OF SPECIALIZATION

Philosophy of Mind, Epistemology, Philosophy of Science

AREAS OF COMPETENCE

Philosophy of Mind, Naturalism, New-Mysterianism

EDUCATION

Ph.D. in Philosophy, Middle East Technical University, Ankara, Turkey, 2016–2022.
Thesis Title: Pseudo-Mysterianism of McGinn: A Constructive and Naturalistic Solution of the Hard Problem of Consciousness is Possible, Supervisor: Assoc. Prof. Dr. Aziz Fevzi Zambak

M.A. in Philosophy, Middle East Technical University, Ankara, Turkey, 2013–2016.
Thesis Title: A Study on McGinn's Mysterianism, Supervisor: Assoc. Prof. Dr. Barış Parkan

B.A. in Philosophy, Middle East Technical University, Ankara, Turkey, 2008–2013

Minor in Psychology, Middle East Technical University, Ankara, Turkey, 2011-2013

PUBLICATIONS

Journal articles

Işıkgil, S. (2017). An Analysis on McGinn's Mysterianism: Reply to Erhan Demircioglu . *Cilicia Journal of Philosophy*, 0 (1), 11-18

Işıkgil, S. (2017). A Study on Plato's Attitude Towards Women. *FLSF*, 12(24), 393-404

CONFERENCES

"Plato's Attitude Towards Women", International Social and Humanities Berlin Conference, May 18 – 21, 2017. Berlin, Germany

TEACHING EXPERIENCE

Teaching assistant at the Middle East Technical University

1. Phil 241: Philosophical Texts I
2. Phil 282: History of Science
3. Phil 341: History of 17th & 18th Century Philosophy I
4. Phil 351: Applied Ethics
5. Phil 343: Philosophy of Plato
6. Phil 408: Philosophy of Mind II
7. Phil 108: Introduction to Philosophy
8. Phil 201: Ethics
9. Phil 390: Rationality Theories of Science
10. Phil 352: Metaphilosophical Studies II

B. TURKISH SUMMARY / TÜRKÇE ÖZET

Beyin üzerine yapılan araştırmaların son yıllarda artış göstermesiyle birlikte bilinç, felsefe, nörobilim ve bilişsel bilimler için ilgi çekici ve merak uyandıran bir araştırma konusu haline gelmeye başlamıştır. Beynin fiziksel aktivitelerinin incelenmesi sırasında kullanılan ileri teknoloji beyin görüntüleme yöntemlerinin bilincin öznel ve niteliksel yanını anlama konusunda insanlığa katkı sağlamaması bilinci giderek tanımlaması zor ve gizemli bir fenomen haline getirmiştir. Bilincin kaynağının fiziksel bir sistem olan beyin olduğunu bilmemize rağmen bilincin beyinden tamamen farklı özelliklere sahip olması kafa karıştırıcı olmaktadır ve aslında bilinç bulmacasının özü de budur. Deney, gözlem ve doğrulama gibi ampirik yöntemlere dayanan bilimsel araştırmaların yaşadığımız dünyanın sorunlarını çözdüğü oldukça açıktır. Ancak, özellikleri itibariyle bilinç gibi karmaşık bir şeyin nasıl olup da kafatasımızın içindeki gri bir et parçasından ortaya çıktığı sorunu diğer tüm bilimsel sorunlardan kökten farklıdır. Beyin, gözlemlenebilir özelliğinden dolayı bilimsel yöntemlerle bilimsel araştırmalara açıkken bilinç, bilimsel yöntemlere uygun varlıklardan farklı olan özellikleri nedeniyle bilimsel araştırmaya direniyor görünmektedir. Bu nedenle bilinç ile beyin arasındaki ilişki, bilincin zor problemi olarak kabul edilir.

Bilincin zor problemi veya daha geleneksel olarak zihin-beden sorunu sinirbilim, yapay zeka ve teknolojideki gelişmelere ve ortaya atılan tüm felsefi teorilere rağmen canlılığını ve önemini korumaya devam etmektedir. Bilinç sorunu, antik çağlardan günümüze kadar hem filozoflar hem de bilim adamları tarafından ele alınmış ve temel amaç, zihinsel durumların doğal, fiziksel dünyada nasıl gerçekleştiğini anlamak olmuştur. Başka bir deyişle, elektronlardan, atomlardan, elektromanyetik alanlardan ve çekim kuvvetlerinden oluşan bir dünyada, fizik yasalarına aykırı olmadan, acı ve zevk gibi zihinsel durumların nasıl ortaya çıktığı anlaşılmalıdır. Bu amaç doğrultusunda geliştirilmiş çözüm odaklı birçok geleneksel yaklaşım ortaya konulmuştur. Geleneksel çözümler iki ayrı kategoride incelenebilir: Bu kategorilerden ilki, bilincin öznel yanını koruyan ancak fiziksel temelde

açıklayamadıkları için bilinç-beyin etkileşimini anlamayan yaklaşımları temsil eder. Diğer kategori, bilinci fiziksel bir temelde açıklamaya çalışırken, bilincin öznel ya da niteliksel yönünü görmezden gelen ya da gereken önemi vermeyen yaklaşımları temsil eder. Bu iki uç düşünce biçimi, zor bilinç sorununu çözmekten acizdir ve bu yetersizlik, bilinç sorununu bir sarkaç gibi iki uç arasında gidip gelen sonsuz bir kısır döngüye sokar.

Öznel/subjektif/fenomenal bilinci bugün bildiğimiz anlamda temel fiziksel varlıklar açısından açıklamanın mümkün görünmediği fikrine ben de katılıyorum. Fenomenal bilinci klasik fiziğin maddesel varlıklarına indirgeyerek açıklamaya çalışan psiko-fiziksel teoriler, bilincin öznel ve niteliksel yönünü göz ardı ederek başarısız olmuştur. Bu başarısızlıklar, Colin McGinn gibi bazı karamsar filozofların bilincin bilimsel bir açıklamasının olabileceği hususunda umutsuzluğa kapılmasına sebep olmuş ve onu asla açıklanamayacak bir gizem olarak görmelerine yol açmıştır. McGinn'in tatmin edici olmayan mevcut teorilere bir alternatif olarak ürettiği tezi bilincin aslında doğal bir fenomen olduğunu iddia eder, ancak bilişsel kapasitelerimiz bunu açıklamak için yetersizdir. McGinn'e göre bu yetersizlik yüzünden elimizden gelenin en iyisini yapsak bile bilinç bizim için sonsuza kadar bir gizem olarak kalmaya mahkum olacaktır. Ancak bu çalışmanın amacı, McGinn'in gizemciliğinin temelsiz ve aşırı bir iddia olduğunu ortaya koymak ve bilincin öznel yanını bir başlangıç noktası alarak, bilinci fiziksel ve bilimsel bir temelde açıklamayı vaat eden doğalcı yaklaşımların olabileceğini öne sürmektir.

Colin McGinn, bilincin hiçbir zaman insanlar tarafından açıklanamayacağını iddia eden bir yaklaşımın öncüsüdür ve bu yaklaşım bu çalışmada "güçlü gizemcilik" olarak adlandırılmaktadır. Gizemciliğin zayıf versiyonu olarak adlandırdığım ve Thomas Nagel tarafından temsil edilen yaklaşım bilincin gelecekte bilimsel bir açıklamasının olabileceği fikrine kapıları kapatmazken bu yaklaşımdan farklı olarak güçlü gizemcilikte bilinç sorununun çözülemezliği kalıcıdır, bu da insanoğlunun bilişsel kapasitesindeki yetersizliğin bir sonucudur. 1989 tarihli makalesinde McGinn, bilinç problemine dair yaklaşımını "benim tercih ettiğim yaklaşım

doğalcıdır ama yapıcı değildir" diyerek tanımlamaktadır. Yapıcı olmayan doğalcılık olarak tasnif ettiği tezini üç ana argümana dayandırmaktadır:

i. Beynin bilinci doğal olarak açıklayan bir niteliği vardır. (Bu argüman bu çalışmada "Doğalcılık Argümanı" olarak adlandırılmıştır)
ii. Biz insan olarak beynin bu niteliğine bilişsel olarak kapalıyız (Bu argüman "Kapalılık Argümanı" olarak adlandırılmıştır)
iii. Felsefi bir zihin-beden problemi yoktur (Bu argüman "Çözümleme Argümanı" olarak adlandırılmıştır)

McGinn'e göre, bilincin kendisi doğal olarak beyinden kaynaklanan biyolojik bir fenomendir, bu nedenle beyin ve bilinç arasındaki ilişki, karaciğer ve safra arasındaki ilişki kadar doğaldır. Beyinde, zihinsel ve fiziksel arasındaki bağlantıdan sorumlu olan *P* gibi bir nitelik vardır. Ancak *P* beynin niteliklerinden biri olmasına rağmen beyin bilince sahip olduğu şekilde bu niteliğe sahip olduğu için *P* de tıpkı bilincin kendisi gibi beynin fiziksel niteliklerinden farklı özelliklere sahiptir. Beynimiz uzamsal bir yapı olmasına rağmen bilinç öyle değildir ve *P* de beynin uzamsal bir niteliği olmamalıdır çünkü şimdiye kadar beyindeki uzamsal hiçbir şey bizi aradığımız bilinç-beyin bağlantısını bulduğumuza ikna etmemiştir.

McGinn, beyin ve bilinç arasındaki bağa ilişkin doğalcı bir tutum takınsa bile, bilinç sorununun çözümüne yönelik yaklaşımının yapıcı olmadığını iddia eder (s. 350). Yani beyin ile bilinç arasındaki bağlantıyı mümkün kılan beynin doğal niteliğini hiçbir zaman açıklayamayacağımıza inanır ve bu iddiasını "bilişsel kapalılık" teziyle desteklemeye çalışır. Bilişsel kapalılık fikrine göre bir organizmanın kavram oluşturma kapasitesi bir niteliği veya bir teoriyi anlamak için yeterli değilse, organizma bilişsel olarak o niteliğe veya teoriye kapalıdır. Her türün kendi bilişsel sınırlamaları vardır, bu nedenle bazı türler için bir teori veya nitelik elde edilebilir ancak bilişsel olarak diğer türlere kapalı olabilir. Örneğin, "bir farenin zihnine kapalı olan bir maymunun zihnine açık olabilir ve bize açık olan maymuna kapalı olabilir" (1989, s. 350). Bir elektronu ele alırsak, insan zihni bir elektronun özelliğini anlamak

için yeterli bilişsel kapasiteye sahipken, maymun zihinleri onun ne olduğunu asla anlayamaz; bilişsel olarak bu özelliğe kapalıdırlar.

Fakat McGinn *P* niteliğini hiçbir zaman kavrayamayacağımız yönündeki fikrini savunurken yalnızca benzerlik argümanına bağlı kalmaz. Ona göre, *P* niteliğini anlamak için bizi çözüme götürecek iki olası yol vardır: birincisi iç gözlem, yani bilincimizi doğrudan inceleyerek, ikincisi ise beynimize dair araştırmalar yaparak (1989, s. 354). Yani tüm dünyayı ya iç gözlem temelli ya da algı temelli yetiler aracılığıyla anlıyoruz. Bu durumda ya bilinç durumlarımız hakkında doğrudan bilgi sahibi olarak ya da beyin üzerinde gözlemler yaparak *P*'yi anlamaya çalışabiliriz, ancak bunların hiçbiri *P*'yi anlamamız konusunda bize yardımcı olmaz.

İç gözlem yoluyla bilincin özelliklerinin farkına varırız; zihinsel durumlarımıza anında erişebilir ve böylece içsel hayatımızdan kavramlar oluşturabiliriz. Fakat maalesef bu, *P*'ye ulaşmak için yeterli bilişsel yeti değildir, çünkü *P*, beyin ve bilinç arasında aracı bir işleve sahiptir, ancak iç gözlem, bilinçli durumları beyne bağlı olduğu şekliyle bize sunmaz; bilinç-beyin bağlantısının sadece bir yanını verir (1989, s. 354). Bir başka nokta da şu ki, *P* kavramını bilinç kavramları üzerinden analiz ederek elde edemeyiz çünkü ancak kendi iç gözlemimizle ulaşabildiğimiz bilinç durumu için kavram oluşturabiliriz. Örneğin, yarasanın ekolokasyon durumu için bir kavram oluşturamayız; bu durum yarasanın kendi öznel deneyimidir. Benzer şekilde, gözleri görmeyen bir kişi daha önce hiç görmediği bir 'kırmızı' deneyimi kavramını kavrayamaz. Aynı şekilde, *P*'nin bir bilinç sorununun çözümü olarak kavramsallaştırılması, içebakış yönteminin ötesindedir.

P'ye ulaşmaya çalışabileceğimiz bir diğer yol da beynimizin incelenmesidir. Ancak McGinn'e göre bu yolu kullanarak da iç gözlemden daha iyi bir sonuç elde edemeyiz çünkü bu yetinin işlevi beyni kavrayışımızı oluşturmaktır. Beynin incelenmesi için duyularımızı kullanmalıyız, ancak duyular bize yalnızca uzamsal nesneleri sunar. Duyular uzamsal özelliklere yanıt vermede etkilidir, ancak uzamsal hiçbir şey bilinç sorununu çözemeyeceğinden, *P* bilincin kendisi gibi uzamsal olmayan bir nitelik

olmalıdır. Bu nedenle McGinn'e göre P'nin bilinç-beyin bağını nasıl sağladığını anlayamayız ve uzamsal özelliklerin araştırılmasıyla bilinci açıklayamayız.

Bir şeye algısal olarak kapalı olmak, ona bilişsel olarak kapalı olmayı gerektirmez, çünkü kavram oluşturma yöntemlerimizle *P* için teorik bir kavram oluşturmaya yönelik umudumuz hala vardır. Ancak McGinn, bu yöntemin de *P*'yi bize verme konusunda yararsız olduğunu iddia ediyor. Gözlem verilerinden bir çıkarım yaparak teorik bir *P* kavramı elde edemeyiz çünkü gözlemlerimize dayalı teorik kavramlar üretirken "homojenlik ilkesini" kullanıyoruz. Homojenlik ilkesi, fiziksel verilerden çıkarım yaparak formüle ettiğimiz teorik kavramın, gözlemimizin öznesi ile yeterli derecede homojen olması gerektiğini beyan eder. McGinn'e göre, bu yöntem gözlemlenemeyen maddi varlıklara dair kavram oluşturmamız aşamasında çalışır. Çünkü bu varlıklar gözlem nesneleri olmasalar da gözlemlediğimiz fiziksel verilerin analojik uzantıları olmaları vasıtasıyla onlar hakkında teorik kavramlar üretebiliyoruz. Fakat *P*, beynin gözlemlenebilir varlıklarının analojik bir uzantısı olarak verilemez çünkü ne bilincin kendisi ne de *P*, beynin uzamsal nitelikleriyle aynı türden varlıklar değildir. Bu durumda homojenlik ilkesine göre *P* sadece algısal olarak değil, bilişsel olarak da bize kapalıdır.

Sonuç olarak, McGinn'e göre, kavram oluşturmamıza yarayan bu iki farklı yeti, içebakış ve beynin incelenmesi, bize beyin ve bilinç arasındaki ilişkinin kısmi bir resmini sunar; bu nedenle beyni ve bilinci birleştiren temel özelliği ortaya çıkarmakta başarısız olurlar. Beyin-bilinç bağlantısı hakkındaki kısmi bilgimiz, kavram oluşturma yetilerimizdeki içsel sınırlamaların bir sonucudur ve ne yazık ki bunları değiştirmenin veya genişletmenin bir yolu yoktur. Bu nedenle, McGinn'e göre beyin-bilinç problemi bizim için bir gizem olarak kalmaya mahkumdur.

Bahsedildiği üzere McGinn, bilinç-beyin sorununa yapıcı bir çözüm bulma konusunda karamsar bir konuma sahiptir. Ancak felsefi kafa karışıklığını ortadan kaldırmak için beyin ve bilinç arasındaki bağı anlamamız gerekmediğine inanmaktadır. Doğuştan gelen bilişsel sınırlarımız nedeniyle bilinç probleminin doğalcı çözümüne bilimsel olarak asla ulaşamayacağımızı bilmenin felsefi bilinç

problemini de ortadan kaldırdığını iddia etmektedir. Bilinç, "sindirim" veya "cinsel üreme" gibi süreçlerden daha karmaşık değildir ve beynin bilinç salımı, karaciğerin safra salgılaması kadar doğaldır. Beyin ve bilinç arasındaki ilişki hakkındaki felsefi kafa karışıklığı, bu bağın mucizevi bir şekilde ortaya çıktığı hissine dayanmaktadır. Ancak McGinn'e göre bu durum mucizevi bir olayın sonucu değil, bilişsel kapasitemizin yetersizliğinin sonucudur. Yani gizem ontolojik değil, bilinci açıklayan bilimin bilişsel sınırlamalarımızın ötesinde olması anlamında epistemolojiktir ve bu sınırlamaların farkında olmak felsefi bilinç-beyin problemini ortadan kaldırmak için yeterlidir.

McGinn, kendi hipoteziyle bilinç-beyin sorununun yapıcı olmayan ama doğalcı çözümünü kanıtlamaya çalışır. McGinn'in bilinç sorununa doğalcı bir çözüm olduğu konusundaki ısrarını mantıklı bulsam bile, bu çözümün bizim için numenal olduğu tezini kabul edilemez bulmaktayım. Beyin ve bilinç arasındaki bağlantının sınırlı bilişsel kapasitemiz nedeniyle bizim için sonsuza dek dünyanın karanlık tarafında kalacak olduğu iddiası sindirilebilir bir iddia değildir. McGinn, gizemciliğini bilişsel kapalılık teziyle inşa eder ancak bilişsel kapalılık tezini haklı çıkarmak için kullandığı ana argümanların sonuçsuz olduğunu göstermek McGinn'in gizemciliğinin de yanıltıcı olduğunu ortaya koyacaktır.

McGinn'in tanımladığı *P* niteliğinin uzamsal olmaması McGinn'in doğalcı konumuyla ilgili bir çelişki yaratmasa da bu durum *P*'nin karakterizasyonunda bazı problemler olduğu gerçeğini değiştirmemektedir. Tamamen farklı özelliklere sahip iki farklı nitelik arasında aracılık yapabilmek için her biri ile yeterince homojen olan bir niteliğin olması gerekir ve McGinn, beynin uzamsal özelliklerinin, bu özelliği sağlamaya yeterli olamayacağında ısrar eder. Ancak, McGinn'in tanımladığı *P* niteliği de aracılık rolü sağlama açısından yetersiz görünmektedir. McGinn, *P*'nin rolü beyin ile bilinç arasında aracılık etmek olduğu için *P* hem beynin hem de bilincin bir özelliği olmalıdır dese de *P*'yi bilinçle tıpatıp aynı özeliklere sahip bir nitelik olarak karakterize eder. Bu durumda, bilinçle aynı özelliklere sahip bir özelliğin beyinle nasıl ilişkilendirilebileceği sorusu, bilincin beyinden nasıl ortaya çıktığı sorusu kadar kafa karıştırıcıdır.

Aracı bir *P* için gerekli temel koşul, bilinç ile *P* ve *P* ile beyin arasında paylaşılan en az bir ortak özellik olması durumudur, ancak McGinn'in tanımladığı *P* niteliğinin beynin özellikleriyle hiçbir ortak yanı yok gibi görünmektedir. Bu durumda neden bilinçle aynı özelliklere sahip bir niteliğe ihtiyacımız var ve bu gibi bir nitelik, karşılaştığımız sorunların sayısını artırmaktan başka ne işimize yarayabilir gibi yeni sorular ortaya çıkmaktadır. Görünüşe göre McGinn'in *P*'si yalnızca aracılık rolünde başarısız olmakla kalmıyor, aynı zamanda bilinç-beyin probleminin yanı sıra *P*-beyin problemi ve *P*-bilinç problemi olarak iki farklı problemi daha gündeme getiriyor. Şimdiye kadar bilinç ile beyin arasında nasıl bir bağlantı olduğunu arıyorduk ancak McGinn'in teziyle birlikte *P*-beyin ile *P*-bilinç arasında nasıl bir bağlantı olduğu problemleri de anlaşılmayı bekleyen durumlar arasında yerini almıştır.

McGinn, gizemciliğinin özü olan bilişsel kapalılık tezini çift aşamalı bir yöntemle kurar. Öncelikle bilişsel kapalılığın olabilirliğini göstermeye çalışır ve ardından iki farklı kavram oluşturma yetisine başvurarak bilişsel kapalılığın gerçekliğini iddia eder. Ancak olabilirlikten gerçekliğe uzanan bilişsel kapalılık tezinin birtakım problemleri vardır.

Her biyolojik türün kendi bilişsel sınırlamaları vardır ve bir türde bilişsel olarak kapalı olan bir problem diğerine açık olabilir ve McGinn, insan zihninin dünyadaki her sorunun çözümünü anlayacak kadar güçlü olduğunun garantisi olmadığını iddia ediyor. Yukarıda örneklendirildiği gibi bu iddia maymunların elektron kavramını anlayamaması veya farelerin trigonometrinin çözümlerini anlayamaması gibi, insanların da beyin-bilinç probleminin çözümünü mümkün kılan nitelik gibi evrenin bazı özelliklerini anlayamadıkları anlamına gelir.

Biyolojik bir sistem olarak insan zihninin evrendeki bazı şeylere bilişsel olarak kapalı olma ihtimalini ben de göz ardı etmiyorum. Evrenin farkında olmadığımız, ne hakkında olduğunu bilmediğimiz, bunlarla ilgili soru soracak kavramsal bilgiye sahip olmadığımız çok ilginç ve önemli özellikleri olabileceğini düşünüyorum. Ancak hakkında düşünebileceğimiz, sorular formüle edebileceğimiz ve bazı yönleriyle ilgili araştırma yapabileceğimiz bir bilinç bu özellikler arasında görünmemektedir.

Hayvanlar ve insanlar arasında benzetmeden yola çıkarak bilinç sorununun çözümüne ilişkin bilişsel kapalılığın insanlık için mümkün olduğunu tartışmak pek de makul görünmemektedir çünkü hayvanların evrenin bazı gerçeklerine bilişsel olarak kapalı olması durumu ile insanların bilinç probleminin çözümüne dair bilişsel kapalılığı arasında temel bir fark vardır. İnsanlar bilinç problemlerini formüle edip anlayabilirken, hayvanlar bilişsel olarak kapalı oldukları alanlarla ilgili problemleri formüle edemez ve anlayamaz. Yani hayvanlardan farklı olarak insan, bilmediğinin farkındadır. Maymunların elektron kavramına kapalı olmaları, farelerin trigonometri problemlerini çözememeleri veya köpeklerin politik problemler üzerinde düşünememeleri şaşırtıcı değildir, çünkü bu vakaların hiçbiri onların dünyasında problem olarak yoktur.

Dennett (1991, 1995) ve Kriegel (2003, 2009) da hayvanlar ve insanlar arasındaki bu ayrıma dikkat çekmektedir. Onlara göre McGinn, insan zihni ile hayvan zihni arasındaki en önemli fark olan dilsel yeteneğimizi göz ardı etmektedir. Kriegel, "bir soruyu anlamak ile onun olası yanıtlarını anlamak arasında kavramsal bağlantılar olduğunu" vurgular (2003, s. 184). Yani, bir problemin kendisini anlamak için doğuştan gelen sınırlamalarımız yoksa, bu durumda bu problemin olası çözümlerine de bilişsel olarak açık olmak durumundayız. Bir problemi anlıyor olmak onun olası çözümleri arasında hangi çözümün doğru olduğunun bilgisine doğrudan ulaşmayı zorunlu kılmaz. Yalnızca çözümünün ne olacağını anlayabilmeyi, doğru araştırma ve gelişmelerle nihayetinde doğru çözüme ulaşabileceğimiz gösterir. Problemi kavrayacak kadar güçlü olan bilişsel kapasitemiz ile problemin çözümünün bizim için anlaşılamaz olduğunu varsaymak pek mantıklı görünmemektedir. Bütün bunlar göz önüne alındığında, insanın bilinç probleminin çözümüne bilişsel olarak kapalı olma durumunu insan ve hayvan arasında kurulan bilişsel bir analojiyle açıklamak oldukça zayıf görünmektedir.

Bunun önüne geçmek için McGinn, insanlığın bilinç sorununun çözümünü anlamak için gerekli bilişsel yetiye sahip olmadığını iddia ederek bilişsel kapalılık ihtimalini insanlık için güçlü bir gerçekliğe dönüştürmeye çalışmıştır. Yukarıda ayrıntılı olarak açıklandığı gibi, kavramlarımızı oluştururken bağlı olduğumuzu iddia ettiği ve *P* ile

ilgili kavram oluşturamayacağımız hususunda ileri sürdüğü temel argüman homojenlik ilkesidir. McGinn, homojenlik ilkesini kavram oluşumuna ilişkin bir tür koruma prensibi olarak kullanır. Fakat ne yazık ki bu ilke kavramsal çıkarımların alanını yalnızca en basit ve en dolaysız bağlantı türleriyle sınırlayarak çalışır. Yani beyni gözlemleyerek ancak uzamsal beyin durumları arasındaki bağlantıları bulabilir ve bu verilerle ilgili teorik kavramları ortaya koyabiliriz ve iç gözlem yoluyla ancak kendi bilinçli durumlarımız hakkında yorum yapabilir ve bunlarla ilgili teorik kavramları ortaya koyabiliriz.

Bu kuşkusuz doğru gibi görünmektedir, ancak gözlemsel verilerimizle ilgili teorik kavramları ortaya koyarken davranışı üçüncü şahıs yoluyla gözlemlemek ve bilince ilişkin akıl yürütme yoluyla doğrudan olmasa bile dolaylı olarak bilinci araştırmak gibi üçüncü bir yola başvuruyoruz. İçe bakış, bilinç kavramımızı şekillendirmede önemli bir role sahip olsa bile, bilincin kendisini yalnızca iç gözlem yoluyla ortaya koymadığı açıktır; örneğin diğer canlılarda bilincin varlığı durumunu davranışlarındaki tezahürü ile inceleyebiliyoruz. McGinn, gözlenen fiziksel verileri açıklamak ve bunlara dair kavram üretmek için yalnızca bu verilerle ilgili teorik özelliklere ihtiyacımız olduğunu iddia etse de (1989, s. 359), bazı gözlemlenebilir verilerde hem bilinç durumlarını hem de beyin durumlarını eş zamanlı inceleyerek kavram üretiyoruz. Flanagan bunu kör görüş olgusuyla açıklar. Kör görme hastalarının birincil görsel korteksinde (V1) lezyonları vardır, ancak görsel uyaranlara bilinçli olarak göremedikleri halde istatistiksel olarak anlamlı tepkiler verebilirler. Yani bu hastalara nesnelerle ilgili sorular sorulduğunda baktıkları yöndeki nesnelerle ilgili soruları doğru yanıtlamaktadırlar. Ayrıca, hiçbir şey görmediklerini söylemelerine rağmen, bazı eylemleri doğru bir şekilde gerçekleştirebilir ve baktıkları yöndeki nesnelerin özelliklerini tanımlayabilirler. Bu durumda, McGinn'in iddiasının aksine, gözlemlenen verileri açıklamak için hem bilincin niteliklerine (veya eksikliğine) hem de görsel korteksteki lezyonu ortaya çıkarmak için beynin niteliklerine başvurmamız gerekir (1993, s. 114). Görme bilinci açısından tam gören kişiler ya da görme engelli hastalar ile görme bilincinin eksik olduğu kişiler arasındaki fark salt fiziksel verilerle açıklanamaz. Bu tür durumlar için

teorik açıklamalar hem sinirsel özelliklere hem de bilinçli özelliklere aynı anda başvurmayı garanti eder.

Bunun yanı sıra insanın belki de hiçbir zaman algılamadığı varlıklar hakkında düşünebilen ve kavramlar sunabilen bir varlık olduğu gerçeği de göz ardı edilmemelidir. McGinn, teorilerimiz oluştururken algımızın nesneleri olmayan özelliklere atıfta bulunduğumuz gerçeğini de kabul ediyor. Ancak yine de homojenlik ilkesi prosedürüne uyan varlıklar hakkında teorik kavramlar oluşturabileceğimiz konusunda ısrar ediyor. Gözlemlediğimiz fiziksel verilerden analojik bir uzantı yaparak bu tür varlıklar hakkında teorik kavramlar oluşturduğumuzu, ancak *P* gözlemleyebileceğimiz uzamsal bir nitelik olmadığı için *P* hakkında herhangi bir teorik kavram oluşturamadığımızı iddia ediyor. McGinn'in homojenlik ilkesi aşırı derecede kısıtlayıcı görünmektedir. Gözlemlerimizden bir çıkarım yaparak gözlemlenemeyen nesneler hakkında kavramlar oluşturuyoruz, ancak çıkarım yöntemimizin ve kavram oluşturma yöntemimizin yalnızca gözlemlediğimiz fiziksel nesnelerle analojiye dayandığını söylemek bilimi karanlık çağlara geri götürmekten başka bir işe yaramayacaktır. McGinn'in homojenlik ilkesi, göreceli olarak makro düzeyde gözlemlenemeyen nesnelerin teorik kavramlarının oluşumu için başarılı olabilir ancak enerji gibi gözlemlediğimiz uzamsal varlıklarla herhangi bir analojiye dayanmadan oluşturduğumuz mikroyapısal düzeyde kuantum fiziği alanından teorik kavramlar da vardır. Kavramları yalnızca McGinn'in iddia ettiği gibi analojik uzantı yoluyla üretiyor olsaydık, bugün sahip olduğumuz birçok bilimsel kavrama kapalı olurduk. Bu nedenle, McGinn'in bilişsel kapanış tezinin temeli olan homojenlik ilkesi oldukça yersiz görünmektedir. McGinn'in bilişsel kapasitemizin sınırlarına ve homojenlik ilkesine dayandırdığı kapalılık tezi yalnızca sonuçsuz bir varsayımdır; bilinç sorununun bilişsel gücümüzün dışında olduğunu kabul etmemiz için bize hiçbir kanıt ya da makul bir neden sunmaz.

McGinn'in bilişsel kapalılık tezini ısrarla savunmasının ve altında başka bir neden olmalıdır. McGinn'in kabul edilebilir herhangi bir veri veya bilimsel kanıt olmadan kurguladığı kapalılık tezinin, insan zihninin yetersizliğinden kaynaklanan bilişsel bir kapalılık değil, McGinn'in problemi sonlandırma arzusunun sonucu olan bir

'psikolojik kapalılık' olduğunu düşünmekteyim. Bir kişi bir sorunla karşılaştığında ve bilişsel sistemi içinde bu konuyla ilgili bir sonuca varamadığında rahatsız hissedebilir ve bu konuda bilgi toplama, araştırmasını durdurma ve sonuçlandırma isteği duyabilir. Yani psikolojik kapalılıktan kastettiğim, bireyin belirsizlik veren konuyu kapatma ve sonuçlandırma ihtiyacı ve isteğidir.

Öyle görünüyor ki, McGinn, bilinç sorunuyla ilgili olarak konuyu kapatma ihtiyacı tüm insanlığın bilişsel yetersizliği üzerine ortaya atılan bir hipoteze dönüşmüştür. McGinn, 1989 ve 1993 tarihli çalışmalarında bilinç-beyin sorununu, oldukça mucizevi ve ürkütücü olan, beyin sıvısının bilinç şarabına dönüştüğü hissi veren, yoğun bir kafa karışıklığı bırakan ve oldukça şaşırtan bir kargaşa olarak tanımlamaktadır. Bu tanımlamalardan da çıkarsanabileceği üzere McGinn'in asıl amacı bilinç sorununun yarattığı yoğun kafa karışıklığını ortadan kaldırmak ve bu sorunla ilgili bir an önce kesin bir sonuca varabilmektir. Bu amaçla, sorunun olası çözümlerine katkıda bulunmak yerine, bilinç hakkında düşünmeyi bırakmayı tercih etmiş ve kısa yol olan bilişsel kapalılık tezini öne sürmüştür. Bu tartışmaların ışığında, McGinn'in herhangi bir delile dayanmayan bilişsel kapalılık tezinin, onun hızlı bir sonuca varma ihtiyacının bir sonucu olan yanıltıcı bir tez olduğunu ve kapalılık tezi üzerine inşa ettiği epistemik gizemciliğinin de sözde-gizemcilik olduğunu düşünüyorum.

Başından beri bu tezin amacı, McGinn'in epistemolojik gizemciliğinin aşırı olduğunu ve fenomenal bilinci fiziksel temelde açıklayabilecek umut verici bir yaklaşımın olduğunu ortaya koymaktır. Bütünleşik Bilgi Teorisi (BBT) olarak adlandırılan bu yaklaşım, bilincin öznel, doğrudan ve dolaysız olduğuna dair sezgisel verileri korumakla kalmaz, aynı zamanda bilinci bir sistemin bilgi entegre etme kapasitesi ile özdeşleştirerek fiziksel bir temelde açıklanabileceğini göstermeyi amaçlar.

BBT modelinin bilincin öznel karakterini fiziksel sistemde nasıl açıkladığını tartışmaya geçmeden önce bilincin dünyanın uzay-zaman birlikteliği içindeki yerini kısaca analiz etmek bu çalışma için faydalı olacaktır. Einstein'ın Görelilik Teorisi bize zaman ve uzamın birbirinden ayrılamayacağını söyler. Başka bir deyişle,

görelilikte uzay ve zaman arasında yeterince net bir ayrım yoktur Sadece zamanda olamadığınız gibi sadece uzamsal olamazsınız, çünkü bu görüşte uzay-zaman birlikteliği sadece yapay olarak ayrı boyutlara bölünebilir ve bu yüzden ya tamamen onun içindesiniz ve dolayısıyla uzay ve zamandasınız ya da hiç içinde değilsiniz. (Lee, 2007, s. 341)

Klasik Kartezyen teori, zihnin zamanda var olduğunu kabul etse de uzamsal olarak varlığını inkar eder. Ancak zaman ve uzam birbirine göre iç içe geçmiş bir süreklilik gibidir ve bu açıdan bilinçli durumlar uzam-zamansal olmalıdır. Zaman ve uzay arasında net bir ayrım yoktur ve bu nedenle bir şeyin uzayda değil, yalnızca zamanda var olması veya bunun tersinin olması makul değildir. Görelilik bize bilincin uzamsal-zamansal olması gerektiğini söylese de bize onun uzay-zamana nasıl entegre edilebileceğini söylemez. Bilincin uzamsal olarak var olduğunu kabul etmek, geleneksel anlamda fizikalizmin doğru olduğunu kabul etmek anlamına gelmez.

Nagel'e göre, mevcut nesnel kavramlarımız bilincin öznel yönünü içermediğinden ne öznel ne nesnel olarak tanımlayabileceğimiz yeni bir kavram türünün oluşumu ile zihinsel ve fiziksel arasındaki uçurumu kapatabilir ve fiziksel dünyada bilince bir yer bulabiliriz. Ancak bunu bir adım daha ileri götürüp şu şekilde söylemek daha doğru olacaktır: Gelişen teknoloji, yapılan araştırmalar ve bilimdeki ilerlemeler ışığında fiziksel dünya anlayışımızı ya da fiziksel varlık denilen şeyden ne anladığımızı revize edersek fenomenal bilinç için sorunsuz bir fiziksel ve bilimsel açıklama üretebilir duruma gelebiliriz. Aslında Nagel'in bilincin öznel karakterini içerecek bir kavramsal değişiklik önerisi, fiziksel ve uzamsal dediğimizde anladığımız şeyin değişimine de bağlı görünmektedir.

McGinn gibi bilinç sorununun sonsuza kadar çözümsüz kalacağını düşünmek yerine ‘fiziksel’ kavramının ne olduğunu düşünmek, ‘fiziksel dünya’ ve ‘fiziksel olan’ hakkında sahip olduğumuz bilginin yeterli olup olmadığı sorgulamasının peşine düşmek daha makul olacaktır. Bizim anladığımız anlamının dışında, mevcut bilgilerimizin ötesinde “gerçekte var olan” fiziksel dünyanın sınırlarının ne olduğu bu tezin konusu değildir ve fizikçiler ve bilim adamları için bir araştırma konusu

olarak şimdilik bir tarafa bırakılmalıdır. Ancak bizim için önemli olan şudur ki, fiziksel dünyaya ve onun nesnelerine dair bilgimiz, yeni araştırmalar, gelişen teknoloji ve bilim dünyasındaki ilerlemeler ile geliştikçe ve değiştikçe bilinç fiziksel dünyanın bir nesnesi olarak bilimsel olarak açıklanabilir hale gelecektir. Bu nedenle, varoluşun tek bir tür temel kategorisi olduğu kabul edilse bile, eğer bu kategorinin şu anda bildiğimiz anlamıyla fiziksel olmanın ötesinde, gerçek manada fiziksel olduğu ortaya çıkarıldığında bilincin açıklanması bizim için hiç de gizemli olmayacaktır.

McGinn, asılsız bir iddiayla bilinç probleminin çözümüne dair tüm kapıları kapatıp insanları gereksiz umutsuzluğa sürükler. Bu çalışma sadece McGinn'in gizemciliğinin çürütülebilir olduğunu göstermekle kalmayıp bir adım daha ileri giderek bilincin zor probleminin yapıcı ve doğalcı bir çözümünün olabileceğini savunur ve bu iddiayı bilincin açıklanmasında umut vaat eden bir yaklaşım olan Giulio Tononi'nin Bütünleşik Bilgi Teorisi ile destekler. Bilincin açıklanmasına ilişkin başarılı bir teori, öncelikle bilincin doğasına dair bir analiz ortaya koymayı, ardından da bu doğayı görmezden gelmek yerine onu koruyan fiziksel temellerin araştırılmasını hedeflemelidir. Tononi tarafından 2004 yılında bilinci açıklamak için geliştirilen Bütünleşik Bilgi Teorisi, böyle bir teorinin destekleyici bir örneğidir. Bu teori, fiziksel temelden yola çıkmak yerine, bilinç hakkında bildiklerimize dayalı bir yaklaşım geliştirmeye çalışır. Tononi, bilincin içsel varlığını kabul eder; herkes kendi bilincini deneyimleyebilir, bilinç kesinlikle özneldir ve bölünmez bir bütündür. Tononi bu gerçeklerden yola çıkar ve insan bilincinin salt nöron ateşlemesinden daha fazlası olduğunu düşündüğünü belirtir. Esas olan nöronun kendisi değil, bu ağlardan geçen bilgidir.

BBT için iki önemli kavram enformasyon (bilgi) ve entegrasyondur (bütünleştirme). Tononi, üretilen bilginin önemini öncelikle “fotodiyot düşünce deneyi” ile açıklamaya koyulur. Bu deneyde hem fotodiyota hem de kişiye boş bir ekran gösterilir ve ekran açıkken “aydınlık”, kapalıyken “karanlık” demeleri istenir. Kişi ekranın açık ve kapalı olduğu durumları bilinçli olarak ayırt edebilirken, fotodiyot bu ayrımı ışığa duyarlı bir sensör ve bu sensöre bağlı olarak ışığı karanlıktan ayıran bir dedektör ile yapar. Bu düşünce deneyine göre kişi ile fotodiyot arasındaki temel fark,

ayrım yapılırken üretilen bilgi miktarıyla ilgilidir. Başka bir deyişle, bir fotodiyot ve bir kişinin ekrandaki "aydınlık" veya "karanlık" durumlarını algıladığında ürettikleri bilgi aralarında temel bir fark oluşturur.

Klasik bilgi teorisi kavramlarına göre her iki durumda da sonuçlar aynı ise fotodiyot ve öznenin ürettiği bilgi aynı yani bir bit bilgi olacaktır. Ancak BBT modeline göre bilgi miktarını ölçmek için kullanılan entropi hesaplaması için farklı bir yol izlenmektedir. Fotodiyot iki seçenekten birini seçer: "aydınlık" veya "karanlık". "Aydınlık" ve "karanlık" durumları eşit derecede olası olduğunda, fotodiyotun logaritma tabanına göre ortalama bilgi entropisi, bir bit bilgiye karşılık gelir. Ancak bir kişi ekrana baktığında sadece bir fotodiyot gibi ışığı algılamakla kalmaz, farklı birçok deneyime de sahip olur. Bir fotodiyottan farklı olarak bir kişinin 'aydınlık' deneyimini yaşamasının nedeni, ekranın aydınlık olması durumunun bir kişinin deneyimleyebileceği sayısız alternatiften sadece biri olmasıdır. Örneğin, ekran sadece "aydınlık" veya "karanlık" değil, aynı zamanda yeşil veya kırmızı da olabilir; ekranda şu ya da bu film, çiçek, dağ ya da manzara gibi birçok farklı kare olabilir. Aslında ekranda görülebilen karelere karşılık gelen olağanüstü sayıda olasılık vardır. Kişi sadece "aydınlık" ve "karanlık" arasında seçim yapmakla kalmayıp aynı zamanda çok daha geniş bir alternatif yelpazesinden seçim yapmaktadır. Bu nedenle, kişinin "aydınlık" veya "karanlık" ifadesi bir bitten daha yüksek bilgi taşır. BBT modelinde bilgi aslında çok sayıda alternatif arasından ayrım yapabilme yeteneğidir (2008, s. 218); kişinin ışığa bilinçli tepkisi ile fotodiyotun tepkisi arasındaki fark, birçok alternatif arasında ayrım yapabiliyor olmaktır. Fakat daha sonra Tononi "dijital kamera" örneği ile çok fazla alternatif arasından ayrım yapabilmenin bilinç için tek başına yeterli olmadığını bizlere gösterir. Bir dijital kamera sahip olduğu milyonlarca fotodedektör sayesinde bir insana kıyasla çok daha fazla alternatif arasında ayrım yapabilir, fakat bu durum dijital kameranın bilinçli bir varlık olduğunu göstermez. Bu durumda bilgi miktarı bilincin varlığı için gerekli olsa da yeterli bir koşul değildir. Ve bu noktada entegrasyonun, yani bilgiyi bütünleştirebilme kapasitesinin önemi ortaya çıkar. BBT'ye göre bilinçli bir sistem bilgiyi bütünleştirebilme kapasitesine sahip bir sistemdir ve bilinç bütünleşik bilgiden ortaya çıkar.

BBT modeline göre bilinci açıklamak için iki ana problemi çözmemiz gerekir. İlk problem, bir sistemin bilince sahip olabilmesini belirleyen koşulların ne olduğu ile ilgilidir (2004, s. 2). Nitekim bu sorunun çözümü bize bilincin ne menem bir şey olduğunu ve hangi koşullar altında bilinçli bir varlıktan ya da bilinç durumlarından söz edilebileceğimizi gösterir. Bu problemin çözümü ile bir sistemin üretebileceği bilinç miktarını bilebileceğimiz gibi bilincin niceliğini de ölçmüş olacağız. Bu sebeple bu problemi "bilincin nicelik problemi" olarak adlandırabiliriz. İkinci problem, bir sistemin ürettiği bilincin niteliğini belirleyen koşullarla, yani bu tezin ana konusu olan bilinçli deneyimin öznel karakterini oluşturan koşulların nasıl anlaşılacağıyla ilgilidir. Bu soruna "bilincin nitelik (veya qualia) problemi" diyebiliriz ve bu problemin çözümü, bir sistemin bilincin niteliğini ya da başka bir deyişle öznel karakterini nasıl ürettiğinin açıklamasını sağlayacaktır.

BBT fenomenoloji ile başlar. Fiziksel açıklamasını fenomenal olanın karakteristik özelliklerinden yola çıkarak yapmaya çalışır. Özetlemek gerekirse, bu teorinin iki temel iddiasından ilki, bilincin, beynin karmaşık dediğimiz sinir sisteminin bilgi bütünleştirme kapasitesine tekabül ettiği ve bütünleşik bilgi miktarının da ölçülebildiği ve Φ (fi) sembolü ile gösterilebildiğidir. Bizler pek çok şeyin bilincine sahip olabilen organizmalarız: Bir tiyatro sahnesini izlerken duvarda resmedilmiş olan figürleri görebilir ve dışarıdan gelen her sesi ayırt edebiliriz. İzlediğimiz veya izleyeceğimiz tüm film karelerini düşündüğümüzde her karenin bizim için belirli bir bilinç durumu vardır. Bu, beynimizin mekanizmasının içsel durumu sayısız olası alternatife ayırdığı ve bu bilginin bilinçli deneyim üretmek için beyin sistemiyle yüksek oranda bütünleştiği anlamına gelir. Beynimizin ilgili bölümleri arasındaki sebep-sonuç etkileşimi nedeniyle bütünlük halinde bir bilinç deneyimi yaşarız. Örneğin ne kadar uğraşırsak uğraşalım, içinde bulunduğumuz bölgenin sadece sağ tarafını bilinçli olarak algılayamayız ve sol tarafını görmezden gelemeyiz; ya da kahkaha atan bir arkadaşımızın yüzüne baktığımızda onun gülüşünü fark etmememiz pek olası değildir. Bilincimizde olan bilgiler alt bölümlere ayrılamaz; tam bir bütün olarak verilir. Bölünmüş beyin denilen sendromu düşünelim. Bu sendrom kontrol edilemeyen epilepsi tedavisi için iki yarım küreyi birbirine bağlayan korpus kallozumun cerrahi olarak kesilmesi sonucu ortaya çıkan bir tablodur. Bu

ameliyattan sonra hastaların bilinçlerinin ikiye ayrıldığı görülür; deneyimleri eşzamanlıdır, ancak bir bütün değildir. Hastalar sağ ve sol görme alanlarındaki uyaranları farklı algılar ve bir bütünlük oluşturamazlar; hastanın sağ yarım küresi sol görme alanının farkındadır ve sol yarım küre sağ görme alanının farkındadır. Ancak hasta kendisine sunulan uyarandan bir bütün olarak habersizdir. Bu durumda, bölünmüş beyin hastaları, "iki kortikal yarıküreyi birbirine bağlayan 200 milyon tel" kesildiği için olağanüstü bir bütünlüğe sahip değildir. (Koch, 2009). Bu örnek aynı zamanda bilinçli olmak için beynimizin nöral kompleksindeki öğeler arasında iletişim olması gerektiğini de ortaya koymaktadır; yani bütünleşik bir varlığa ihtiyacımız var.

Bu çalışmada yanıtlamaya çalıştığımız asıl problem bilincin öznel yanının bilimsel olarak açıklanıp açıklanamayacağı sorusudur. Tononi, BBT teorisi ile bunun bilimsel olarak açıklanabileceğini savunur ve bunu, beyindeki kompleks adını verdiği sitsemin içindeki elementlerin birbirleriyle olan bilgi ilişkilerinin yapısı ile tanımlar. Tononi'ye göre, iki farklı kompleks için veya farklı zamanlardaki aynı kompleks için tüm deneyim nesneleri ve bilgi entegrasyon değerleri aynı olsa bile, her iki kompleks için veya farklı zamanlardaki aynı kompleks için bilgiyi bütünleştirme şekli ve yolu farklı olacaktır ve bu durum bilincin özel ve öznel yönü yansıtmaktadır.

BBT sorunları olan bir teori olabilir, çünkü bilincin niceliğini ölçmek oldukça karmaşık olarak kabul edilebilir. Ancak teorinin eksikliğinden ziyade avantajlarına baktığımızda BBT, bilinçli deneyimlerin öznelliğini göz ardı etmeden bilinci fiziksel sistemde açıklamaya çalıştığı için bilinç sorununa bütüncül yaklaşan bir çalışma olarak değerlendirilebilir. İkinci avantajı ise bilincin sadece insanlar için değil, insan dışı varlıklar ve cansız varlıklar için de kapsayıcı bir şekilde ele alınması gerektiğini savunmasıdır. BBT için, bilinç sadece bir sistemin bilgi entegre etme kapasitesine bağlıdır ve bu demektir ki bir sistemin bilinç sahibi olabilmek için güçlü bir benlik duygusu, dil yeteneği, duygu veya çevresel koşullara bağlılık gibi koşullara sahip olmaya ihtiyacı yoktur. Dil ile ilgili olmadığı için sözlü bildirimin olmadığı durumlarda bilincin varlığından söz etmek mümkün hale gelir ve bu iddia ile hayvanlarda ve bebeklerde bilincin varlığının yanı sıra koma veya uyurgezerlik gibi

bazı nörolojik olguları da açıklar. Ayrıca bu teori, REM uykusunda giriş ve çıkış sinyallerinin zayıflamasına rağmen bilincin korunmaya devam ettiğini iddia etmektedir (2004, s. 20). Bu örnekler fi değerinin düşük olduğu durumlar olarak tanımlanabilir. Diğer bir deyişle, BBT'ye göre bilinçli deneyim, fi değerinin maksimum olduğu durumunda gerçekleşir. Sistemin entegrasyon kapasitesi arttıkça sinerji ve bilinç de artar. Kişiler fi değerinin maksimumun altında olduğu durumlarda deneyimlerini bildirmezler ancak bu fi değerinin sıfıra ulaştığı anlamına gelmez.

Bunların yanı sıra, matematiksel bir kavram olan bilginin ölçülebilir bir fiziksel varlık olduğunu kanıtlamaya çalışan önemli fizikçiler vardır ve bilginin fizikselliği bize fenomenal bilinci, hurafelere ya da umutsuzluğa yer vermeden, açıklama şansı verir. Örneğin, Landauer 1961 tarihli çalışmasında termodinamik ve bilgi arasındaki bağlantıyı vurgulayarak ölçülebilir bilginin fiziksel bir varlık olduğunu savunur.

Vopson, 2019'da yayınlanan bir makalede Einstein'ın genel görelilik teorisine dayanarak teorisini kurar ve bilginin fiziksel bir varlık olduğunu öne süren yeni bir kütle-enerji-bilgi denkliği ilkesi formüle eder. Teori, her bilgi parçasının sonlu ve ölçülebilir bir kütlesi olduğunu öne sürer (2020). Vopson'un teorisi, bilginin maddenin beşinci hali olduğu anlamına gelir ve bu fikir fiziğin eksik olan kısmını yeni ve heyecan verici bir şekilde tamamlar. Şu anda tanımlayamadığımız anlamda bilginin fizikselliği fiziksel dünyaya ait mevcut bilgimizin ve anlayışımızın gelişip değişmesine katkıda bulunmaktadır.

Vopson'ın yanı sıra ünlü fizikçisi Max Tegmark da bilincin bütünleşik bilgi kuramı çerçevesinde ifade edilebileceği görüşünü paylaşıyor. Kuantum mekaniği ve bilgi teorisini birleştirerek, bilinçteki gerçekliğin doğası hakkında bazı soruların bilimsel deneysel yöntemlerle araştırılabileceğini düşünüyor. Landauer ve Vapson gibi Tegmark da bilgiyi ölçülebilir fiziksel bir varlık olarak açıklar.

Fiziksel bir varlık olarak bilginin açıklamasını ve ispatını fizikçilere bırakıp böyle bir bilginin bu çalışmaya katkısına geri dönersek, bilginin şu anda anladığımızdan farklı bir anlamda fiziksel bir varlık olduğu iddiası bize fiziksel teorilerle çelişmeden

fiziksel dünyaya ve uzamsal olana dair algımızı genişletmenin mümkün olduğunu, temelde maddenin dörtten fazla hali olduğunu ve bilincin böyle bir fiziksel varlıkla açıklanabileceğini söyler.

Bütünleşik Bilgi Teorisi'nin bilince karşılık gelen belirli bir niceliğin hesaplanabilirliğine dair olan iddiasının sonuçsuz olduğu düşünülebilir. Zaten BBT'nin bilincin açıklanmasına dair nihai bir teori olduğunu iddia etmek de bu tezin amaçları arasında değildir. Ancak BBT'nin bilinç sorununu bütünsel olarak açıklama çabası olduğu açıktır ve bilinçli deneyimlerin öznelliğini yok saymaz. Ayrıca bilginin fiziksel olduğuna dair son yıllarda artan çalışmalar ve ortaya koyulan hipotezler fiziksel dünyaya ait mevcut bilgilerimizin değişip genişleyebileceğine ve bu doğrultuda bugün fiziksel dünyaya ait olmadığı düşünülen fenomenal bilincin bir fiziksel dünya nesnesi olarak tanımlanabileceğine dair umutlarımızı yeşertir. Yapılan araştırmalarla ileriki dönemlerde BBT teorisinin yanlış olduğu kanıtlansa bile, McGinn gibi bilincin açıklanamayacağına inanan gizemcileri çürüttüğü ve yeni nesil araştırmacıların bilincin açıklanması konusunda izleyeceği doğru yolda onlara katkıda bulunacağı ve ışık tutacağı açıktır.

C. THESIS PERMISSION FORM / TEZ İZİN FORMU

ENSTİTÜ / INSTITUTE

Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences ☐

Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences ☒

Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics ☐

Enformatik Enstitüsü / Graduate School of Informatics ☐

Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences ☐

YAZARIN / AUTHOR

Soyadı / Surname : Işıkgil
Adı / Name : Sena
Bölümü / Department : Felsefe / Philosophy

TEZİN ADI / TITLE OF THE THESIS (**İngilizce** / English): PSEUDO-MYSTERIANISM OF MCGINN: A CONSTRUCTIVE AND NATURALISTIC SOLUTION OF THE HARD PROBLEM OF CONSCIOUSNESS IS POSSIBLE

TEZİN TÜRÜ / DEGREE: **Yüksek Lisans** / Master ☐ **Doktora** / PhD ☒

1. **Tezin tamamı dünya çapında erişime açılacaktır. /** Release the entire work immediately for access worldwide. ☒
2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two years. *** ☐
3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. ***** ☐

***** *Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

Yazarın imzası / Signature ..... **Tarih** / Date .....
(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)
(Library submission date. Please fill out by hand.)

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.