AN ASSESSMENT OF THE STATUS OF THE MULTIPLE REALIZABILITY THESIS IN COGNITIVE SCIENCE

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

BY

EMIN UMUT BAYSAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

MAY 2011

Approval of the Graduate School of Informatics

_____

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of master of science.

_____

Prof. Dr. Deniz Zeyrek
Head of Department

This is to certify that I have read this thesis and in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____

Dr. Ceyhan Temürcü
Supervisor

Examining Committee Members

Prof. Dr Deniz Zeyrek (METU, COGS) _____

Dr. Ceyhan Temürcü (METU, COGS) _____

Assist. Prof. Dr. Annette Hohenberger (METU, COGS) _____

Assist. Prof. Dr. Mehmet Cem Kamözüt (Ege, PHIL) _____

Prof. Dr. Ayhan Sol (METU, PHIL) _____

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

Name, Last name: Emin Umut Baysan

Signature: _____

ABSTRACT

AN ASSESSMENT OF THE STATUS OF THE MULTIPLE REALIZABILITY THESIS IN COGNITIVE SCIENCE

Baysan, Emin Umut
M. Sc. Department of Cognitive Science
Supervisor: Dr. Ceyhan Temürcü

May 2011, 61 pages

It has been argued that there are physically different ways of instantiating mental properties, the nature of which is the subject matter of cognitive science. This claim has been known as the Multiple Realizability Thesis (MRT). It has been suggested that the MRT shows that a reduction of mental properties to physical

properties is impossible, as there cannot be one-to-one correspondences of mental properties to the properties of the brain. Moreover, it has been argued that the latter point shows that physical explanations are not relevant to the explanations of cognitive science, as they would lack the generality of psychological explanations.

This thesis will try to explain from which assumptions of a traditional cognitive science perspective the MRT follows. It will also discuss several responses that have been introduced against both the MRT and the anti-reductionist conclusions that are assumed to follow from it. The responses include a challenge to the scientific status of cognitive science. According to this challenge, the MRT entails that the subject matter of cognitive science, namely mental properties, lack a similarity in the physical level, hence an instance of a mental property is not informative about another instance.

While discussing these theories, a revision of the MRT will be proposed. According to this revision, the MRT is compatible with the assumption that there could be an underlying similarity between different physical realizers of a given mental property. It will be argued that by means of this revision, both the challenge to the scientific status of cognitive science, and the argument for the irrelevance of physical explanations will fail.

Keywords: Multiple Realizability, Reduction, Special Sciences, Philosophy of Cognitive Science

ÖZ

ÇOKLU GERÇEKLEŞİM TEZİNİN BİLİŞSEL BİLİMDEKİ KONUMUNUNUN BİR DEĞERLENDMESİ

Baysan, Emin Umut
Yüksek Lisans, Bilişsel Bilimler
Tez Yöneticisi: Dr. Ceyhan Temürcü

Mayıs 2011, 61 sayfa

Doğası, bilişsel bilimin çalışma alanını oluşturan zihinsel özelliklerin ortaya çıkması için çeşitli fiziksel yollar olduğu iddia edilmektedir. Bu iddia, Çoklu Gerçekleşim Tezi (ÇGT) olarak bilinmektedir. ÇGT'nin, zihinsel özelliklerin fiziksel özelliklere indirgenemeyeceğini gösterdiği söylenmektedir. Çünkü, ÇGT'ye göre, zihinsel özelliklerin beyinde birebir fiziksel karşılıkları yoktur. Dahası, bu en son bahsedilen noktadan dolayı, bilişsel bilimde fiziksel açıklamaların yeri olmayacaktır,

çünkü sözü edilen fiziksel açıklamalar, psikolojik açıklamaların sahip olduğu genel-geçerlik özelliğinden yoksundurlar.

Bu tez, ÇGT'nin, geleneksel bilişsel bilim bakış açısının getirdiği hangi varsayımlardan türediğini açıklamaya çalışmaktadır. Ayrıca, hem ÇGT'ye, hem de ondan türetilen indirgemecilik karşıtı sonuçlara cevap olarak sunulmuş çeşitli iddialar da tartışılmaktadır. Bunlar arasında, bilişsel bilimin bilimsel konumunu sorgulayan bir iddiaya da yer verilmektedir. Bu iddiya göre, ÇGT, bilişsel bilimin çalışma alanını oluşturan zihinsel özelliklerin, fiziksel açıdan bir benzerlik taşımadığını göstermektedir. Bu da, aynı zihinsel özelliğin çeşitli örneklerinin birbirleri hakkında bilgi verememesine neden olmaktadır.

Bu iddialar tartışılırken, ÇGT'nin bir revizyonu sunulmaktadır. Bu revizyona göre, ÇGT, bir zihinsel özelliğin çeşitli gerçekleştiricilerin birbirine fiziksel açıdan bir noktaya kadar benzerlik gösterebilecekleri varsayımı ile tutarlıdır. Sözü geçen revizyona göre, hem bilişsel bilimin bilimsel konumuna olan itirazın, hem de bilişsel bilimde fiziksel açıklamalarının yeri olmadığı iddiasının geçerli olmadığı gösterilmektedir.

Anahtar Kelimeler: Çoklu Gerçekleşim, İndirgeme, Özel Bilimler, Bilişsel Bilim Felsefesi

Burcu'ya

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor Ceyhan Temürcü for his support, guidance, comments, criticisms and encouragement. Without him, I wouldn't be able to specify the problems that are discussed in this work, and realize the gaps in the solutions that I provide.

I would like to thank Didem Gökçay for her support as a teacher, and the members of the faculty at the Cognitive Science Department at METU for their helps and the courses they have offered. I would also like to thank Cem Kamözüt and John Bolender for their inspirations on me as analytical philosophers. Moreover, I am grateful to Dinçer Çevik, Volkan Çifteci and Işık Sarıhan, with whom I shared the excitement of working on intriguing problems.

I would like to express my gratitude to TÜBİTAK (The Scientific and Technological Research Council of Turkey) for the generous financial support it has provided from 2008 to 2010.

And finally, I would like to thank three important people in my life for their comprehensive support: My father Candan Baysan, my mother Gül Tekay Baysan, and Burcu Aydoslu.

TABLE OF CONTENTS

ABSTRACT……iv
ÖZ……vi
DEDICATION……..viii
ACKNOWLEDGEMENTS…….ix
TABLE OF CONTENTS…….....x
CHAPTERS
1 INTRODUCTION…….……1
1.1 Scope and Aim of the Thesis……..1
1.2 Structure of the Thesis……3
2 THE PHILOSOPHICAL CONTEXT: REDUCTIONISM…….……...……...5
2.1 Introduction: Reduction……..5
2.2 Approaches to Mind and Body……...6
2.1.1 Substance Dualism……...……7
2.2.2 Behaviorism……8
2.2.3 Identity Theory as a Paradigm for Reductive Physicalism…….8
2.3 Concluding Remarks…….10
3 COGNITIVE SCIENCE AND MULTIPLE REALIZABILITY……11
3.1 Introduction……...11
3.2 The Computational Theory of Mind (CTM)……....12
3.2.1 Computation, Mental States and Functionalism……12
3.2.2 The Language of Thought Hypothesis (LOTH)……13
3.2.3 The Status of the CTM in Cognitive Science……....16
3.3 The Multiple Realizability Thesis (MRT)……16
3.3.1 Multiple Realizability……17
3.3.2 Functional Isomorphism……18

3.3.3 Evidence for the MRT……...19
3.4 Anti-Reductionism……21
3.4.1 Special Sciences…….21
3.4.2 Physical Explanations and Cognitive Science……...23
3.5 Concluding Remarks…….25
4 RESPONSES TO ANTI-REDUCTIONISM…….……...…27
4.1 Introduction…….……..27
4.2 Kim's Argument against Special Sciences……...27
4.3 Universality and Physical Explanations……...……30
4.4 Sober on the Relevance of Physical Explanations……31
4.5 Cognitive Science Challenged……..33
4.5.1 The Challenge……33
4.5.2 Responding to the Challenge: A Revision…….34
4.6 Concluding Remarks…….36
5 AGAINST THE MULTIPLE REALIZABILITY THESIS…….……38
5.1 Introduction……...38
5.2 Argument from "Qualia"……..38
5.3 Bechtel & Mundale: The Practice of Neuroscience…….....40
5.3.1 Psychological and Neural Properties…….……40
5.3.2 Generalizing Across Species…….42
5.3.3 Differences in Human Brains……43
5.3.4 Why Has the MRT Been Popular? ……...…44
5.4 Shapiro Against the MRT……45
5.4.1 Functional Isomorphism Revisited……45
5.4.2 Conditions for Multiple Realization……..46
5.4.3 Are there Genuine Multiple Realizations? ……47
5.5 A Rejoinder to Shapiro and Bechtel & Mundale……..49
5.6 Embodied Cognition and the MRT……...51
5.7 Concluding Remarks…….55
6 CONCLUSION…….……...56
6.1 A Summary……...56
6.2 A Conclusion…….58

BIBLIOGRAPHY……..60

CHAPTER 1

INTRODUCTION

1.1 Scope and Aim of the Thesis

The aim of this thesis is to assess the current status of the Multiple Realizability Thesis (MRT) in cognitive science. Briefly, the MRT is the thesis that different sorts of physical systems could have similar or same mental properties. That is, a humanly mental life, for instance, is not necessarily restricted to human-like biological organisms. Although, this is a philosophical claim, there are at least three reasons why its discussion is in the scope of cognitive science:

1-The MRT was first formulated as a thesis against various forms of Reductionism, including Theoretical Reductionism, which suggests that higher-level theories could be reduced to lower-level theories. A theory is taken to be a higher-level theory with respect to a lower-level theory where the properties of the former have realizations by the properties of the latter. As human beings, we are biological organisms, and as biological organisms, we are composed of physical systems. Hence, our mental properties are realized by our physical organization. Thus, the reductionist argues that cognitive science, conceived as the study of the mental life of human beings (and perhaps other relevant systems which have mental properties), can be reduced to a lower-level science, such as physics. The MRT came as a response to this claim, as it showed that there is no one-to-one correspondence between a mental property and a physical property. That is, if the same mental property is realized by different systems by means of different physical properties, one cannot reduce a mental property to a given physical property. Moreover, quite ambitiously, it has been claimed that,

because of the multiple realizability of mental properties, physical explanations do not have any relevance to the explanations of cognitive science. Therefore, the discussion of the MRT will reveal us insights as to the status of cognitive science in a wider scope of science, its reducibility to physics, and even its relations to other sciences.

2-The MRT is widely assumed to be an extension of the Computational Theory of Mind (CTM), which has been one of the working hypotheses of the standard cognitive science for decades. According to the CTM, the mind is best understood as a computational device. It is assumed that different physical systems could perform the same computations. Thus, mental properties have no, or few, restrictions with respect to the physical properties, insofar as mental properties are taken to be computational properties. If there is a deductive connection between the CTM and the MRT, one would argue that the MRT could be taken as a working hypothesis of the standard cognitive science. Therefore, a discussion of the MRT will provide us new points of view with respect to more general problems of cognitive science.

3- The MRT has also been used as a thesis for showing that mental properties are not scientific kinds. There is a lively discussion among the philosophers of mind with respect to the status of mental properties in science. Under the assumption that mental properties are the properties which cognitive science has been working on, this problem is a problem of cognitive science itself. The elaboration of the claim that mental properties are not scientific requires a technical background which, hopefully, will be given in what follows. Yet, for the purposes of this introduction, it should just be noted that a discussion of this problem will give the cognitive scientist some intuitions about the philosophical grounds on which the cognitive science is standing. For a cognitive scientist, the question of whether cognitive science is a genuine science or not would seem to be trivial, however, this should not let us ignore questions about its foundations.

This thesis will provide a philosophical discussion of these three main problems in order to provide tools for confronting them. The thesis will discuss and

try to reply several objections to the MRT and its implications against Reductionism while trying to propose a revision about the extent of the MRT itself.

1.2 Structure of the Thesis

Chapter 2 will introduce the philosophical context in which the MRT was first formulated. It will expand on the notion of reduction as a relation between theories, and provide a brief sketch of different metaphysical theories regarding the relation between the mind and body.

Chapter 3 will introduce the MRT. It will take the MRT as an extension of one of the basic working hypotheses of the standard cognitive science, namely the Computational Theory of Mind (CTM). It will then present the special sciences argument, which uses the MRT as a premise to the conclusion that Reductionism is false. This chapter will also discuss the arguments for the irrelevance of physical explanations in cognitive science.

Chapter 4 will discuss the responses to the anti-reductionist consequences of the MRT. Moreover, an argument will be introduced which is taken by some to pose challenges to the scientific status of cognitive science. The challenge will be responded in this chapter by means of proposing a revision with respect to the extent of the MRT itself. I will propose that such a revision will demonstrate that the argument for the irrelevance of the physical explanations is likely to fail.

Chapter 5 will provide a sketch of the accounts which directly try to refute the MRT itself. One of such responses to the MRT gains its strength from the Embodied Cognition (EC) research program, which however, refutes some basic assumptions of the standard cognitive science.

The conclusion chapter will introduce a brief summary of the discussions and will emphasize the revision of the MRT that is proposed in Chapter 4. Briefly, it will be concluded that, the MRT could not be refuted as long as the working hypothesis of the standard cognitive science is held. However, dealing with the challenge to the cognitive science raising the questionability of the scientific status of the mental properties will require a revision of the MRT itself. The revision will show that the extension of the MRT should be restricted to some degree, and this restriction will provide us some grounds for establishing the relevance and the importance of physical explanations with respect to cognitive science. This means that, although

total refutations of the MRT fail, some consequences of the MRT, such as the claims about the irrelevance of the physical explanations in cognitive science, must be dispensed with.

CHAPTER 2

THE PHILOSOPHICAL CONTEXT: REDUCTIONISM

2.1 Introduction: Reduction

With the development of the positivist philosophy of science in the early 20th century, and its emphasis on the importance of empirical data in sciences, physics was assumed to be the ultimate science to which all theories would be reduced. A classical account of such an idea was introduced by Oppenheim and Putnam (1958), where they argued that, in an ideal scientific practice, the concepts and the laws of all sciences could be unified by means of reducing them to a basic and elementary paradigm. As Fodor states, the positivist philosophers of science took this ideal as a motivation to be achieved "in the long run" (1974, p. 97). The assumption was that physics might be incomplete or it might not be as successful as it was supposed to be, but with a complete and successful physical paradigm, all theories of other sciences would be reduced to theories of physical sciences.

The crucial concept regarding the discussion of the unity of science is the concept of "reduction", and it needs to be scrutinized by mentioning the classical accounts to which the reductionist philosophers of science appealed. What philosophers mean by reduction is generally the type of theoretical reduction introduced by Ernest Nagel (1961), which is called the Nagel model (Kim, 1996, p. 212). Reduction, according to the Nagel model, is a relation between two scientific theories, where a theory is taken to be a set of basic laws and all statements that can be logically derived from those laws. In the Nagel model, there are two theories, T_1

and T_2. If T_2 (target theory) is going to be reduced to T_1 (base theory), laws of T_2 are supposed to be logically derivable from the laws of T_1.

Nevertheless, as T_1 and T_2 are different scientific theories, their terminologies, or, 'scientific properties' that they use may be different, and this is almost exclusively the case with all theories that philosophers of science mention regarding reduction. Therefore, for reduction in the sense of the Nagel model, one needs statements which bind the theoretical terms of T_1 and T_2. These binding statements are called 'bridge principles' or 'bridge laws' in this model.

To present the issue in a formal way, assume that *(1)* $F \rightarrow G$[1] is a law of T_2, (2) F and G are not among the expressions of T_1 while F^* and G^* are, and *(3)* $F^* \rightarrow G^*$ is a statement of T_1. The bridge principles that are needed are then *(2a)* $F \leftrightarrow F^*$, and *(2b)* $G \leftrightarrow G^*$. From *(3)*, *(2a)*, and *(2b)*, one can derive *(1)*. This means that, *(1)* could be reduced to *(3)*. If all of the statements of T_2 can be reduced to statements containing expressions of T_1, this would be a typical case of reduction according to the Nagel model, in the sense that a theory is reduced to another by means of showing that all laws of the former are derivable from the laws of the latter in addition to the bridge principles. This means that the reduced theory is implicitly contained in the reducing theory. In other words, the reduced theory is not an independent theory with its own domain.

2.2 Approaches to Mind and Body

Reductionism is a hot topic not only in cognitive science but also in philosophy of mind. The discussion revolves around the question of whether minds can be reduced to brains, or whether the properties of cognitive science can be reduced to the properties of the brain or the rest of the body. However, in the scope of theoretical reduction, one could also ask whether cognitive science, which is a science about mental properties, is reducible physics, which is a science about the physical properties of the world. Even before considering the reducibility of minds to brains, or cognitive properties to neural properties, there was the question of how these minds and bodies are related. Different answers to this question will come along with different prospects for reduction.

---

1 *F* and *G* are predicates. Thus, $F \rightarrow G$ reads as “If something is F, then it is G”.

2.2.1 Substance Dualism

Prima facie, a mind and a body seem to be distinct from each other. On the one hand, we have an idea of ourselves as having mental lives, yet on the other hand, we know that we have bodies which are material things extended in space. This view is traditionally called Substance Dualism, and by definition, it is the view that a human being is composed of two distinct substances, namely the mind and the body. When it comes to the question about the reducibility of the mind to the body, Substance Dualism suggests that there cannot be such a reduction, as the mind and the body are completely different substances. Being a substance, in the context of Substance Dualism, is the capability of having independent existence. Hence, the mind could exist without the body.

However, the following question about theoretical reduction still seems to be available: Can we reduce cognitive science, as a science of mind, to physics? The Substance Dualist has to give a negative answer to this question, as if different substances could exist independently from each other, the bridge laws that have been introduced in the discussion of the Nagel model would not be available. That is, there would be nothing to guarantee that there is a necessary relation between a mental and a physical property.

Substance Dualism has some problems, and one of the major problems it encounters is the problem of mental causation. That is, the causal relation between mental and physical events is an everyday phenomenon. According to the formulation which is given by Kim (1996, p. 125), there are three types of mental causation: (i) *mental-to-physical causation,* as in the desire to write something and its effect on the fingers; (ii) *physical-to-mental causation,* as in the physical sensations, such as severe burns, causing psychological states, such as pain; and finally (iii) *mental-to-mental causation,* as in believing something and consequently deciding to do something because of that belief. The first two types of mental causation pose severe problems to substance dualism, because, if minds and bodies are distinct as the substance dualist thinks, the causal relations between, say, desiring something and moving fingers, or severe burns and consequent pains are mysterious. Thus, from a scientific perspective (if it is one of the aims of science to study the causal phenomena) Substance Dualism is problematic.

2.2.2 Behaviorism

In the first half of the 20th century, with the positivist emphasis on the importance of observation in sciences, Substance Dualism was widely disregarded among psychologists and empirically minded philosophers of mind. That epoch witnessed the rise of Behaviorism as a scientific and philosophical program. Behaviorism has been taken to have different versions such as Methodological Behaviorism, and Logical Behaviorism. The former was a paradigm in psychology, whose supporters argued that psychologists should only talk about the observable relations between stimuli and behavioral responses. On the other hand, the latter was a school of thought among some philosophers, and those who believed in it argued that Substance Dualism is scientifically false as a matter of logic. According to this latter view, a statement about a mental state (such as a belief) could be translated as a statement about the sensory inputs (such as seeing something) and motor outputs (such as reporting something).

To state briefly, it can be understood that Behaviorism, as a psychological paradigm and a philosophical view, suggests that what should be studied consists of the relations between observable phenomena, and it eschews the talk of the relations between internal states of a person.

For some empirical and methodological reasons, behaviorism became more and more unpopular by the middle decades of the 20th century. First, Chomsky's (1959) attack on and refutation of behaviorism from an empirical point was successful: He showed that it is impossible, even in principle, that behaviorism could explain the capacity to understand and generate language. Second, as cognitive scientists, we want to talk about internal mental states, and since Behaviorism does not allow us to do that, it does not seem to be a tenable alternative.

2.2.3 Identity Theory as a Paradigm for Reductive Physicalism

The refutation of Behaviorism does not entail that Substance Dualism is true. There is, at least, one other alternative, namely the Identity Theory, which identifies mental properties with properties of brain. Thus, according to the Identity Theory, which was formulated by philosophers such as Smart and Place, mental states such as feeling of pain, or believing that it is going to rain, are states of the brain (or the relevant parts of the nervous system). The famous analogy is that just like it was

discovered that the lightning is identical with an electrical discharge, it was also discovered that the mind is identical with the brain.

Considering reduction with respect to the Nagel model, it can be said that identity theory is a reductive theory, because, with the identification of mental properties and properties of brain, one would have bridge principles connecting the terminology of cognitive science with the properties of the science of the brain as a physical science. For this reason, in what follows, this view will be called Reductive Physicalism.

Reductive Physicalism also rejects Substance Dualism, but the motivation of this rejection is quite different from the motivation of Behaviorism. According to Reductive Physicalism, the reason why Substance Dualism is false is not a matter of logic, as behaviorists argue, but a matter of fact. This means that, it is logically possible, but factually false that there are minds in a substance dualist's sense in addition to bodies.

Another important difference between Behaviorism and Reductive Physicalism is that Behaviorism avoids speaking of internal states when it comes to psychology. However, Reductive Physicalism allows the speech of internal states as long as these internal states are the states of the brain. Place (1956) argued that the acceptance of the existence of an inner process does not show that a mental process is not a brain state, as these inner states could be inner states of the brain.

What can be said about the reports of person's psychological states, rather than the states themselves, such as sensations? According to Smart's (1959) identity theory, mental states (such as sensations) are *nothing over and above* states of the brain in the same sense that nations are nothing over and above citizens. Yet, the phrase "nothing over and above" does not suffice to say that statements about mental states are translatable to statements about brain states. Here the point is that identity and translatability are not the same things.

The last point is relevant to a possible objection to Reductive Physicalism. As it can be foreseen, one might object to such identity statements about psychological properties and say that mental properties cannot be identical with properties of the brain, because someone completely ignorant of his properties of brain could perfectly say that he is in pain (as a psychological property). However, Putnam, who is also

known to be an ardent opponent of Reductive Physicalism, argues that this is "to collapse the two notions of 'property' and 'concept' into a single notion" (1975, p. 430). Concepts, as Putnam argues, are identified by means of synonymy classes, however, identifying two properties does not require that the two properties have the same meaning.

The last point is best introduced with this analogy: "'[T]emperature is mean molecular kinetic energy' appears to be a perfectly good example of a true statement of identity of properties, whereas the concept of temperature is the same concept as a concept of mean molecular kinetic energy is simply false" (ibid., p. 431). Therefore, as Putnam rightly suggests, that properties are identical does not show that concepts about these properties are identical. And this shows why the Identity Theory does not suggest straightforward translations as the translation of "I have pain" to "I have C-fiber stimulation".

Moreover, Smart argues that the principle of simplicity favors the Identity Theory over Substance Dualism. As science shows us, among many, psychological events and physical events (events in the brain) are *at least* correlated. Naturally, philosophers and scientists want to explain these correlations in the most elegant way, and elegance in science comes with simplicity. Finding these correlations is an empirical matter; however, once these correlations are discovered, there is one next move, namely the move to provide a good explanation of this correlation. The reductive physicalists made this move by simply saying that psychological states and brain states are identical.

2.3 Concluding Remarks

In this short chapter, the notion of reduction, in the context of cognitive science, has been introduced. It has been shown that the prominent account of reduction, namely the Nagel model, makes use of bridge principles. Furthermore, implications of different positions in the philosophy of mind have been discussed, and it has been shown that Reductive Physicalism accounts for bridge principles as statements of identity between psychological and physical properties. The next chapter will discuss why and how these identity statements are problematic from the classical cognitive science perspective.

CHAPTER 3

COGNITIVE SCIENCE AND MULTIPLE REALIZABILITY

3.1 Introduction

In the previous chapter, the concept of reduction, and its relevance to philosophy of mind and cognitive science have been discussed. Moreover, Reductive Physicalism has been introduced as the theory that psychological properties are identical to properties of the brain. If this theory is true, one would expect to see that a theoretical reduction would also follow so that cognitive science as a whole would be reduced to the science of the brain. Because, according to the model of reduction at hand, namely the Nagel model, the reductionist needs bridge principles, and the identification of mental properties with the properties of brain would provide perfect bridge principles. This chapter is mainly about the thesis that there are no such identity statements, as an implication of the classical cognitive science perspective.

The following words from Ned Block set the agenda of this chapter:

> Whatever the merits of physiological reductionism, it is not available to cognitive science point of view ... According to cognitive science, the essence of the mental is computational, and any computational state is "multiply realizable" by physiological or electronic states that are not identical with one another, and so content cannot be identified with any of them (1990, p. 146).

In the rest of this chapter, this quote will be elaborated by means of providing the original arguments for each point made in the quote. There are, at least, three notions which are important as for the relevant discussion: (i) *reductionism*, (ii) *the mind as a computer*, (iii) *multiple realizability of computational states*. In the

previous chapter, issues related to (i) have been discussed. Thus, in this chapter accounts about (ii) and (iii) will be introduced.

In 3.2, what philosophers of mind and cognitive scientists mean when they say that the mind is computational will be discussed. For this, the Computational Theory of Mind (CTM), which has been more or less a working hypothesis in classical cognitive science, is of utmost importance. In 3.3, what multiple realizability means, and why it is argued that mind is multiply realizable will be discussed. In 3.4, the Anti-reductionist argument from the MRT will be introduced.

3.2 The Computational Theory of Mind (CTM)

3.2.1 Computation, Mental States and Functionalism

According to the CTM, human beings have representations as cognitive codes physically instantiated in their brains, and instantiation of other representations and production of behavior are outcomes of operations that are carried out on these representations. This is taken to be what computers do, thus the CTM proposes that cognition is a type of computation (Pylyshyn, 1984, p. xiii). To illustrate the point that the behaviors of humans are consequences of some operations on cognitive codes, what the supporters of the CTM mean by these operations should be introduced. For this, there are some important notions that have to be mentioned. The first is the notion of a Turing Machine (TM), named after Alan Turing, who is taken to be the theoretical father of the modern computer.

A TM is an abstract device designed so as to carry out an operation by using symbols such as *0* and *1*. It has a potentially infinite tape on which it can print symbols, a head that scans and prints symbols, and a finite number of states it can go into. The machine follows rules which are sensitive to the states it is at and the symbols it scans. Following these rules, it can erase a symbol, move to another symbol on the tape, print another symbol, and change the state.

Another notion which was crucial for the development of the CTM was that of a Probabilistic Automaton (PA). A PA is a kind of a TM which works in accordance with probabilistic rules. That is, whereas a rule of a TM would be like (i) if the machine is in state *a*, and reads *1* on the tape, it erases *1*, writes *0*, moves the scanner one slot left, and changes the state to state *b*, a PA would have a rule such as (ii) if

the machine is in state *a* and …, *there is such and such probability* that it will do such and such.

Putnam argued that the notion of a PA can be extended so that it would allow for processing sensory inputs, internal states, and motor outputs (1975, pp. 433-4). That is, there would be a machine table of a PA such that it specifies a rule for determining the probability the motor output and the new internal state for every possible combination of sensory inputs and internal states. This extension gives Putnam the reason to argue that if an organism is capable of having a mental state, then it is a PA since mental states are in a similar relation with sensory inputs, motor outputs and other mental states.

The idea that psychological states can be defined with respect to the states of a PA is closely related to the view called Functionalism. What the machine does is described in accordance with the causal relations which it is engaged in, and the causal network at hand is assumed to be describable in terms of a functional organization.

The basic assumption that mind is some sort of a PA has been an important step for the development of the CTM, and it has set the general agenda of a research program. This research program was based on the assumption that the mind could be studied in computational terms, and the way the mind really worked could be discovered just by finding out what kind of programs were implemented in the brain.

3.2.2 The Language of Thought Hypothesis (LOTH)

The CTM is an influential theory in the domain of cognitive science[2], and it is more or less the working hypothesis in many studies. It, mainly, is the thesis that thinking is processed in a mental language (Fodor, 1975). The fact that the nature of the mental states is computational means that psychological processing could be explained in terms of a syntactic operation. Such a view is the grounding assumption of the Language of Thought Hypothesis (LOTH) (Fodor, 1975; Fodor & Pylyshyn,

2 Shapiro (2007; 2011) provides a distinction between standard cognitive science and the Embodied Cognition research program. The difference between these two will be discussed in 5.5. Yet, it should be noted that what is called "cognitive science" in this chapter refers to the standard cognitive science in this context.

1988). The LOTH can be taken as an indicative theory for a computational approach to the mind, so it is worth discussing in the context of the CTM.

Before discussing the LOTH, there is a view which has to be introduced, namely, Representationalism. It is the view that postulating the existence of representational states encoding the states of the world is crucial for the study of cognition (ibid., 1988, p. 7). Fodor and Pylyshyn argue that realism with respect to representations is crucial in a theory of cognition, because, in the psychological level of explanation, the task of the scientist is to find out how the psychological representations are processed[3].

The notion of a representation, and realism about it, are important with respect to the LOTH, because, the LOTH gives an account as to how these representations are processed in the mind. According to Fodor and Pylyshyn, a theory in cognitive science should postulate the LOTH, in the sense that it should be based on the assumption that these representations are processed by means of a syntactic structure. This means that the structure of the representations is important for psychological processing. In other words, such processes are structure sensitive: The way mental states are transformed is based on the structural properties of mental representations (ibid., p. 13).

The idea that the processing is sensitive to the structural properties of representations brings together several features a theory of cognition should postulate. These are:

(i) Productivity;

(ii) Systematicity;

(iii) Compositionality;

(iv) Coherence.

What these features mean, and how they are grounded on empirical findings of cognitive science will be elaborated under this section. Hence, that the CTM is some kind of a working hypothesis for cognitive science will be grounded, considering that the LOTH implies the CTM, and moreover that the LOTH demands

3 This view is contrasted with Eliminativism, which states that there is no place for representations in a theory of cognition.

that these four features should be available in a theory of cognition. That is, a good theory of cognition should account for these four phenomena:

(i) Productivity: Thought is assumed to be productive in the sense that it is able to comprehend and produce propositions taken from an infinite set of propositions. And, this is managed by means of using finite resources. Thus, it is assumed a recursive operation is being held on these finite resources so as to be capable of producing an infinite number of propositions. According to Fodor & Pylyshyn, this is a feature that has to be available in the correct comprehension of the nature of the mind, and the LOTH attributes this capacity to the mind (ibid, p. 33). Productivity is an outcome of the distinct nature of the memory and the program that runs the processes. That the memory and the program are distinct enables the altering of the memory without any change in the computational structure of the program.

(ii) Systematicity: It is suggested that there is a connection between comprehending one thought and comprehending other thoughts. Here, Fodor and Pylyshyn appeal to an analogy from linguistic competence again. So, to say that linguistic capacities are systematic is to say that the ability to produce/understand a sentence is intrinsically connected to produce/understand another. A person who can produce the sentence *John loves the girl* can as well produce the sentence *the girl loves John* (ibid., p. 39). The same is true for thoughts in general: The ability to think that John loves the girl comes with the ability to think that the girl loves John; these two thoughts are intrinsically connected to each other (ibid., p. 39).

(iii) Compositionality: The facts of systematicity are not arbitrary from a semantic point of view: the sentences *John loves the girl* and *the girl loves John* have also close semantic relations. In both sentences, *loves* has the same semantic function. But, there is no semantic similarity between these two sentences and the sentence *a car is more expensive than a bike.* This relation between syntactic and semantic roles comes from the principle of compositionality: That language is systematic entails that a lexical item of a language has same or similar functions in every sentence where it is used. According to the LOTH, what is true for language with respect to its grammaticality is true for the psychological processes in general with respect to the facts of compositionality.

(iv) Coherence: The mind is assumed to be able to make semantically and syntactically plausible inferences. As for the coherence issue, Fodor & Pylyshyn claim that since it is a logical law that *P & Q entails P*, there must also be some kind of a tendency of mind (presumably as a consequence of a psychological law) to think that *P* when *P & Q* is comprehended (ibid., p. 46). According to the LOTH, if such a psychological law is going to be accounted for, one has to appreciate the syntactic nature of psychological processes, because this logical structure can be preserved by syntactic processes.

3.2.3 The Status of the CTM in Cognitive Science

Now that it is clear that the LOTH is an indicative of computational approach to mind, one advantage of viewing mind as some sort of a computational device should be noted. For this, the notion of recursive decomposition is important. The computational approach to the mind makes it possible that the processes of the mind which are complex are decomposable to simpler processes by means of applying simple procedures repeatedly (or recursively). Therefore, with such a recursively decomposable account, the mind can be studied by working on the programs which work in a TM-like manner. That the mind works in a similar way to how a TM works, that is, that it can be broken down to the repeated application of the same simple operations is an appealing idea, and it shows how competence can be acquired without comprehension (Dennett, 2009).

The gist of the notion of a TM with respect to designing effective computers by means of simple operations attracted the philosophers of mind, such as Putnam, so he developed an understanding of mind in computational terms. And computational theories of mind, such as the LOTH, started to develop, and as the founders of those theories argued, they were consistent with the best scientific theories about human psychology, such as Chomsky's linguistic theory. This resulted in a foundation of a new research program, which is called "cognitive science".

So far, the first component of Block's view stated in 3.1, namely that (1) according to cognitive science, the mind is computational, has been elaborated. That is, it has been shown how the mind was interpreted as a computational device by some philosophers, and how some theories of cognitive science were developed in accordance with this view. In the next section, the second component of Block's

view, namely that (2) any computational state (hence, any mental state) is multiply realizable, will be discussed in detail.

3.3 The Multiple Realizability Thesis (MRT)

3.3.1 Multiple Realizability

Trivially, that something is multiply realizable means that it can be realized in multiple ways. An artifact, for example a clock, is multiply realizable in the sense that, there are more than one ways in which a clock can be realized. It can be made out of silver, or out of gold. So, it is clear how artifacts can be realized in multiple ways. However, as philosophers such as Putnam argue, computational states are multiply realizable as well. And since the CTM argues that mental states are computational states, the argument that mental states are multiply realizable follows naturally, since any computational state is multiply realizable. Moreover, this seems to be an implication of a cognitive science point of view, because the CTM is taken to be an implication, even the working hypothesis, of cognitive science itself. In what follows, arguments for the Multiple Realizability Thesis (MRT) from philosophers of cognitive science will be introduced.

In the philosophy of cognitive science literature, there are several accounts in which it is argued that this kind of multiple realizability applies to the properties of cognitive science, namely to the psychological properties. The prominent accounts of this are found in Putnam (1960; 1967a; 1967b), Fodor (1974; 1997), and Block & Fodor (1972). Furthermore, in some accounts, it is argued that this view is more like a consensus in the contemporary philosophy of mind, as LePore and Loewer argue when they say that the MRT "is practically received wisdom among philosophers of mind" (1989, p. 179).

Piccinini (2003) maintains that a technical definition of the term *multiple realizability* can be made as follows:

> *Multiple realizability* about a state or system X under description D1 is the thesis that (i) there is another description D2 of X, (ii) both D1 and D2 have explanatory value within some science of X, and (iii) X under description D2 is a realization of X under description D1, i.e. D1 also applies to classes of systems that D2 does not apply to. For short, … D2 realizes D1 (ibid, pp. 157-8, emphasis in the original).

According to this definition, *metal* is multiply realizable, and it can be realized by *iron*, or any other object that can be described as *metal* as Piccinini notes (ibid., p. 158). Because, the same thing can be described as metal or as iron, and the description of the former applies to pieces of other metals, whereas the latter does not apply to them. Moreover, both descriptions are used in a science (e.g. chemistry). However, this definition of multiple realizability is too general and it is difficult to follow its implications to cognitive science.

3.3.2 Functional Isomorphism

Another technical definition of multiple realization which would be more appropriate for showing how it is assumed that mental states are multiply realizable comes from Putnam (1975). This definition recruits another technical notion, namely the notion of functional isomorphism. "Two systems are functionally isomorphic if there is a correspondence between the states of one and the states of the other that preserves functional relations" (ibid., p. 291, emphasis deleted). A definition of multiple realization can be derived from this understanding of functional isomorphism in the sense that, according to this understanding, two systems can have different physical constitutions and be functionally isomorphic. That is, "computer made of electrical components can be isomorphic to one made of cogs and wheels or to human clerks using paper and pencil" (ibid., pp. 292-3). In other words, the function *F* can be acquired by different constituent structures, such as different physical constitution, hence *F* is multiply realized.

The MRT is formulated as follows: If mental states are thought to be similar to the states of a PA, as Putnam thinks, then what applies to the computer example above applies to psychological states as well. Therefore, one can argue that psychological properties are multiply realizable.

The conception of multiple realizability based on the notion of functional isomorphism is appropriate for the following example: Suppose that we find what the function of the property of believing that it is raining is. This requires a specification of some kind of a causal network whose one element is the belief at hand. The belief that it is raining is probably caused by seeing the cloudy weather and hearing the raindrops. And it probably causes behaviors such as staying indoors or putting on a raincoat. In addition to the perceptual causes and behavioral effects of this belief,

there are causal connections with other mental states. For example, it is connected to the belief that whenever it rains, if you are outdoors, you get wet. Or, it may cause the decision to cancel a plan for going on a picnic. The constituents of this causal network can be extended.

The important point in favor of the MRT which is grounded on the notion of functional isomorphism is that the same causal network can be acquired by means of different physical constitutions. That is, just as a computing machine can be realized either by electrical components or cogs and wheels, as Putnam suggests, a belief can be realized by either a human brain, or some electronic components. The point is that as long as the two systems (namely the human brain and the electronic system) are functionally isomorphic, they are realizing the same mental property, namely the property of believing such and such.

3.3.3 Evidence for the MRT

Empirical evidence in favor of the MRT comes from the intuitive idea that the same mental properties are realized by different organisms (Putnam, 1967; 1975; Block & Fodor, 1972). The classical example by Putnam is the multiple realizability of the property of being in pain, as a psychological property. Since, according to Putnam, pain is a functional state of a whole organism, instead of being a physical-chemical state of the brain, this functional state can hold for different species. Moreover, this account applies not only to the actual species on the earth, but also to the possible organisms that one might discover. As Putnam argues, one might meet Martians and discover that they are functionally isomorphic to human beings, but realizing the property of being in pain in a radically different physical-chemical way than human beings (1975, p. 293).

In a similar fashion, Block and Fodor (1972) argue that the MRT can be defended from a Darwinian point of view. According to *convergence*, which is a Darwinian notion referring to the phenomenon that different lineages independently develop similar traits to solve same problems, it can be argued that similar (or same) psychological traits have been developed by different organisms which have physiologically different structures. In other words, "[p]sychological similarities may often reflect convergent environmental selection rather than underlying physiological similarities" (ibid., p. 161).

Furthermore, it is suggested that the MRT could be held considering the neurological evidence that even the same organism can realize the very same psychological property by means of different brain structures. Block and Fodor state that some psychological functions can be realized by various operations of the brain (ibid., p. 160). Any evidence showing that even the human brain can realize the same psychological properties in different ways would provide support for the MRT.

That the human brain is capable of realizing the same mental property in different ways was put forward by Block and Fodor (1972) as an empirical consideration supporting the MRT. It is known that when some parts of the brain lose their function due to lesions, some other parts of the brain take over the function. This results from the fact that the functional architecture of the brain is dynamic and open to modification by new experiences, and this is known as the plasticity of the brain (or neural plasticity) (Gilbert, 1999, p. 598)[4].

There are two paradigm cases of neural plasticity[5]. First, in the cases of synaptic plasticity, as an outcome of past association of individual neurons, particular synaptic connections are facilitated. And, such connections are inhibited due to the lack of such associations. If synaptic connections could be formed and deformed by experience, it could be argued that different synaptic relations can be responsible for the realization of the same psychological function, hence, an example of multiple realizability could be given. Second, in the cases of cortical functional plasticity, a psychological function is realized by an area of the brain at a given time and by a different area at another time. Such cases occur when sensory input is increased or reduced due to amputation, and the size or the shape of the corresponding cortical areas change consequently. As the shape or the size of the brain area which is responsible for the realization of a given psychological function changes across time, this could be given as an example of multiple realizability.

4 That the human brain has plasticity is an empirical hypothesis. It is not the scope of this thesis to provide a conclusive argument for neural plasticity, or an empirical evaluation of that theory.

5 See Polger (2009).

Examples for such kind of multiple realizations also include the cases in which the neural mechanisms for psychological functions, such as face recognition, change through development (Johnson, 2005). In infants of 2 months, face recognition is processed by the sub-cortical system which is called "con-spec". In older infants, children and adults, the same function is processed by the mature cortical system. This can be taken as an example of multiple realizations, because the same function is realized by different systems of the brain.

3.4 Anti-Reductionism

3.4.1 Special Sciences

What does the MRT tell us with respect to reduction? If the MRT is true, then the mental properties are not identical with the properties of the brain. Instead, they are realized by the properties of the brain in human beings, but they are also realized, or realizable, by different properties in different systems. Thus, the first problem is that, there are no identity statements between psychological properties and physical properties, and according to the Nagel model of reduction, without these bridge principles, reduction is not available.

One might say that although there are no identity statements, once can have bridge laws between psychological and physical properties in the form of disjunctive laws. That is, instead of statements like $S \leftrightarrow P$, one can have statements in the form of $S \leftrightarrow P_1 \vee P_2 \ldots \vee P_n$, where each P refers to one kind of realizer of a psychological properties. Fodor's (1974) response to this proposal introduces the notion of a special science, and in what follows, first this notion will be explained, and second, Fodor's argument will be introduced.

The idea that mental properties, which are scientific properties (that is, the properties of cognitive science), are multiply realizable is strongly related to another idea, namely that cognitive science and other sciences whose properties are multiply realizable are special sciences. A special science is a higher-level science which uses its own taxonomy and which makes uses of scientific properties which are realized by properties of lower-level, or basic sciences, such as physics.

The notion of a special science constitutes the backbone of the one of the arguments against Reductive Physicalism. In order to introduce how this argument is

related to the MRT, the special science arguments of Fodor (1974; 1997) and Putnam (1975) should be discussed.

It should be noted that a rejection of Reductive Physicalism is compatible with what is called Token Physicalism. As Fodor defines it, the latter "is simply the claim that all the events that the sciences talk about are physical events" (1974, p. 100). Therefore, Token Physicalism is not a strong thesis as what might be called Type Physicalism, because, according to Type Physicalism, the properties of special sciences are physical properties themselves. Token Physicalism does not entail Type Physicalism. That is, that the tokens of events that are explained by the special sciences are identical to tokens of events that are explained by physics does not mean that types of events of special sciences are identical to types of events of physics.

It would be fair to admit that Token Physicalism does not entail reductive physicalism. But Fodor also has the burden of proof to show that special sciences are not reduced to physics at all, and this has to be related to the MRT in some way. For this, Fodor uses another example, this time from a real scientific practice, namely economics, which is a special science itself.

Consider Gresham's law, which can be stated informally as *bad money drives out good*, as a law of economics as a special science. Granted that Token Physicalism is true, one should be justified in believing that a given event *E,* which is about monetary exchanges, is a physical event, and *E* has a physical description *D* in the vocabulary of physics. Moreover, *E* falls under the laws of physics by means of *D*. However, there are a large number of events in addition to *E* which realize the condition of monetary exchange. Thus, any description *D** which would attempt to describe all events realizing monetary exchange would be a disjunction of an indefinite number of physical events (ibid., p. 103). Monetary exchange is, hence, multiply realized.

According to Fodor, that properties of special sciences are multiply realizable poses another problem for reductive physicalism, and this can be explained in a formal way. The following is a summary of Fodor's (ibid., pp. 108-12) argument with a slightly different formalism.

Consider a Nagel model reduction for a special science law. Suppose that

(1) $S \rightarrow R$

is a law of special science. *S* and *R* are predicates of the special science, therefore they are multiply realizable by physical properties. Thus, we have two bridge laws:

(2a) $S \leftrightarrow P_1 \vee P_2 \ldots \vee P_n$;

(2b) $R \leftrightarrow Q_1 \vee Q_2 \ldots \vee Q_m$.

Moreover, there are statements of physics which connect each *P* predicate with a *Q* predicate, some of which are:

(3a) $P_1 \rightarrow Q_2$;

(3b) $P_2 \rightarrow Q_4$;

(3c) $P_n \rightarrow Q_m$, and etc.

A statement at the level of physics must be like this:

(4) $(P_1 \vee P_2 \ldots \vee P_n) \rightarrow (Q_1 \vee Q_2 \ldots \vee Q_m)$

Recall the Nagel model's criterion for reducibility: If (1) can be derived from (2a), (2b) and (4), then (1) could be reduced to (4). This seems to be an appropriate case of reduction, because such a derivation is valid.

The problem is that, although the derivation is valid, (4) does not have any scientific significance at all, meaning that, the reduction of (1) to a statement of a basic science is not successful from a scientific point of view. The reason why this reduction is not successful is that (4) itself is not a law, despite the fact that each (3) statement is a law. Fodor illustrates this with the following analogy:

> I think, for example, that it is a law that the irradiation of green plants by sunlight causes carbohydrate synthesis, and I think that it is a law that friction causes heat, but I do not think that it is a law that (either the irradiation green plants by sunlight or friction) causes (either carbohydrate synthesis or heat). Correspondingly, I doubt that 'is either carbohydrate synthesis or heat' is plausibly taken to be a natural kind predicate (ibid., p. 109).

In other words, Fodor argues that statements which make use of the disjunction of scientific predicates are not laws, although the statements that use the basic predicates in the disjunctions may be laws.

3.4.2 Physical Explanations and Cognitive Science

So far, Fodor's argument that special sciences cannot be reduced to basic sciences has been introduced. Putnam (1975) developed another argument from special sciences for showing the irrelevance of physical explanations to special sciences. Since cognitive science is a special science in the sense that has been introduced above, Putnam's argument boils down to the claim that physical explanations are irrelevant to cognitive science explanations.

As mental properties are realized by physical properties, some other macro-level properties of physical objects (e.g. shape of an artifact) are realized by micro-level physical properties (e.g. micro-physical properties of the atifact). Putnam (ibid., pp. 295-7) gives a thought experiment by means of making an analogy between a psychological explanation and an explanation about an event related to the shape of an artifact. The thought experiment goes as follows:

Putnam asks us to conceive of a wooden board which has two holes on it, first one being circular and second one being square. Both the diameter of the circular hole and one side of the square hole are 1 inch. A cubic peg whose each side is 15/16ths of an inch will not pass through the circular hole, but it will pass through the square hole. Now, the question is what the explanation of this fact is.

As Putnam suggests, there are, at least, two possible explanations of this fact. The first one is the micro-physical explanation: The peg, as a physical object, is a cloud of atoms. So, it has a description which makes use of all the details about its atomic structure. The same is true for the board as well. It is, in principle, possible to compute all possible trajectories of the peg and to deduce from the laws of particle mechanics or quantum electrodynamics that it cannot pass from one hole, but it can pass from another hole.

Unlike this complicated and detailed explanation, the second alternative appeals to the macro-physical story: Both the peg and the board are rigid, and, as a matter of a geometrical fact, one hole is bigger than the surface of the peg while the other hole is not. Thus, the peg passes from the former, not the latter. As Putnam argues, the second explanation is superior to the first one, because, it is more general. The reason why it is more general is that the properties that are mentioned in the second explanation are multiply realizable. That is, each of the properties *being rigid, being smaller than the surface of the peg,* and etc. can be realized by a variety of

different atomic structures. In other words, the second explanation "is a correct explanation whether the peg consists of molecules, or continuous rigid substance, or whatever" (ibid., p. 296).

Thus, Putnam argues that the explanations appealing to the micro-physical features of the peg and the board are irrelevant. Even, he argues that "in terms of the purposes for which we use the notion of an explanation, … [micro-physical explanation] is not an explanation" (ibid., emphasis deleted). He suggests that if we are still willing to say that it is an explanation, we have to admit that it is a methodologically "terrible" explanation (ibid.): It cannot apply to other systems of similar macro-physical structures, and it is a demand of scientific method that the explanations be applicable to other systems. He argues that the second explanation applies to a large number of facts, whereas the first one can apply only to this particular event, and fails to be a general explanation (ibid., p. 297).

Needless to say, Putnam argues that the same is true for the class of psychological explanations. The physical explanations would be irrelevant, or terrible, because they would not have the power of being applicable to the class of interesting systems. That is, since psychological properties are multiply realizable, the explanations appealing to the physical features would just be coincidently true, thus not general. To conclude, the psychological explanations do not hold in virtue of the physical substance that our minds are made of: "We could be made of Swiss cheese, and it wouldn't matter" (ibid., p 291).

3.5 Concluding Remarks

In this chapter, the CTM and the MRT were introduced. The issues were discussed within the context of the development of the cognitive science perspective, which entails a computational approach to the mind. It has been shown that the defenders of the MRT argue that it shows that Reductive Physicalism is false. The anti-reductionist claims from the MRT which have been introduced in this chapter can be divided into three groups: The first claim is that a given psychological property cannot be identified with a given physical type, because, as the MRT suggests, there are multiple physical ways of realizing that psychological property. The second claim is that bridge laws, because of multiple realizability, are disjunctive, and this renders reductionism as an unattainable scientific ideal. And

finally, it is argued that physical explanations about psychological phenomena are not as good as psychological explanations, because, psychological explanations are more general than physical ones due to the availability of multiple realizations. The second and the third arguments for Anti-reductionism also constitute the argument for special sciences. In the following chapter, objections to the special sciences argument will be discussed. Arguments against the MRT itself are saved for Chapter 5.

CHAPTER 4

RESPONSES TO ANTI-REDUCTIONISM

4.1 Introduction

In this chapter, some arguments against the anti-reductionist implications of the MRT will be reviewed. First, Kim's argument will be introduced. He tries to show that the real implication of the MRT is that mental properties, namely the properties of cognitive science, are not scientific kinds. Second, Batterman's account which shows that the MRT does not demonstrate the irrelevance of physical explanations will be introduced. Third, Sober's argument against Fodor's and Putnam's Anti-reductionism will be discussed. Finally, a rejoinder to Kim will be provided. For the rejoinder, a revision of the MRT will be proposed..

4.2 Kim's Argument against Special Sciences

Kim (1999) questions the scientific status of the mental properties. He starts his analysis by making an analogy between *jade* and *pain*. He says that *jade* was formerly considered to be a mineral kind, but it is not a mineral kind anymore, because it is discovered that it comprises two different minerals which have different molecular structures, *jadeite* and *nephrite*.

Then the following statement was formerly taken to be a law:

(L) *Jade* is green.

It was assumed that (L) has been confirmed by a vast amount of observations. Now that it is discovered that jade is not a mineral kind, it is not controversial to say that (L) is not a law. Rather, it is the conjunction of these two laws:

(L1) *Jadeite* is green.

(L2) *Nephrite* is green.

The reason why (L) is not a law is that it does not pass the test of 'projectibility', meaning that it does not have "the ability to be confirmed by observation of 'positive instances'" (ibid., p. 520). It does not pass this test, because, the following scenario is conceivable: We can re-examine the past observations and find out that all the positive instances of (L), that is, all the millions of observed samples of green *jade*, turn out to have been samples of *jadeite*, and none of *nephrite*! If this should happen, we clearly would not, and should not, continue to think of (L) as well confirmed" (ibid., emphasis in the original).

Kim anticipates a possible objection: It might be objected that this case can be applied to anything, for example, to emeralds. We might say that being *emerald* is the disjunction of being *African emerald* and being *non-African emerald*. So, does this mean that *emerald* is not a kind? Kim says that the case is not analogous to the *jadeite-nephrite* case. In the case of *emerald*, we do not have heterogeneously disjunctive kinds, that is, all African and non-African emeralds behave in similar ways. However, "samples of *jadeite* and samples of *nephrite* do not exhibit an appropriate "similarity" with respect to each other to warrant inductive projections from the observed samples of *jadeite* to unobserved samples of *nephrite*" (ibid., p. 521, emphasis in the original).

Kim takes it that *jade* is multiply realized by *jadeite* and *nephrite*, and this is why *jade* is not a nomic[6] kind. Then, what is the relevance of the jadeite-nephrite case to the debate between Reductionism and Anti-reductionism? Kim argues that the case of *pain* is analogous to jadeite-nephrite case, because the property of being in *pain* is also multiply realizable in different organisms (human beings, reptiles and Martians) by different neurological properties, namely by N_h, N_r, and N_M. If *jade* and *pain* are analogous with respect to their disjunctive characters, then it follows that pain also does not pass the test of projectibility. That is, generalizations about human pains do not apply to reptilian pains.

6 Nomic: Scientific, in this context. A nomic kind is a kind which appears in scientific laws.

What happens, then, if pain does not pass the test of projectibility? According to Kim, the consequence is that *pain* is not a nomic kind. This shows that pain "is not the sort of property in terms of which laws can be formulated; and "pain" is not a predicate that can enter into a scientific theory that seeks to formulate causal laws and causal explanations" (ibid., pp. 522-3). And if the MRT is true for all mental properties, then, it follows that all mental properties are nonnomic.

This shows that cognitive science is not a science *per se*, simply because the properties of cognitive science, namely the mental properties, are not scientific kinds. Kim does not mean that psychology is a pseudo-science like alchemy: "The crucial difference, from the metaphysical point of view, is that psychology has physical realizations, but alchemy does not" (ibid., p. 524).

That psychological properties have physical realizations also makes it possible that psychological properties are reduced locally. By local reduction, Kim means the reduction of human pain to N_h, reptilian pain to N_r, and Martian pain to N_M, etc. Indeed, he says that, "multiple local reductions, rather than global reductions, are the rule, even in areas in which we standardly suppose reductions are possible" (ibid.). Moreover, the availability of local reductions also makes it possible that local psychologies can be studied by means of investigating the physical properties of the local reduction bases.

Kim labels this kind of reduction as "functional reduction" (2002; 2005). He summarizes how functional reduction can be managed as follows: (i) The property to be reduced is specified by means of providing a functional description. (ii) The realizers of the functional property, that is, the physical properties that fit the causal role, are found. (iii) An explanatory theory is provided which gives an account of how the physical realizer plays the functional role that was specified.

It seems that there are, at least, two advantages of this account of reduction: First, it shows the Reductive Physicalist a way of negotiating with the functionalist without giving up reduction to physical types. Secondly, it emphasizes a pragmatic point in the sense that it lets us focus on mental properties of individuals or species, rather than give a universal account of a psychological property. However, this latter point could also be seen as a disadvantage, as science seeks generality.

Thus, Kim's argument can be taken against Fodor's (1974) Anti-reductionism that was introduced in the previous chapter. Recall that Fodor's claim was that, because of the availability of multiple realizations, laws of cognitive science cannot be reduced to laws of a basic science such as physics. However, Kim suggests that, because of the availability of multiple realizations, the scientific status of mental properties is in question.

4.3 Universality and Physical Explanations

A response to the Anti-Reductionist argument comes from Batterman (2000), but his argument can also be interpreted as a response to Kim's argument. Batterman makes use of Block's "Disney Principle", according to which "laws of nature impose constraints on ways of making something that satisfies a certain description" (Block, 1997, p. 120). In Walt Disney movies, anything, regardless of its structure, can talk, think and etc. But, obviously, this is not the case in reality. As Block argues, you have to be more than, say, a teacup to have a mental life. Block states that it would "be amazing if laws of nature imposed *no constraints at all* on what can think or be conscious" (ibid.). Here, he implies that there is the intuitively plausible requirement that there be similarity between different realizers of psychological properties. Batterman admits that there has to be a similarity as such, and he tries to provide a physical explanation for this similarity, by means of using a technical term from the field of physics, namely the notion of "universality".

Scientific reasoning, as Batterman suggests, demands the explanation of behaviors of a system. There are dominant parameters which are important for the specification of a system's behavior. One demarcates such parameters from other several parameters which could be interpreted as details yet are still important for a *full* description of the system. That there are such several parameters which could be interpreted as irrelevant for a general description is responsible for the reproducibility of experiments.

The dominant parameters are assumed to be universal by physicists. A straightforward example is the observed behavior of pendulums. Pendulums "all have periods (for small oscillations) that are directly proportional to the square root of the length of the rod from which the bob is hanging" (Batterman, 2000, p. 120). This is true for all pendulums of different colors, sizes, etc. Thus, the universal

structure is the mathematically defined equation, whereas color parameters are just the "noise" to the explanation of the behavior of a pendulum.

Batterman suggest that multiple realizability is just a version of universality. Consider the example of being in pain as a psychological property. There must be a description of this property which is universal in all pains, whereas some of the (presumably physical) details will be taken to be the noise in the explanation. At this point, his claim seems to be against Reductive Physicalism, because here, Batterman echoes the anti-reductionists' concerns about the irrelevance of the detailed micro-level descriptions. However, the crucial point in Batterman's account is that he does not say that all the physical details are irrelevant to the explanation of the behavior. What is universal can be still a physical property, not a functional one. He says that philosophers usually "ignore the possibility that there are physical reasons why the details of the makeup of the individual realizers may be largely irrelevant for the upper level behavior of the system" (ibid., p. 124, emphasis deleted).

Now, at this point, Batterman's account can also be interpreted as it is against Anti-Reductionism, for he says that the dominant parameters, which are physical, are relevant for a general explanation. Because of the distinction between dominant and irrelevant physical parameters, this does not mean that every physical detail is important for a good explanation. Therefore, Batterman's suggestion seems to be a moderate position between Reductive physicalism and Anti-reductionism.

4.4 Sober on the Relevance of Physical Explanations

Sober (1999) provides an argument against the argument from special sciences. Recall that Putnam argues that, because of multiple realizability, the physical explanations of psychological phenomena are not as general as the psychological explanations, hence one should dispense with the physical explanations. Putnam (1975) uses a thought experiment about pegs and boards to make his argument more clear (see 3.4 for Putnam's thought experiment in detail).

First, Sober suggests that although the micro-details are not interesting for Putnam, they may be interesting for others. That something has more details than what we want to hear does not mean that those details are irrelevant. A genuine case of an irrelevant detail would be as follows: Suppose that the peg which does not go through the first hole is green. An explanation which would make use of the colour

would be irrelevant because the colour has no causal role in the peg-board story. The colour of the peg is causally irrelevant. Sober suggests that the details that Putnam includes among the irrelevant ones are not objectively irrelevant. They are irrelevant because we do not want to hear about them. However, details such as colour of the peg are objectively irrelevant, because they have no causal role in the story. Therefore, Sober's point is that anything counts as an explanation as long as it explains the *causes* of the phenomenon under consideration. This also shows that, in some context, a micro-physical explanation may be better than a psychological explanation.

Second, Sober (1999) introduces another argument against the anti-reductionist argument from the special sciences perspective. Here, he is mainly against the general anti-reductionist framework about the irrelevance of the physical information. This argument is mainly based on Sober's supposition that laws of special sciences are probabilistic (ibid., p. 555). Here, his argument revolves around a law stated as follows: *Smoking causes lung cancer*. This statement is actually a probabilistic statement, because, what it says is that smoking makes lung cancer highly probable.

Now, since *smoking* and *lung cancer* are special science properties[7], these are supposed to be multiply realizable. Suppose that *smoking* is realized only by two carcinogenic ingredients: A_1 and A_2. Sober states that the realization relations between types of A (that is, A_1 or A_2) and *smoking* are also probabilistic. The probability of A_1 realizing the higher-level properties of nicotine is not necessarily equal to the probability of A_2 realizing such higher-level properties. That is, "[i]t can easily turn out that one of these ingredients is more carcinogenic than the other" (ibid.). The same is true for *lung cancer*: Suppose that *lung cancer* is only realized by B_1 and B_2 which are different types of tumors. It may turn out that the probability of having lung cancer when one has one type of tumor is greater than having another type of tumor. Now, there are two different probabilistic relations with respect to the

7 The question whether there is such an actual special science which has *smoking* and *lung cancer* as its properties is not important here. As long as *smoking causes cancer* is some kind of a law, it is a law of a special science.

law that smoking causes lung cancer: First, the law itself is a probabilistic law. Second, the relations of the properties with their realizers are probabilistic. Sober argues that, from these considerations, one can show how micro-level explanations may be better than general explanations.

In the case of a person who is suffering from lung cancer, what would the explanation of this disease be in this case? According to one explanation, we can say that the person has cancer because he smoked cigarettes. On the other hand, according to another explanation, we can say that he has cancer because he inhaled the ingredient A_1. According to Putnam's argument, the second one would not be an explanation, because since there is multiple realizability, another ingredient, namely A_2, could have caused cancer. Sober, however, argues that second explanations has its own virtues. Remember that *lung cancer* is also multiply realized by tumors B_1 and B_2. And, also note that A_1 and A_2 differ from each other with respect to their probabilities causing B_1 and B_2, and there are different treatments of B_1 and B_2. Therefore, knowing what kind of ingredient (A_1 or A_2) is crucial for specifying the treatment of lung cancer. As Sober suggests, "it is a good thing for cancer research that the multiple realizability argument has not won the hearts of oncologists" (ibid., p. 556).

Thus, Sober finds a proper way of showing how micro-level explanations may be better than macro-level explanations. In some cases, the micro details are more important than the general picture.

4.5 Challenging Cognitive Science

4.5.1 The Challenge

As discussed in 4.2, Kim argues that, because of the MRT, mental properties are not projectible. In other words, statements about psychological properties cannot be confirmed by positive instances. Take the classical example of being in pain: Since *pain* is multiply realizable, different realizers of it might be vastly different from each other. Therefore, as Kim states, the statements about human pains do not apply to statements about human pains. And this allegedly shows that psychological properties are not scientific kinds.

One can take Kim's argument as a challenge to any cognitive scientist, because cognitive science is supposed to be a science talking about scientific

properties, and the scientific properties it is supposed to be dealing with are mental properties. If Kim is right, the cognitive scientist will have to dispense with mental properties in order to keep its scientific status. What the cognitive scientist has to do here is to refute Kim's point. The following is an attempt to this end.

4.5.2 Responding to the Challenge: A Revision

Kim's argument starts with an assumption: Mental states are multiply realizable by *vastly different* physical states. Then, it is concluded that such mental states cannot be nomic if their physical realizers are vastly different from each other. Here, I propose that if a slightly different version of the MRT is assumed for the sake of argument, the conclusion Kim derives will not follow. The slightly revised version I propose is, at the beginning, silent about the extent to which realizers are different from each other. That is, we should only propose that mental states are multiply realizable. That is, at a first attempt, we should not consider the possibility of, say, Martians which could have the same mental states as human beings.

Since Kim's argument as a challenge to cognitive science is being examined, why should one bother dealing with organisms which are vastly different than human beings? Why should a cognitive scientist deal with the Martian pain? The primary subject of cognitive science should be, more than anything else, the mental life of human beings[8]. And, as far as the scientific progress shows, the physical structures of human beings are significantly similar to each other in the sense that, all else being equal, generalizations about a given human being apply to another human being.

Here, it is not difficult to foresee this inevitable objection: Cognitive scientists do not only study the human psychology; there are studies about other species. Even, advancements in Artificial Intelligence (AI) could be attributed to the developments in cognitive science. However, these do not refute the point made here. Because when studying the psychology of a given species, it should be noted that the

8 It could be argued that cognitive science deals with "intelligence", not only human intelligence. However, the discussion of Kim's challenge requires the discussion of "mental states". Although it could be argued that the mental states of a possible Martian and a human being are both indications of intelligence in both organisms, it should be clear that such indication of intelligence does not entail that these mental states are the states of the same mental kind.

evolutionary connections are likely to tell the cognitive scientist something about human cognition.

As for the AI case, it should be noted that there are two sorts of studies: First, there are researchers who are willing to develop some machines which could perform some tasks. Second, there are researchers who are developing some machines, or models, in order to understand the nature of cognitive capacities of human beings. The first case is not relevant to our purposes here, because, as long as the aim is to develop a machine which could perform a given task, the machine is not supposed to perform the task *in our way*. That is, it could still be argued that it is performing the task in an *intelligent* way; however, its way and our way are presumably vastly different from each other[9]. The second case is relevant; but since, in most cases, the purpose in modeling cognition should be to understand the human mental life in a more comprehensive way, one would not develop a model whose constitution is vastly different than a human being. Here, it might be objected that the machine will be different than a human being in so far it is not built out of flesh and bones. But the point is that, regardless of the physical substance, the operations that are carried out in the machine should be significantly similar to the operations that are carried out in the human brain[10].

The revision I mentioned is not an *ad hoc* move. It is well-motivated by the practice of cognitive science. The examples of multiple realizations abound. But let us take one example from the field of developmental cognitive science. As mentioned in 3.3.3, through development, the brain realizes the psychological function of face recognition by different systems (Johnson, 2005). Now, suppose also that the MRT is true in a wider sense that it is possible that an organism which is dramatically different from human beings (e.g. a Martian) realizes the same function in different ways.

Suppose that a cognitive scientist is studying face recognition by means of analyzing several organisms that are capable of recognizing faces. Take these

---

9 See footnote 7.

10 See 4.3 for the discussion of Block's (1997) Disney Principle for the intuitively plausible idea that things are likely to resemble each other in order to satisfy a functional description.

organisms as O_1 (an infant of 2 months), O_2 (an older infant), O_3 (an adult human being) and O_4 (a Martian). Although face recognition in each case is multiple realized, it is likely that there is some continuity between O_1 and O_2, O_2 and O_3. However, such continuity is less likely to be found between O_4 and any of the other three organisms. This, at first glance, seems to vindicate Kim's challenge to cognitive science, as it seems that there is the possibility of a very different realizer of the same function. However, the scientist is situated in a context, and the developmental cognitive scientist is situated in a context where, mostly, the differences in the brain in different stages of development matter. That is, the developmental cognitive scientist is, albeit contingently, is supposed to neglect the possibility of a Martian as such.

With these considerations in mind, Kim's argument will fail to be a challenge to the cognitive scientist. I have tried to show that the assumption Kim makes in the first step of his argument could be revised, and after such a revision, his conclusion does not follow. His argument might still be a challenge to a "cognitive science" whose aim is to study every possible organism (and machine) that realizes some higher-level properties, but one would not expect to find such a science, as this is practically an impossible ideal.

I propose that such a revision is also likely to suggest that the claims both of Batterman (2000) and Sober (1999) are true in the sense that the physical information in cognitive science would be explanatory. That is, if the revision is held, one would expect to see more resemblance between different realizers of mental properties than Putnam has thought. Therefore, the physical explanation in a given instance of a mental property is likely to provide information about the next possible instance; hence physical explanations would be relevant.

Therefore, the revision I propose here is important for two independent reasons: First, it shows that we could dispense with the Anti-reductionist worries about the irrelevance of the physical explanations. Second, a possible challenge to the scientific status of cognitive science is shown to fail.

4.6 Concluding Remarks

In this chapter, several arguments against the anti-reductionist argument from the special sciences perspective have been introduced. First, it has been shown that

Kim argues that if the MRT is held, the scientific status of mental properties, hence cognitive science, is in danger. Second, it has been argued that Batterman's moderate position between Reductionism and Anti-reductionism can be favored, since it shows that both camps have something right. The important thing that Batterman shows is that what the realizers of a given psychological property share in common may be something physical. This understanding leaves room for physical explanations in the cognitive science, because, if Batterman is right, physical information is as important as the information from more general psychological explanations. Third, Sober's similar conclusion that physical explanations have their own virtues has been introduced. Finally, the challenge Kim sets has been taken. For this, I have proposed a revision of the MRT, such that a claim of multiple realization should be silent about the extent to which realizers are physically different from each other. I have argued that the practice of cognitive science and intuitions about satisfying a functional description imply that the different realizers of a mental property are likely to resemble each other. If that is true, first, a conclusion similar to the claims of Sober and Batterman could be held. Second, Kim's challenge that cognitive science deals with mental properties which do not resemble each other in physical ways fails.

CHAPTER 5

AGAINST THE MULTIPLE REALIZABILITY THESIS

5.1 Introduction

Chapter 3 explained the Multiple Realizability Thesis (MRT) in detail, and discussed the arguments that followed from it. Chapter 4 has discussed several objections to the anti-reductionist conclusion that seem to be implied by the MRT. The present chapter introduces some arguments which are against the MRT itself.

The first response I will discuss makes use of the notion of "qualia". The second argument to be introduced against the MRT is developed by Bechtel and Mundale (1999), and it is to the point that the practice of neuroscience implies that both the MRT and the implications that are drawn from it are false. The third argument to be introduced belongs to Shapiro (2000; 2004), and it mainly rests on an original interpretation of the notion of multiple realization. Then, a rejoinder to Bechtel and Mundale, and Shapiro will be provided. Finally, the consequences of the Embodied Cognition research program for the MRT will be explored.

5.2 Argument from "Qualia"

Zangwill (1992) suggests that one important tool for the demonstration of the MRT is imagination. The argument from imagination, as Zangwill suggests, tries to establish the MRT from such facts that we can imagine a possible organism (such as a Martian) which does not have a nervous system similar to a human being, yet feels pain. According to Zangwill, this is not conclusive. The reason is that the imagination might just support an argument which is against the MRT as well: We can imagine two functionally isomorphic organisms one of which feels pain while

the other lacks the qualitative aspect of feeling pain[11]. So, according to this argument from imagination, we can consider, or imagine, two functionally isomorphic organisms which, however, differ from each other with respect to the qualitative natures of their psychological states. Now, the problem is that this supposition would show that the MRT is false, because, under the assumption that mental states do indeed have a phenomenal, qualitative character, having a functional description is not sufficient for having a mental state.

It should be noted that this sort of an argument is not originally developed by Zangwill. This argument was originally developed by Thomas Nagel (1974), and a different version of it was provided by Jackson (1982; 1986). However, these arguments were not developed against the MRT, instead, they were against Reductive Physicalism. It was argued that a physical explanation of mental states would lack the qualitative aspects of mental states. But, as Zangwill suggests, the same point can be made against the MRT.

Nagel (1974) uses an example which shows that the complete information of physics cannot give us access to the subjective experience of a bat; in other words, no matter how much one is informed about the physical structure of a bat, one can never know *what it is like to be a bat*. This argument, as Nagel suggests, shows that a reductionist program would miss the qualitative nature of psychological states, no matter how successful it is in explaining the physical nature.

On the other hand, Jackson's (1986) objection is based on a thought experiment, so it may be taken as a better example for an *argument from imagination*. Suppose that there is a scientist called Mary who has lived in a black-and-white room, and was educated through black-and-white books, a black-and-white TV, and so on. Mary has never experienced colors other than black, white and shades of gray. As a super-scientist, she knows all the facts about physics, chemistry and neurophysiology. Then, she knows every single fact about what is going on in the brain of a person when she sees a red object. Jackson maintains that if the identity claim of Reductive Physicalism is true, then Mary is supposed to know all there is to know. However, Jackson argues that when Mary is released and sees a ripe tomato,

11 The term “qualia” is used for to the qualitative aspects of mental states.

she will learn something, which in turn means that, prior to her release, she did not know everything. Therefore, as Jackson suggests, identity claim of Reductive Physicalism is false. That is, although she knows everything about the physical processes in the brain when a person sees red, prior to her release from the black-and-white room, she cannot know the qualitative aspect of seeing red.

These are two arguments from imagination against reductionism. One can see how these arguments may work against the MRT as well. In establishing the MRT, the first step is defining the states of an organism in a functional way. But these arguments would show that the functional description of an organism must leave out qualitative aspects of experience. Based on a very similar motivation, Zangwill thinks that an argument from imagination would show that a functional description of a mental state lacks the qualitative aspect of mental states. And, he suggests that this is important, because the very same tool that the MRT endorses, namely imagination, refutes the MRT itself.

However, Zangwill's claim can be criticized as follows: The MRT is also defended by arguments which are not from imagination. That is, without imagining a Martian, it can be argued that two organisms, even two human beings, are in the same mental states without being in the same physical states. One might still argue that such argument appeals to the functional description of a mental state, and this definition leaves out qualia, yet, this is completely a different point. That is, the point is not that the same tool for establishing the MRT, namely imagination, could be used against the MRT; without resort to imagination, the MRT is still defended.

5.3 Bechtel & Mundale: The Practice of Neuroscience

One prominent objection against the MRT comes from Bechtel & Mundale (1999). Their account is mainly based on the emphasis on the argument that the implications of the MRT are not consistent with the practice of neuroscience. In this section, a reconstruction of their criticism of the MRT will be reviewed.

5.3.1 Psychological and Neural Properties

The first argument Bechtel and Mundale provide emphasizes how psychological and neural properties have been connected to each other within the

practice of neuroscience[12]. As Bechtel and Mundale argue, as far as the history of neuroscience shows, "the appeal to function, especially psychological function, is an *essential* part of both the project and its tools" (ibid., p. 177). That is, regardless of the truth values of the theories provided by neuroscientists, the practice itself is linked with psychological functions.

There are several examples of studies on how psychological functions and brain regions are conceptually linked. For instance, there is the research program of Gall, which resulted in the development of the proposal that there are brain locations for 27 different psychological functions including memory of words, memory of things, mathematical ability, love of offspring etc. (ibid. pp. 183-4). The significant thing with respect to the present discussion is that Gall distinguished between cranial areas according their psychological functions with the assumption that the function was important for the specification of the region. As Bechtel and Mundale suggest, many of what Gall has said were incorrect as the recent theories of brain show, but he set a strategy to study brain (ibid.).

Following Gall's agenda, researchers have developed ways to find regions in the brain for various psychological functions. One method was based on the idea that damage to a region in the brain would result in the deficit of its psychological function. Thus, Broca in 1861 found the relation of "damage to the third convolution of the left frontal lobe with less of articulate speech" (ibid.). Here, Bechtel and Mundale draw an early conclusion that the availability of deficiency cases rejects the MRT on some grounds. They argue that, when scientists work on these deficits, they assume that there is no multiple realization within the brains of different human beings, and they predict that damage to an area of a brain in any subject, all else being equal, result in a deficit of a psychological function being performed by that area in other (undamaged) brains. Also, since there are cross-species comparisons in

12 Therefore, it may be said that this argument is not against the MRT *per se*, but against one conclusion that the supporters of MRT draw from it, namely that physical properties are irrelevant to psychological properties (the argument about the irrelevance of physical properties are discussed in 3.4). Although arguments against the reductionist implications of MRT (instead of the MRT itself) are introduced in Chapter 4, and it is this chapter's task to discuss the arguments against the MRT itself; however, the present discussion is relevant to a wider account against the MRT.

these cases, they argue that researchers are also implicitly rejecting the MRT about different species (ibid.)[13].

In addition to the studies which show how damage to a region leads to deficits in functions, there are alternative approaches such as using "mild electrical stimulation to excite the suspected area in normally functioning organisms and determine what psychological activity is thereby increased" (ibid., p. 185). For example, Ferrier, as well as using deficit studies, relied on such stimulation studies and "argued that visual processes were localized in the angular gyrus … and the occipital lobe, while auditory processes … [were] localized in the temporal cortex" (ibid., p. 187).

It should be noted that, for the sake of the argument, it is not important whether these empirical claims about the brain are true or false. What matters is that they show, first, how function and the description in the physical level are interlinked, and second, that the MRT is implicitly rejected in these studies. Yet, the defender of the MRT would still say that, regardless of what the scientific practice of brain implies, the truth of the MRT lies on the possibility that the same function of the brain could be held by a brain, or any other system, which is in a different physical state.

5.3.2 Generalizing Across Species

One argument Bechtel and Mundale use against the MRT is that scientists assume commonalities of brain structures across species. As they argue, "it is the very *similarity* (or more precisely, *homology*) of brain structures which permits us to generalize across certain species" (ibid., p. 178, emphasis in the original). For this claim, they provide historical examples from neuroscience, and it should again be noted that since this argument is about the practice of neuroscience, the truth or falsity of the theories mentioned are not critically relevant to the purposes here[14].

In 1909, Broadmann developed a map of brain with 47 areas in the human cortex (ibid., p. 179). This study is also relevant to the discussion in 5.3.1 in the

13 Their argument about the availability of generalizations across species will be elaborated in 5.3.2.

14 See footnote 12.

sense that "Broadmann's goal of identifying different regions of the brain was to distinguish areas of the brain which performed different functions" (ibid.). But the important point about the generalizations across species comes from how he developed his map and what he stated afterwards. The theory he proposed was that "the cortex generally consisted of six different layers of neurons (which manifested themselves in stain preparations due to different cell types and densities)" (ibid.). He arrived at this conclusion by working by comparisons between different species. Working on different species, although he drew attention to differences between different species, his primary motivation was the assumption that there were commonalities in brain regions.

Here, it should be noted that this account of Bechtel and Mundale has to be revised to some extent. As Sungsu Kim (2002, pp. 608-9) argues, the cases of homology do not provide arguments either for or against MRT. Here, one needs to bear in mind the distinction between homology and homoplasy: Traits inherited from common ancestors are homologies, whereas, traits that arise out of distinct origins are homoplasies. If a mental property is realized in a human being and in a Martian, this would be a case of homoplasy. If such a mental property has the same physical realizer in the human being and the Martian, this would count as evidence against the MRT. Homologies, on the other hand, do not provide evidence for or against the MRT. Thus, the generalizations across species mentioned by Bechtel and Mundale (1999) do not show that the MRT is false. This, clearly, is a problem and renders their argument against the MRT inconclusive.

5.3.3 Differences in Human Brains

It is argued by the supporters of the MRT that brains of different human beings are different. Also, as the neural plasticity cases show, even the same brain could realize the same psychological function differently over times due to some lesions in some areas (for a discussion of this idea, see 3.3.3; for a critique, see 5.4.3). This allegedly demonstrates that it is even impossible to generalize across human brains.

Bechtel and Mundale think that this problem resolves itself from a practical point of view: It is known that there are differences between the brains of different individuals. But, researchers are using a common atlas for mapping different brains

onto it (ibid., p. 190). Also, positron emission tomography (PET) and magnetic resonance imaging (MRI), the noise ratio is low, and this enables researchers to average across brains. Such averaging cancels out, or neglects, the individual differences. Bechtel and Mundale argue that this practice of averaging shows that there is less variability than the MRT assumes (ibid.).

However, the same problem with their argument applies here as well. That is, the possibility that brains which are physically in different states realizing same mental states still creates intuitions which favor the MRT: Less variability is still variability.

5.3.4 Why Has the MRT Been Popular?

Finally, Bechtel and Mundale offer a diagnosis as to why the MRT has been so popular and widely accepted in the field of philosophy of mind and cognitive science. They suggest that there are two main reasons:

(i) Different criteria are used when identifying psychological properties and differentiating neural properties. Philosophers use a coarse grain when grouping mental states under the same kind, while using a fine grain for distinguishing between physical states as different kinds (ibid., p. 179). If one uses coarse grains to identify psychological properties on the one hand, and neural properties on the other hand, then one-to-one mappings between psychological and neural properties (in other words, identities) would hold. The same is true for fine grains as well.

(ii) In the discussions as to the plausibility of the MRT, there is a no well defined frame of reference for the criterion of distinguishing between kinds of states. For example, in some context, the mental states of a human being and of an octopus might be taken as similar, but in other context, these might be taken as different. The food seeking behavior of human beings and the octopus could be taken as different states if we are dealing with the particular way an organism seeks food. On the other hand, without defined frame of reference, both states could be taken as the same mental state (such as the state of feeling hunger) (ibid.).

This diagnosis seems to emphasize an important point. When saying that *X* is multiply realized by *Y* and *Z*, we assume that *X* is the same kind in two realizations, and *Y* and *Z* are different kinds. But, what do we appeal to, other than our intuitions, in order to argue that *X* is the same kind? That is, how do we know that pain is the same pain in a human being and an octopus? And, how do we know that the physical states which are responsible for the realizations of pain in a human being and an octopus are genuinely different than each other? These are intriguing questions, and they are similar to the worries that are raised by Shapiro. Thus, an evaluation of these questions will be provided after Shapiro's objections to the MRT are introduced.

5.4 Shapiro Against the MRT

The final response to the MRT to be discussed in this work belongs to Lawrence Shapiro (2000; 2004). First, he argues that Putnam's functional isomorphism argument does not establish the truth of the MRT (2000). Second, he analyzes the concept of multiple realization and shows that the supporter of the MRT has a dilemma according to this analysis (2000; 2004). Third, he discusses the examples for the MRT, and shows that, according to his formulation, these examples are not cases of multiple realization (2002; 2004**)**. In what follows, all these arguments will be discussed.

5.4.1 Functional Isomorphism Revisited

As introduced in 3.3.1, Putnam tries to establish the MRT from the notion of functional isomorphism (1975). Indeed, his way of arguing for the MRT rests on the definition of multiple realization by means of functional isomorphism. He suggests that two systems can be functionally isomorphic but can be made of different physical substances. Hence the functional kind in question is multiply realizable. Nevertheless, Shapiro claims that to be functionally isomorphic is not sufficient for being multiply realized.

The details of Shapiro's own definition of multiple realization will be given in 5.4.2, but for the purposes relevant here, it should just be noted that, according to Shapiro, in order for a kind to be multiply realized by different systems, the realized kinds have to be same in all realizations (2000, p. 638; 2004, p. 46).

Here is an example of functional isomorphism (2000, p. 638-9): Conceive of a simple pocket calculator designed for making simple arithmetic operations. Now,

conceive of a mainframe computer which can run very sophisticated programs in addition to calculating simple arithmetic operations. As Shapiro claims, although the calculator and the mainframe computer are functionally isomorphic with respect to simple calculations, they are not realizations of the same kind (ibid., p. 639).

Moreover, conceive of a functional isomorphism between a mouse trap and the marks on a paper (ibid.): Suppose that a symbol written on a paper, say "A", represents that the mouse is present, "B" represents that the mouse is in trap, and "C" represents that the mouse is dead, and so on. Shapiro argues that there is a isomorphism between the states of a real mousetrap and the symbols written on the paper. But, as he suggests, it is clear that such an isomorphism does not make the paper a mousetrap. The piece of paper does not realize the functional kind *mousetrap*. That is, being functionally isomorphic does not guarantee being the realization of the same kind. The same point applies to the AI examples for the MRT, but the elaboration of this argument is saved for 5.4.3, in which the alleged cases of multiple realization will be evaluated.

5.4.2 Conditions for Multiple Realization

Shapiro takes it that it is crucial for multiple realization that what is being realized is to be the same kind in all realizations (2000, p. 638). Yet, there is one further condition that has to be held for providing a case of multiple realization. The realizers must be different from each other (Shapiro, 2004, p. 46). That is, if two organisms' ways of realizing a psychological property are not different from each other, the psychological property in question is not multiply realized.

Thus, if physical properties X and Y multiply realize the functional kind M, these two must be held:

(i) X and Y are realizations of the *same* functional kind M (M in X and M in Y are functionally identical);

(ii) X and Y are *different* kinds (They realize M in distinct ways) (ibid.).

Both of these conditions bring complexities for the MRT. The first problem is related to the first condition. That is, how could one decide whether a function is identical in two different realizations? Consider the example of an analog and a digital watch. We intuitively think that analog and digital watches are clear examples of multiple realization. "It is the kind *watch* that is being multiply realized, with the

analog and digital watches each realizing *watch* in distinct ways … [But] why are analog and digital watches both realizations of the same thing?" (Shapiro, 2002). One might say that they are realizing the same function, namely showing the time. However, it might be said that the function of the digital watch include indicating tenths of seconds, whereas the function of the analog watch is indicating only the hour (ibid.). Thus, the problem is that, one can take the function of some artifact in one way or another.

The second problem for the MRT is linked to the second condition of being multiply realizable. Now, consider another artifact, namely the corkscrew. Consider the question "whether two corkscrews, alike in constitution and mechanism but distinct in color, count as alternative realizations of the kind *corkscrew*" (2000, pp. 643-4). According to the second condition above, these would not count as multiple realizations, because the way the function is obtained is not different in two realizations. "[D]ifferences in color make no difference to their performance as a corkscrew" (ibid., p. 644).

Here, one could compare this account with Batterman's account of multiple realization mentioned in 4.3. According to Batterman, pendulums of different colors and sizes are multiple realizations of the kind *pendulum* (2000). But here, in Shapiro's formulation, this is not the case, since the colors or the sizes of pendulums do not change the way the function of the pendulum is realized (2004, pp. 58-9). With this agenda, according to Shapiro, there is a way of distinguishing trivial cases and important cases of multiple realizability (ibid., p. 55). Pendulums or corkscrews in different colors are, thus, trivial cases of multiple realizability. But, the question is this: Are there genuine cases of multiple realizability?

5.4.3 Are there Genuine Multiple Realizations?

For this, now, let us leave the talk of artifacts, and consider the real scope of the MRT. Are there multiply realizable psychological properties? Let us go through a thought experiment allegedly showing the truth of the MRT (Shapiro, 2000, p. 645): Imagine that neurons of a brain is replaced with silicon chips one by one. Once all neurons are replaced, we end up having a silicon brain. It is suggested that the silicon brain would function the same way with the original brain, because during the replacement, the functional organization of the neurons are kept.

However, Shapiro asks: Why do we think that the physical difference between the original and the silicon brains is an interesting one? (ibid.). He suggests that if "each neuron's contribution to psychological capacities is solely its transmission of an electrical signal, and if silicon chips contribute to psychological capacities in precisely the same way, then the silicon brain and the neural brain are not distinct realizations of a mind" (ibid.). According to Shapiro, this thought experiment is not successful in showing that minds are multiply realizable, because, the second condition for being multiply realized is not held in this example.

But, what about the other intuitive examples of multiple realizations? Are they false examples as well? Consider the alleged example of the multiple realization of eyes in the mammalian and the octopus. Shapiro argues that in both the mammalian and the octopus eye, the ways the image is processed are quite similar. "Each eye has a single lens that causes an inverted image to fall on a retina, where light is then transformed into electrical signals" (ibid., p. 646). He suggests that "the information that the lens of the octopus's eye is composed of one kind of protein and the lens of the mammalian eye of another, or that the two eyes utilize different kinds of visual pigments" is irrelevant (ibid.).

Another argument for the MRT is provided from the cases of neural plasticity. The cases include the changing of the brain as a result of experience while still establishing the same psychological function (The examples and definition of neural plasticity can be found in 3.3.3). Shapiro mentions that there are studies which show that the "brain of a violin player or a person proficient with Braille will differ in predictable ways from the brains of those who do not use their fingers for very fine discriminations … The motor cortex in these specially trained individuals will have more space dedicated to representing the inputs from those fingers used in these operations than the motor cortexes of individuals not so trained" (2002). But, the question is whether this is an example of multiple realizability.

According to Shapiro, it is not. He suggests that the brains of the trained individuals and the untrained individuals are not performing the same function (ibid.). One could say that the motor cortexes of the trained and the untrained individuals are doing the same thing in the same sense that he could take the analog and the digital watch are the same functional kinds. That is, in a very broad sense,

without a good frame of reference, they are doing the same things. But in a finer sense, the trained and untrained motor cortexes are realizing different kinds. And, insofar as the kind is not the same in different realizations, there is not multiple realization.

One further argument for the MRT is provided from the Artificial Intelligence (AI) point of view. The reason why Shapiro is not convinced by AI arguments does not directly derive from the conditions posited in the aforementioned analysis of the notion of multiple realization. In 5.4.1, the reason why functional isomorphism does not guarantee the sameness of kinds has been introduced. For the very same reason, that some programs are functionally isomorphic from one stance with the mental properties of a human being does not entail that the realized kinds are the very same kinds in these realizations. For example, a cognitive model might be functionally isomorphic with a human being in a given psychological task. Shapiro argues that "the fact that we can describe the operation of a system at a high level of abstraction does not imply that the system is multiply realizable … Once we see the distinction between a description of a system and the system described, the temptation to move from claims of functional isomorphism to the truth of MRT loses its allure" (2000, p. 639).

5.5 A Rejoinder to Shapiro and Bechtel & Mundale

In 5.3, it has been said that the grains problem introduced by Bechtel and Mundale (1999) needs to be evaluated. Since Shapiro's argument against the MRT echoes this problem, a common rejoinder to both accounts would be appropriate. It is suggested that multiple realizability is established by means of using different criteria for identifying mental states on the one hand, and differentiating physical states on the other hand. Talking about functions allows, to some extent, appealing to vague terms. That is, as Shapiro suggests, what the function of a system is ambiguous in some cases[15], and as Bechtel and Mundale claim, when we identify functional kinds, we are using coarse grains[16]. Moreover, as Shapiro says, we distinguish physical

15 Is the function of the digital watch to show the time, or to keep the telephone numbers as well?

16 We say that both analog and digital watches are realizations of the kind *watch.*

states from each other quite easily[17], or as Bechtel and Mundale suggest, we use fine grains.

However, although these are important questions, once can still ask the following question: Is there only one possible physical way to have a given mental state? Our intuitions and pre-scientific observations of the biodiversity imply that the answer is negative to this question. Not to mention the possibility of the existence of physical systems which are completely different from each other, there are actual cases of multiple realization. Shapiro might not be convinced that the mammalian eye and the octopus eye are multiple realizations, but would not he still be convinced that mammalian psychology and octopus psychology in general are cases of multiple realizations? He would argue that the latter is not a case of multiple realization, because the realized kinds are different in each cases; One is the *mammalian psychology*, the other is the *octopus psychology*. But, one might still ask Shapiro: Why not take *psychology* as the realized kind in two realizations? Since both the mammalian psychology and the octopus psychology are psychologies, it seems that we have a genuine case of multiple realization.

Moreover, and more importantly, that there are physical similarities in different realizations of a psychological property does not show that the MRT is false. All it does show is that a strong interpretation of the MRT for showing that physical explanations are not relevant at all is false. That is, if there are some physical similarities between different realizers, there is a good chance that, contrary to what Putnam proposes, the physical description of a psychological event would apply to another psychological event of the same type. Recall that this last point is compatible with, or an extension of the revision I have proposed in 4.5.2.That is, the MRT could be true, but there could still be an underlying similarity between the physical realizers of a mental property. If Shapiro's example is correct, although his conclusion is not tenable, a physical explanation of how a mammal sees would be similar to a physical explanation of how an octopus sees. This does not mean that the

17 We suppose that the mamalian brain and the octupus brain *must* perform in different physical ways.

MRT is false; it just shows that the MRT is not sufficient for dispensing with physical explanations.

5.6 Embodied Cognition and the MRT

There is a research program in cognitive science which is called Embodied Cognition (EC). EC tries to explain mental phenomena by means of appealing to the relation of mental properties to the body and the environment. As Wilson summarizes, the main claim of the EC is that "cognitive processes are deeply rooted in body's interactions with the world" (2002, p. 626). As Shapiro points out, it "departs from more traditional cognitive science in the emphasis it places on the role the body plays in an organism's cognitive processes" (2007, p. 338). Here, the "more traditional" cognitive science refers to the science which assumes the truth of the ideas that are introduced in the 3.2. Yet, here, it is appropriate to point out two basic assumptions the traditional view endorses, which according to the EC are false:

1- Cognitive processes are processes which are similar to the processes one finds in a computer. Our sense organs are input devices which take information from the environment and transform it to syntactic codes. So, the nervous system can manipulate these codes so as to produce inner codes or outputs. Thus, cognition, as Pylyshyn (1984) argues, is computation over representations of the world.

2- Cognition is isolated from the external world. Cognitive processes are held only on the inner codes. This vies is called methodological solipsism (Fodor, 1980), because the things which cause the inputs, and the things which are caused by the outputs are not relevant for the understanding the cognitive processes.

According to the EC, these two assumptions are rejected on these grounds Shapiro, 2011, p. 26:

(a) If the world is given to the mind and is available whenever the mind needs to consult to, then a rich representation of the world to act on is not required;

(b) The body and the environment is part of the cognitive processes, so we cannot insulate cognition from these two.

The endorsement of (a) and (b) requires more than speculation. For this, more particular claims, and evidences and examples for establishing these claims should be introduced. Thus, Wilson (2002) outlines the six views which are associated with the EC, three of which will be relevant for establishing (a) and (b).

1. Cognition is situated. It takes place in an environment, and it makes use of perception and action inherently.
2. Cognition is time pressured. The real-time interaction is inherent in cognition.
3. We off-load cognitive work onto the environment. Because of our cognitive limits (e.g. on working memory), we make use of the environment.
4. The environment is part of the cognitive system. Mind and world are connected to each other in such a way that studying mind without understanding its relation to the environment is impossible.
5. Cognition is for action. Our cognitive states should be understood with respect to their contribution to behavior.
6. Off-line cognition is body based. Even abstract cognitive abilities, such as thinking about things which are remote in time and space, are grounded in mechanisms that are inherently related to the body (ibid., p. 626).

The claims 3, 4 and 6 are used for establishing (b) and (a), and it is now time to see what kind of evidences and examples are given for 4, 3 and 6 respectively:

4- The environment is part of the cognitive system.

Clark and Chalmers (1998) provide the following thought experiment[18]. There are two people: Inga and Otto. Inga has a normal memory, and her memory includes the information that the museum is on 53rd Street. She decides to visit an exhibit, and remembers that the museum is on 53rd Street. Otto, however, has

18 In the mentioned paper, Clark and Chalmers (1998) argue for Extended Cognition (ExtC), which is, for metaphysical reasons, different from the EC. While the EC claims that the information about the bodily and the environmental features are explanatorily as crucial as the information about the inner processes, the ExtC claims that such features are part of the processes. That is, the latter is a stronger claim than the former. For the sake of the point made here, the difference between these theories is not important, as the ExtC entails the EC.

Alzheimer's disease, and he uses his notebook to remember addresses. His notebook includes the information that the museum is on the 53rd Street. He decides to go to the same exhibit, and he checks his notebook to see the address. Both Inga and Otto go to the same exhibit. There is no difference between Inga's and Otto's beliefs about the address of the museum. The only difference is that, Inga's belief is in her brain, whereas, Otto's belief is in his notebook. "The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin" (ibid., p. 13). That Otto's case is not a normal case is not of importance here. He could have a normal memory, and yet still could not remember the address without checking the notebook. This thought experiment, as Clark and Chalmers argue, shows that the mind reaches beyond the boundaries of the head. That is, the relevant objects in the environment, and perhaps the environment as a whole, could be taken as part of the cognitive system. If this is true, obviously, the solipsism assumption attributed to the traditional cognitive science could be rejected.

3 & 6 - Off-line cognition is body based & we off-load cognitive work onto the environment.

As Wilson (2002) points out, following examples are given for supporting these two claims. In cases where subjects play Tetris[19], the "data suggest that players use actual rotation and translation movements to simplify the problem to be solved" (ibid., p. 629). One can think examples from one's daily life. One example is moving in a room so as to see what kind of furniture would suit the room. Another example is laying out the pieces of a toy train on a table in a slightly organized way so as to understand how to set it. These are examples for showing how off-loading work onto the environment makes the cognitive processes more efficient. An example for showing that off-line cognition, such as abstract thinking, is body based is the efficient use of fingers when counting (ibid., p. 632).

Recall that these claims are used for rejecting the assumption that cognitive processes are carried out by means of making computations over representations.

19 A computer game in which the player, by using arrow keys, has to locate falling blocks of different shapes to the ground.

Although these examples are not conclusive for showing that this assumption is false, what they do is showing that building a rich representation of the world is not required, and indeed is not efficient, for carrying out some or many (or perhaps all) cognitive process, insofar as the body and the environment are available for the use of cognition.

Thus, if (a) and (b) are endorsed, the thesis that mind is a self-contained organ separate from the body is false. Shapiro argues that, the examples that are raised by the scientists of the EC show that it is true that mental properties are informative about the physical properties of the body and one can predict the properties of the body from the facts about the mind (2004, p. 225).

What does this latest claim tell us with respect to the purposes of this thesis? First, it should be recalled that the MRT is used by Putnam (1975) for showing that physical explanations are irrelevant for cognitive science explanations (see 3.4 for details). Yet, if the claims of the EC are true, Putnam's claim is false. That is, if one can predict physical facts from the facts about psychology, there is a closer link between physical explanations and psychological explanations than what Putnam has thought. Second, and more importantly, the EC's way of explaining things casts doubt on the MRT itself. According to the MRT, same mental properties could be realized by different physical systems. This means that the world impose few constraints on what can realize a given mental property. Yet, if the strength of the connection of the mind to the body, as the EC proposes, is considered, it could be suggested that even the way our bodies are poses constraints on what kind of mental abilities we are capable of. Let alone being made out of Swiss cheese, even having an extra finger on the right hand might have inconceivably drastic effects on the mental lives of human beings.

However, one must take this sort of argument against the MRT with considering the following problem. The reason why the EC's ways of explaining things are contradicting with the MRT is that, as it was mentioned above, the EC's claims are rejecting the assumptions that are central to the CTM. Therefore, an elaboration of the incompatibility of the predictions of the EC and the MRT requires tracing the tension back to the incompatibility of the CTM (or the traditional cognitive science) and the EC. This is beyond the scope of this thesis.

5.7 Concluding Remarks

In this chapter, four objections have been introduced against the MRT. Zangwill's argument from "qualia" against the MRT has been found not to be convincing. The arguments provided by Bechtel and Mundale, and Shapiro have been evaluated in one rejoinder, and it has been argued that what they can show is that the MRT does not have the strong implications contrary to what Putnam has suggested. However, the arguments provided cannot show that the MRT itself is false. The rejoinder echoes the revision I have proposed in 4.5.2. Finally, the Embodied Cognition research program (EC) was introduced so as to provide an argument against the MRT. Yet, it has been argued that the reason why the explanations provided by the EC are contradicting with the MRT is that the basic assumptions of the CTM are rejected by the EC.

CHAPTER 6

CONCLUSION

6.1 A Summary

This thesis attempted to assess the status of the Multiple Realizability Thesis (MRT) in cognitive science, which has been formulated as a philosophical thesis for refuting Reductive Physicalism. According to the latter, mental properties can be reduced to physical properties.

What the MRT seems to tell us is that such a reduction is not possible: there are several ways in which things can realize mental properties, hence it is a fallacy to argue that one can reduce a given type of mental state to a given type of physical state. What is more interesting is that this thesis was formulated as an extension of the Computational Theory of Mind (CTM), which has been one of the working hypotheses of the traditional cognitive science. The gist of the CTM is that mental life could be explained in terms of computations. Being computationally equivalent, or being functionally isomorphic, does not require being physically isomorphic, thus, the MRT holds.

We have seen that there are two original positions which are directly related to cognitive science:

(i) Reductive Physicalism: Cognitive science can be reduced to a basic science such as physics.

(ii) Anti-Reductionism: Cognitive science presumes the CTM, which establishes the MRT, and the MRT refutes Reductive Physicalism.

Although, in the last chapter, we have seen that one can find theories in cognitive science which are not endorsing the CTM, such as the Embodied Cognition research program, it is a fair generalization to suggest that what is called the tradition cognitive science assumes the CTM.

Of course, inevitably, there have been objections to Anti-reductionism:

(iii) From the MRT, it does not follow that Reductive Physicalism is false.

(iv) The MRT is itself false.

The elaboration of the ideas from (i) to (iv) correspond to the structure to the chapters of this thesis (from Chapter 2 to 5). I have provided brief evaluations on these ideas where applicable throughout the thesis, yet, for the conclusion I will be proposing here, it is best to summarize the points that are made in each chapter briefly.

Reductive Physicalism

There is the idea that cognitive science can be reduced to the science of the brain. Such a reduction requires a model, and one model which the reductionist can appeal to is the Nagel model, which suggests that reduction is a relation between theories, and one needs bridge laws connecting the properties of the two theories for deducing the statements of one theory from the other. When it comes to the reduction of the cognitive science to the science of brain, which is assumed to be a physical science, the reductionist appeals to the identity of the properties of cognitive science with properties of the brain.

Anti-reductionism

According to Anti-reductionism, such identities cannot be held. The MRT just shows that there is no one-to-one mapping between the properties of cognitive science and the properties of the brain. Therefore, the reductionist's only way of providing bridge laws is to propose that these laws are disjunctive. Yet, as it has been shown, having disjunctive bridge principles does not work either.

Against Anti-reductionism

The MRT does not show that reductionism is false. Among several accounts, Kim's account proposes that what the MRT shows is that mental properties are not scientific kinds. It shows that the same mental property is realized by different physical properties, and this means that the investigation of one realization of a given

property does not provide us information about another realization of the same property.

This should be taken as a challenge to any cognitive scientist, as the task of a cognitive scientist is to explain what mental properties are. If these are not scientific properties, what is the point of studying them as the subject matter of a given science? I have argued that this challenge could be replied easily: One does not have to assume that the MRT shows that there are vastly different physical realizers of mental properties. There can be an underlying similarity between different realizers, thus, it is likely that physical properties of an instance of a mental state would project information about another instance of the same mental state.

Against the MRT

Several arguments against the MRT were formulated. The objections against the MRT were rejected, and it has been argued that the best objection against the MRT boils down to the claim that there are physical similarities between different realizers of psychological properties. But this does not show that the MRT is false; it only allows us to dispense with a strong anti-reductionist implication of the MRT which says that physical explanations are not relevant at all when it comes to the study of psychology.

6.2 A Conclusion

A challenge to the scientific status of cognitive science has been raised by an entailment of Kim's argument. A cognitive scientist might not be interested in dealing with this challenge, as cognitive science, trivially is a science. Yet, I took this as a challenge that could be replied. The reply is that, the properties of cognitive science, despite being multiply realizable, are likely to exhibit a natural similarity.

The very same reply to this challenge takes us to a position which is somewhat critical to the anti-reductionist argument from the MRT. As long as there is some kind of a unity, or presumed similarity among the subject matter of cognitive science, the extent to which the properties of cognitive science are multiply realizable is restricted. This does not mean that the MRT is false. But, it shows that one can still have physical similarities between the realizers of psychological properties.

The last point does not mean that Reductive Physicalism is true. However, it definitely shows that the strongest version of Anti-reductionism, which says that physical explanations are irrelevant to cognitive science, is false.

BIBLIOGRAPHY

Batterman, R. W. (2000). Multiple realizability and universality. *British Journal for the Philosophy of Science* 51 (1):115-45.

Bechtel, W. P. & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural States. *Philosophy of Science* 66 (2):175-207.

Block, N. (1990). Can the mind change the world? In G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam.* Cambridge: Cambridge University Press.

Block, N. (1997). Anti-reductionism slaps back. *Philosophical Perspectives* 11:107-32.

Block, N. & Fodor, J.A. (1972). What psychological states are not. *Philosophical Review* 81:159-81.

Chomsky, N. (1959).Review of Skinner's Verbal Behavior. *Language,* 35*:*26-58.

Clark, A. & Chalmers, D.J. (1998). The extended mind. *Analysis* 58 (1):7-19.

Dennett, D. (2009) Darwin's “strange inversion of reasoning”. *Proceedings of National Academy of Sciences of the United States of America* 106:10061-65.

Fodor, J.A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese* 28 (2):97-115.

Fodor, J.A. (1975). *The Language of Thought.* Cambridge, Massachusetts: Harvard University Press.

Fodor, J.A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences* 3: 63-109.

Fodor, J.A. (1997). Special sciences: Still autonomous after all these years. *Philosophical Perspectives* 11:149-63.

Fodor, J.A. & Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture. *Cognition* 28:3-71.

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32:127-36.

Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy* 83:291-5.

Johnson, M. (2005). Developmental cognitive neuroscience: An introduction. Oxford: Blackwell

Kim, J. (1996). *Philosophy of Mind*. Westview Press.

Kim, J. (1999). Multiple realization and metaphysics of reduction. In J. Kim & E. Sosa (eds.), *Metaphysics: An Anthology*. Blackwell Publishers.

Kim, J. (2002). Preécis of Mind in a Physical World. *Philosophy and Phenomenological Research* 65 (3):640–43.

Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton University Press.

Kim, S. (2002). Testing multiple realizability: A discussion of Bechtel and Mundale. *Philosophy of Science* 69 (4):606-10.

LePore, E. & Loewer, B. (1989). More on making mind matter. *Philosophical Topics* 17:175-92.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace & World.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review* 83:435-50.

Oppenheim, P. & Putnam, H. (1958). The unity of science as a working hypothesis. In H. Feigl et al. (eds.), *Minnesota Studies in the Philosophy of Science, vol. 2*. Minneapolis: Minnesota University Press.

Piccinini, G. (2003). *Computations and Computers in the Sciences of Mind and Brain*. *Dissertation*. Dissertation, University of Pittsburgh.

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

Polger, T.W. (2009). Evaluating the evidence for multiple realization. *Synthese* 167 (3): 457-72.

Putnam, H. (1960). Minds and machines. In S. Hook, (ed.) *Dimensions of Mind*. New York: New York University Press.

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

Putnam, H. (1967b). The mental life of some machines. In H. Castañeda (ed.), *Intentionality, Minds and Perception.* Detroit: Wayne State University Press.

Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers, vol. 2*. Cambridge, Mass: Cambridge University Press.

Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science.* Cambridge, Mass: The MIT Press.

Shapiro, L.A. (2000). Multiple realizations. *Journal of Philosophy* 97 (12):635-654.

Shapiro, L.A. (2002). Neural plasticity and multiple realizability (html document). Retrieved from http://philosophy.wisc.edu/shapiro/papers/neural%20plasticity%20spp.htm

Shapiro, L.A. (2004). *The Mind Incarnate*. Cambridge, Mass: MIT Press.

Shapiro, L.A. (2007). The embodied cognition research programme. *Philosophy Compass* 2 (2): 338–46.

Shapiro, L.A. (2011). *Embodied Cognition*. New York: Routledge.

Smart, J.J.C. (1959). Sensations and brain processes. *Philosophical Review* 68:141-56.

Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science* 66:542–564.

Wilson, M. (2002), Six views of embodied cognition. *Psychonomic Bulletin & Review* 9 (4): 625–36.

Zangwill, N. (1992). Variable realization: Not proven. *Philosophical Quarterly* 42 (167):214-19.