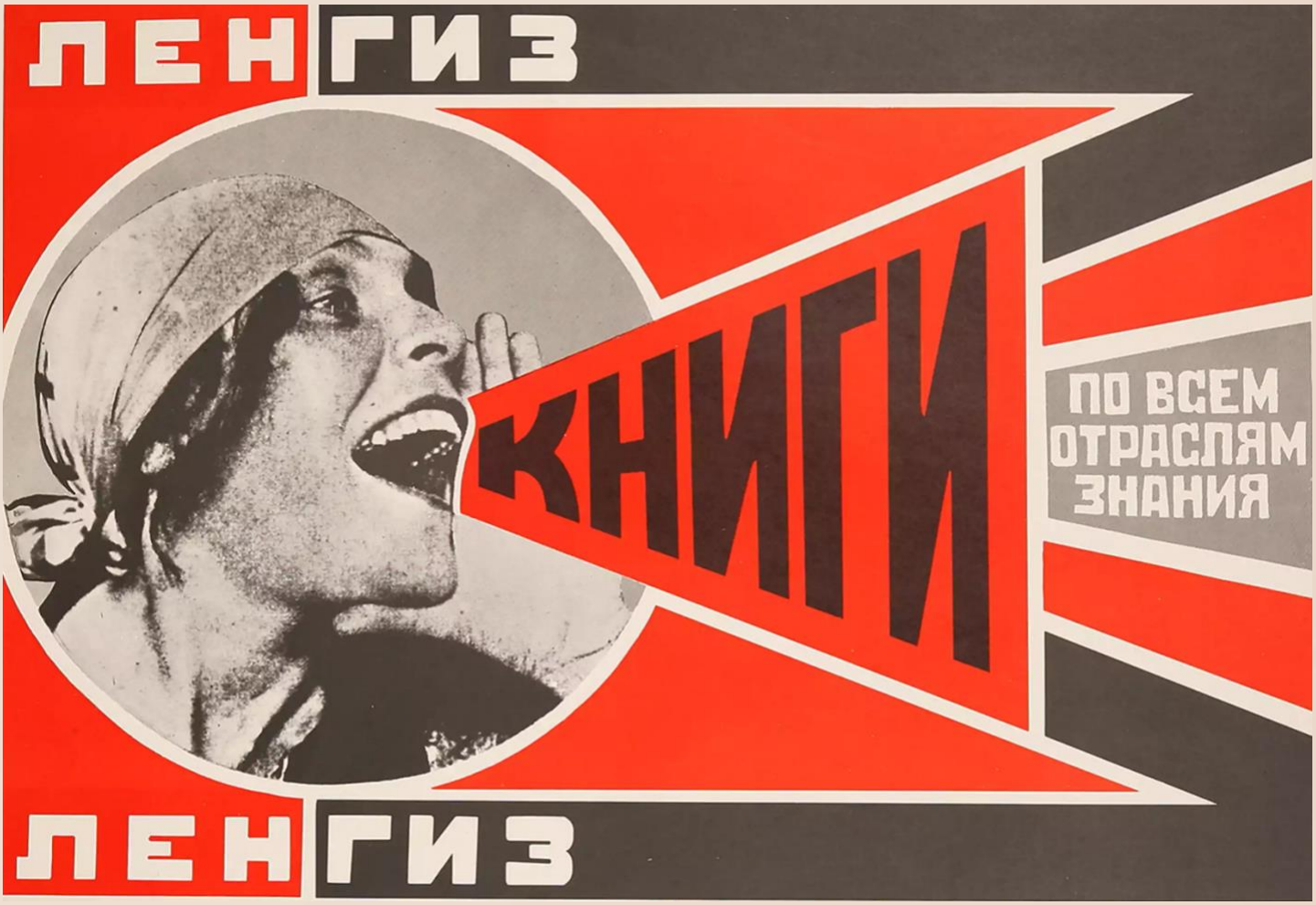
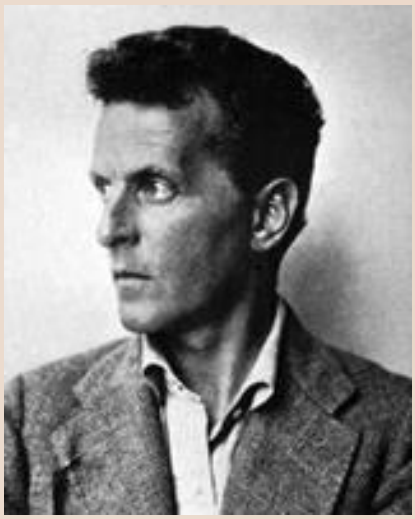
INQUIRIES INTO COGNITION



Wittgenstein's Language-Games and Peirce's Cognitive Semeiosis for the Philosophy of Cognition



by ANDREY PUKHAEV

Pontificia Università Gregoriana
Facoltà di filosofia

Andrey Pukhaev

INQUIRIES INTO COGNITION

Wittgenstein's Language-Games and Peirce's Cognitive Semeiosis for the Philosophy of Cognition

Dissertation for the Doctorate
in the Faculty of Philosophy
of the Gregorian University

Rome 2013/2023

Artwork by Aleksander Rodchenko, 1924
Set in 11 point Palatino Linotype

PREFACE

This work is a revision of the concept of the mind and substituting for it 'cognition,' redefined as the natural ability of constructing infinite strategies out of the limited rules within families of different language–games. Consequently, philosophy of cognition will be suggested instead of philosophy of mind. This conclusion comes as the logical consequence of the analysis of different philosophical psychologies from Plato to Kant and different philosophies of mind from the 19th century theories to today's paradigms of the discipline.

This work is methodologically deconstructive, reconstructive and instructive from the perspective of five fundamental questions in the philosophy of mind: ontological, metaphysical, epistemological, semantic and logical (the problem of identity). Part I is the deconstruction of the most influential theories from the perspective of these questions. Part II is reconstructive from the perspective of the philosophical influences of Ludwig Wittgenstein (1889 – 1951) and Charles S. Peirce (1839 – 1914). The last chapter is instructive by providing new theoretical foundations for substituting the philosophy of cognition for the philosophy of mind.

This doctoral dissertation is the result of my studies in philosophy at the Gregorian Pontifical University. It has been made possible through the encouragement, guidance and supervision of the academic moderator, Prof. Gennaro Auletta. Many of the ideas proposed here are indeed further elaborations of his innovative theories on the same and similar subjects. His patient work of supervising the progress has influenced this dissertation in many ways and deserves my profound and sincere gratitude.

Many ideas in this study were taught and inspired by Prof. Jaakko Hintikka, my MA moderator, and by Prof. Judson Webb at Boston University. I would like to express my sincere gratitude to Prof. Hintikka for his illuminating introduction to the works of Wittgenstein and Peirce, and to Prof. Webb for his instruction on the philosophy of mind and computationalism.

David Mayer proofread this work for English. I am very grateful for his tenacious and hard work. I would also like to thank Peter Knecht for controlling the consistency of my translations from German into English.

CONTENTS

INTRODUCTION

Philosophy of mind finds itself in a somewhat peculiar situation. On the one hand it has the well–defined, independent subject of the human mind to explain. On the other hand, its methods are not of its own. Explanations of the mind demand a theoretical basis rooted in ontology and metaphysics: the questions of existence and of the nature of the mind. Furthermore, when these questions are confronted, the questions of epistemology and semantics will follow, the questions of *how* we come to know *that* the mind exists as some reality (i.e. an emergent property, or function, or substance, etc.), and *how* we can state them in the form of a meaningful discourse. Finally, whatever position one takes on the relationship between matter and form or body and mind, the question of the identity between them must be answered as well.

The five questions sketched above do not have the same formulations throughout the history of philosophy, just as the definitions of metaphysics, ontology, epistemology, logic and semantics vary significantly over time and their authors. Throughout this work I will change the formulations of the questions in keeping with the originality of the philosophical discourse of those who provided the answers for us.

These five questions are only the foundations of the philosophy of mind. The issues of consciousness, free will, causation, artificial intelligence (AI), thought, volition, emotion, intentionality, memory, action, perception, kinaesthesis, qualitative experiences (qualia), the self and personhood, are the core of the discourse of the philosophy of mind. How these will be treated, however, depends essentially on how the five foundational questions would be answered. Formulations, answers, considerations and revisions of these five questions are the purpose of this present work, while the treatment of the above mentioned topics within the philosophy of mind will be done very superficially and will rather depend on the context of the selected methodological questions.

The usual way of doing the philosophy of mind was either through the dualism/monism of the ontological and metaphysical affirmation of human nature, or by rejecting the dualism/monism choice in favor of some alternative approach. The alternatives to dualism/monism are very scarce and, perhaps without exception, come as the consequence of the radical revision of ontology and metaphysics.

Dualism treats the mind and body as two realities, substances, forms, functions, properties, terms or concepts. It comes in three varieties. Substance dualism considers mind and body as two distinct realities, substances and sometimes as forms (i.e. in later Plato). Property dualism treats mind as a form, function or some property of the body/brain. Propositional dualism endorses ontological monism, physicalism, but allows of speaking of the mind as an independent concept.

Monism endorses the oneness of the human person by some kind of reduction or elimination. It comes in two forms: idealism and physicalism. Generally, idealism affirms the reality of the mental by reducing all physical states as mere manifestations of the mental. Physicalism, in its reductionist form, reverses the idealist thesis in saying that mental states are either properties or functions of the brain states. In its eliminativist form, physicalist denies the existence of the mind entirely. However, even in eliminativism, metaphysical, epistemological and semantic questions are treated one way or another.

Each of the above briefly outlined varieties of dualism and monism has many forms in the history of the philosophy of mind. The main method of this work will be a series of inquiries, an investigation into what seems to be the most influential and the most interesting variety of dualism and monism. The methods of inquiry will be borrowed from Ludwig Wittgenstein and Charles S. Peirce. The major concept of this inquiry's method is that of language–games and *semeiosis*, which explains the title of this work.

The way of proceeding for each part, chapter and section is to examine first the primary sources and then give an assessment of the ideas and conclusions of the authors. After that the secondary literature is presented and evaluated, and the second assessment in the light of the secondary literature on the author or topic is given.

The work is divided into two parts. At the end of each part I propose a conclusion which coincides with the thesis of this work. Hence, the two parts are meant to propose and prove two theses. The first thesis is the rejection of the dualist/monist method and language in the philosophy of mind as inconsistent with our current understanding of either 'physical' or 'mental' in reference to human subject. As a way of proof I conclude that none of the five questions posed to the philosophy of mind can be answered in either dualism or monism. The second thesis is the alternative approach to the study by suggesting a redefined notion of cognition in terms of the language–games and *semeiosis*. This conclusion, or thesis, follows in many ways from the studies of Wittgenstein (Chapter III) and Peirce (Chapter IV), but in some aspects is novel and independent in respect to these two authors. What is consistent with Wittgenstein is the use of his notion of the language–games and with Peirce is his theory of signs interpreted as the way these language–games work semantically and epistemically. What is novel and independent in relation to Wittgenstein is a different notion of identity, and in relation to Peirce is a suggested non–temporal notion of continuum. Finally, both terms, identity and continuum, are considered essentially relevant to the larger notion of cognition viewed as a four–dimensional language–game.

Part One is negative and deconstructive. I examine each theory from the perspective of the above questions. Chapter I looks into the terminological evolution of the *psuchê* in Plato and in Aristotle. The term is then compared with the use of the *anima* in the Middle Ages, especially in Thomas Aquinas. The central topic in that chapter is the Cartesian and post–Cartesian period in their uses of the Latin *mens* and the English *mind*, contrasted with the Kantian *Seele*.

Chapter II adds the logical question of identity to the previous four. The same method is used for examining the leading contemporary theories in the philosophy of mind. This chapter draws many conclusions based on the recent works not only in the

philosophy of mind but also in the philosophy of language, mathematics, logic, cognitive science, cognitive biology and some research in cognitive neuroscience.

The main conclusion of Part One is already inspired by the method of philosophy from the later works of Wittgenstein, in his criticisms of metaphysics and with his alternative view in epistemology and semantics. The conclusion is that none of the five questions can be answered within the methodological framework that extends from dualism to monism, and that logical identity treated as mathematical equality is inapplicable to the mind/body relations. This, indeed, is both the negative and deconstructive aspect of Part One.

As a consequence, Part Two proposes the alternative approaches of Wittgenstein and Peirce in their rejections of both dualism and monism. The central theme of Chapter III is Wittgenstein's notion of language–games applied to his, so called, Private Language Argument (PI §§243–315). The attempt to re–state the five questions within the dualistic/monistic methodological framework turns out to be, by Wittgenstein's philosophical criteria, nonsense. Ontology and metaphysics turn out to be critical analysis of our language and meaning acquisition, use and expression are essentially bound to the use of the language–games by us, their players. Language–games are presented in their three–dimensionality as a description of our cognition.

Chapter IV examines four accounts of Peirce's semeiotics and five kinds of his theory of continuum. The use of language–games is interpreted as semiotic activity, the use of signs (*semeiosis*) and notion of continuum as the fourth and final dimension of language–games. Wittgenstein's elimination of the notion of identity as tautology or as a contradiction in the previous chapter is contrasted with Peirce's use of John Duns Scotus' concept of virtual identity. This notion of identity, as mentioned above, is elaborated further and somewhat independently from Peirce's own definitions.

The final Chapter V suggests a new philosophy of cognition in which the five questions of methodology are transformed into the three questions: What is cognition (ontology and metaphysics)? What is identity (logic)? What is continuum (epistemology and semantics)?

It is important to stress what this work is *not* about. First, it is not a comprehensive treatment of the history of the philosophy of mind. The method suggested is very clear: a series of inquiries by four or five questions into the leading theories of the philosophy of mind, and inquiries through and by the language–game into what cognition is. The result is a refutation of the philosophy of mind for the suggested philosophy of cognition.

Second, none of the presentations of the two authors follows closely any of the accepted interpretations of their philosophies. This is particularly true of Wittgenstein. Nor are their treatments exhaustive. Instead, some themes in their philosophies are stressed more than others for the purposes of this work. In sum, this work is not an introduction to the philosophy of mind, nor it is a systematic commentary on Wittgenstein or Peirce in some form of comparative study. It is inquiries to problems and suggested solutions. It is a logical and philosophical–linguistic examination, investigation of these problems.

A final note is on the treatment of Wittgenstein's texts. Initially, for the economy of research and time, I decided to use the available translations by Wittgenstein's

students G. E. M. Anscomble, R. Rhees and G. von Wright, as well as some more updated critical editions of Wittgenstein's works by P. Hacker and J. Schulte. However, some of these translations use terms that originate more from the translators' own philosophical ideas than what the original texts say. This became more apparent when I consulted with several secondary source criticisms of the PI English translations. Consequently I was making my own translations trying to get the most resolute and accurate rendering of Wittgenstein's language. These fragmented translations grew in volume so much that the use of the printed translations of Wittgenstein's texts would seem to be doubling and out of place. I decided to make all German translations my own, with very few exceptions, now for the sake of consistency if not accuracy. Whenever the authorship in the translation's footnote is not recognized, the translation is mine. Otherwise, the source of the quoted translation is provided.

Some exegesis of Wittgenstein's works demand more precise numbering of paragraphs, sentences and words than Wittgenstein's own division into sections. This, however, was not an issue for Peirce's texts. Thus, some sections in PI are divided further into numbered paragraphs, sentences and parts of the sentences, usually divided by the semicolon or coma. Thus, a section is given a numeral (i.e. §1), paragraphs within the section are given superscript letters ($§1^{a}$), sentences within each paragraph are given superscript numerals ($§1^{a1}$, $§1^{a1–3}$), and divisions within sentences split by semicolon or coma, are given superscript roman numerals ($§1^{a3i}$).

Wittgenstein's own punctuation in the German text is preserved, while the punctuation in the English translation follows the way that the printed editions have done it. In both the English translations in the text and in the German text in the footnotes, all abbreviations of Wittgenstein works, for the sake of consistency and in order to avoid confusion, are made in English and not in German.

PART ONE

CONCEPTUAL EVOLUTIONS OF SOUL, MIND AND BODY: FROM PHILOSOPHICAL PSYCHOLOGIES TO PHILOSOPHIES OF MIND

CHAPTER I

The Tales of Two Philosophical Psychologies: An Introduction and Historical Background

Introduction

This chapter will claim that in the course of the history of philosophical psychology and of the philosophy of mind one can observe a twofold evolution. First, it is the evolution of terms from the Greek *psuchê* to the Latin *anima* to the Latin *mens*, or the English *mind*. Second, behind this linguistic evolution there was a broader evolution of philosophical ideas and changes in metaphysics and epistemology.

To show this twofold evolution I will select a very limited number of but, as I believe, the most influential, philosophers of the field and will examine their texts through the lenses of four questions: ontological, metaphysical, epistemological and semantic.

The ontological and metaphysical questions are: 'Does the mind exist and is it distinct from the body?' A strictly metaphysical question would be: 'What is the nature of the mind?' The ontological question inquires whether mind and body exist and whether their relationship is real. The metaphysical question reflects on the nature of mind and body as well as their relation. Ontology is concerned with the distinction between substance dualism and monism: idealism and physicalism or materialism. Metaphysics is concerned with the questions of causality and the identity between mind and body. In the history of philosophical psychology and the philosophy of mind, the ontological question was often formulated as the question about the existence of other minds. The metaphysical question can be reduced to the mind/body relationship which is the problem of the explanatory gap.

The epistemological question follows from the ontological and metaphysical questions: 'What are the criteria for our knowledge of the mind–body relationship or identity?' The epistemological question is concerned with the possibility of the knowledge of the mind–body relationship, or the possibility of having such a theory to be coherent. Questions of perception, perceptual content, consciousness and qualitative experiences are all concerned with epistemology. Most theories of perception are representational in nature and postulate some kind of mental imagery (i.e. ideas, forms, impressions, sense–data). In relation to the ontological and metaphysical questions, the epistemological question asks how these mental images are related to the physical world. The old question of skepticism also originates from these inquiries.

Since the mind–world relationship is expressed in language, the last question is semantic: 'What is the relationship between mind–dependent language and mind–independent world?' or, 'What is meaning, provided that there is any?' Semantics is concerned with meaning, intentionality of language and reference.

Since *psuchê*, *anima* and *soul* have different meanings from the 'mind' (and they themselves have carried different meanings throughout time), I will ask these questions exchanging the word 'mind' with *psuchê*, *anima* and soul when appropriate, and will define how each term is applicable within the context of a selected author. These linguistic differences are important enough to change the nature of the questions, yet regardless which term is used, it is used because it was believed to be responsible for cognitive functions in general. Therefore, whether it is *psuchê* or mind, in different philosophies at different times they all explained the same cognitive apparatus, albeit differently. This is what makes them relevant to any contemporary study in the philosophy of mind with an interest in its history.

These questions are asked in order to test the most important philosophical theories on their ability to account for cognitive activity. The choice of questions is based upon the generally agreed definition of the contemporary philosophy of mind. For instance, Brain McLaughlin, in one of the most used current philosophical dictionaries, says the philosophy of mind is "concerned with the nature of mental phenomena and how they fit into the causal structure of reality".[1] If I may equate 'mind' with the "mental phenomena" and 'world' with the "causal structure of reality", then my questions above are nothing more but reformulations of one of such generic definitions of the philosophy of mind existing today. In addition, the epistemological question inquires into the possibility of knowledge of such a relation, and the semantic on its expressibility and meaning. Thus, the historical review in this and the following chapter is no more than a theoretical supplement for the contemporary philosophy of mind.

The first chapter is divided in two parts: First and Second Psychologies. By 'psychology' I mean 'philosophical psychology', which is a general protoscientific theory concerning mental phenomena. The First Psychology precedes Descartes, the Second Psychology covers modern philosophies of mind. The first chapter thus presupposes a philosophical evolutionary process that started with the pre–Socratic and Plato's *psuchê–logoi*, reminiscent throughout the ancient and medieval philosophies of *psuchê* and *anima*.

[1] McLaughlin (1999: 684).

Mostly for etymological reasons, philosophers prior to Descartes did not have, properly speaking, a 'philosophy of mind' but rather a philosophical psychology: a philosophical discourse on the *psuchê* as an explanatory agent of life and cognitive activity. From Descartes to the late–19th century those faculties of the *psuchê* pertaining to life were attributed to the laws of matter, and cognitive capacities to the mind. Yet, since the term *anima* (and its cognates in French *âme* and in German *Seele*) was used throughout modern philosophy in parallel with *mens* (*esprit, raison* and *Geist, Verstand, Vernunft*), I call this second period the Second Psychology. Contemporary use of the term 'mind' as some kind of emergent function of the brain was not in use until the late 19th century.

The following historical discourse on *psuchê* and mind is far from being comprehensive and the number of philosophers considered is very limited. The goal is not to give a history of philosophical psychology and the philosophy of mind but to test the foundations of the philosophy of mind as defined above alongside past theories of mind and soul.

The goal of the first chapter is not to systematically present philosophical psychology in its historical development, but rather to show the causes of its failure to answer questions emerging from their deeper philosophical commitments. These questions are essential in understanding the mind/body relationship and are indispensable for any philosophy of mind. If the above four questions are the consequence of the history of metaphysics and epistemology, then they must be asked and examined next to the theories which generated them.

1. The First Psychology: From *Psuchê* to *Anima*

1.1 Dualist versus Naturalized Ontology of *Psuchê*: Plato and Aristotle

1.1.1 *Plato: Dialectical Dualism*

Just as in the case of the Pre–Socratic philosophers, properly speaking, Plato had no theory of mind separate from his psychology. Within that context of the general discussion on the *psuchê*, Plato is the first to make the mind (*nôus*) as a distinctive part of the *psuchê*. But instead of the mind in a proper (contemporary analytical) sense we should rather speak of the *self* distinct from other faculties of the *psuchê*. For Plato, what we really are, is *psuchê*, and what we should become, is its highest faculty, *nôus*.

The mind in the contemporary sense and the Platonic self should not be confused even though the Greek term *nôus* does not accommodate such a distinction and some authors insist on Plato being the author of the modern concept of the mind.[2] But overall in the ancient world until the Middle Ages, the Greek notion of *nôus* (and consequently Latin *mens*) had a much broader meaning and application than now. If today by the mind we mean, very roughly, some emergent function of the brain, in antiquity *nôus* extended beyond physics and biology into metaphysics, ethics and even cosmology

[2] Lovibond (1991: 35).

(especially in Pre–Socratic philosophy). Through the intellectual capacity a person participated in the fundamental principles of the cosmos which often coincided with some notion of divinity.

The Greek notion of the *psuchê* can hardly be translated into its English linguistic equivalent 'soul'. While Greek *nôus* extends beyond its biologico–psychological use of today, *psuchê* played a much more concrete role in the Greek world. *Psuchê* was that which distinguished animate from non–animate objects and as a term was responsible for the explanation of life. Although it was Aristotle in his *De anima* who provided the first systematic account of life and, hence the discourse on the *psuchê*, already in the theological tradition of the Greek poets and in the Pre–Socratic theories this distinction was present. Because of this linguistic ambiguity I will use the Greek *psuchê* and *nôus* in relation to the Greek authors.

Prior to Plato, questions of living beings and intellect were not among the main questions. Just as Pythagoras, Plato inquires what is real in the world we live in. In the perceived changeable world the immortal, unchangeable and transcendent Forms are the only reality there is. What makes living things as such is the *psuchê*. People can come to the knowledge of the Forms by the use of the highest faculty of *psuchê*, the intellect.

The intellect, the utmost faculty of the *psuchê*, is capable of contemplating the Forms due to its equal nature with the Forms, viz. its immateriality. In *Phaedo* 76b4 – 84b8, Plato speaks of the immortality of the *psuchê* on the grounds of the affinity between *psuchê* and the Forms. In contemplating the Forms, *psuchê* tends to leave the body and reside within the Forms' realm which is *the only* reality there is. And just as the world of the Forms is the only real world, the true nature of humanity in each of us is with the *psuchê* alone. Plato's epistemology depends on his metaphysics and on his cosmology.

The early Platonic dialogues treat *psuchê* primarily as an organ of desire and longing and only secondarily as an arbiter of moral discernment and reason.[3] What a person really is, is that part of the *psuchê* that treats 'good living' as the supreme goal of life.[4] Choosing the good, the true and the beautiful represents a truly human way of life through the middle dialogues as well.[5]

Psuchê in *Paedo*, an early Platonic dialogue, is that which distinguishes animate from inanimate and is responsible for all life, from plants to animals, including humans. That notion would include *nôus* as well, but *psuchê* is not responsible for all of the cognitive functions of the *nôus*. I.e., fear, desires, pleasures and *eros* are only bodily functions.[6] These would be the inferior faculties of the *psuchê* in the *Republic* IV (appetite), while the mind in the *Republic* IV (reason) corresponds to the notion of the *psuchê* in *Paedo*.

[3] Plato (1953: 47e).
[4] Plato (1953: 48b).
[5] Plato (1935: 505d11).
[6] Plato (1953: 80b–83d. 94d).

Many before Socrates thought the unity of the *psuchê* as the condition of its immortality and indestructibility. Plato in *Phaedo* also shared this theory.[7] In the middle and later dialogues this unity is sustained but amended by drawing a three–fold distinction of the one *psuchê*'s faculties: the rational (*logistikon*), the passionate (*thūmoeides*), and appetitive (*to epithumêtikon*). His *polis* and *ethos* were envisioned to have the same distinctions as the human *psuchê*. Just as the class of philosophers–rulers is to control the army and the multitudes of craftsmen for the sake of the overall benefit in the one *polis*, so the ever–passing appetites, desires and emotions are to be controlled by the constant reason. The changeable world's impressions and transient opinions (*endoxa*) are to be transformed into knowledge (*episteme*) which always transcends the immediate and the visible.

Plato's tripartite theory of *psuchê* is also in accord with his earlier *psuchê*–somatic dualism. In his later dialogues this unity is only reinforced by his matured ethical and political theories. We read in *Theaetetus* that *psuchê* is to synthesize perceptual data; an idea in harmony with his later political thoughts in the *Republic*.[8]

But the later–Platonic *psuchê* includes also elements of the body, material elements without which sensory perception would be unthinkable. The notion of the *nôus*, on the other hand, is now connected only to reason.[9] To be human is to transcend the material, thus to become our true, genuine selves. But pure immateriality is where humanity no more. Plato's notion of god as a pure *psuchê* is the goal toward which one must strive but never achieve, a work in progress, a potentiality, a continuous *dunamis*.[10] Plato's view of humanity is that of a 'divine becoming' rather than of 'human being'.

Only in its affinity to the eternal Forms can *nôus* be that which makes the world (that of Forms and that of physical objects) and itself be known. And only because of the universal nature of the Forms in the cosmos can *nôus* come to know anything and itself. Notions of causality and perception in this sort of metaphysics become notoriously obscure, often explained in the language of myth and mysticism, making Plato's psychology rather unattractive to the contemporary philosophers of mind.

Plato's theory of Forms dominates his ontology and metaphysics on every level including his psychology. Relations between the three parts of the *psuchê* depend on the overall relations between subject and Forms. However, the mystical and mythological narratives in his dialogues that dominate in every period prevent a concise and clear understanding of how this relation works.

Particularly, it is not clear as to what ontological status has *nôus,* being that part of the *psuchê* in somewhat direct relation to the world of the Forms. However, it is rather clear that *nôus* alone cannot be accountable for all cognitive functions of the subject. In this respect the more positive accounts of perception in Plato's later dialogues are not sufficient for clarifying metaphysical questions on the nature of the *psuchê*, especially on the interactive dynamics between its diverse parts.

---

[7] Plato (1953: 78b–79b).

[8] Plato (1977: 184d).

[9] Plato (1953: 90a).

[10] Plato (1926b: 899b5).

The epistemological question has little chance of an answer until the ontological and metaphysical doubts are resolved first. Socratic 'knowing thyself' will rely on knowing the Forms corresponding to the objects of our knowledge. In closer analysis it is difficult to see the notions of 'subject' and 'self', although being central to Plato's psychology, as unified concepts. What is the self if not the unity of the *psuchê*, but what is *psuchê* if not ever–in–conflict unity of diverse parts? Aristotle will take on the insufficient theory of perception and unclear interaction between the different parts of the *psuchê* with much greater attention and precision of discourse.

In the context of the semantic question on meaning, the periodization of the Platonic dialogues is of significant importance as well. If in the earlier dialogues the Socratic method of inquiry was meant to dismiss *endoxicated* (opinionated) common answers more than provide the true ones, in the later dialogues we have not only the answers but also clear definitions of the concepts in question. Alas here, as in any other problems, everything will rely on the Forms and on the problem of interaction between the subject and the Form. Thus, the semantic question, also, will depend on the problematic accounts of Plato's ontology and metaphysics.

1.1.2 *Aristotle: Matters of Forms*

When Aristotle proposed his theory of *psuchê* as the universal principle of life which distinguished living from non–living things, he went against two main rival theories of the *psuchê* prominent before him: the pre–Socratic reductive materialism (especially the atomism of Leucippus and Democritus)[11] and dualism (Pythagoreans and Plato)[12]. For the first group, *psuchê* is a composition of atoms and the void just as much as material beings. *Psuchê* might be composed of a different matter (Heraclites), perhaps fire (Anaximander), but it remains material nonetheless. Aristotle refuses materialism on the basis of its inability to explain the distinction between living and non–living beings according to the atomic theory. If both are composed of atoms then what accounts for the locomotion of plants and animals?

Aristotle's psychology takes a different turn from pre–Socratic monism and Plato's dualism. For him, the question about 'how many' (one or two, monism or dualism) is incorrectly phrased. He accepts Plato's explanation that the *psuchê* is the principle of life of a body but extends it from human to animal and vegetative lives.

Forms, furthermore, are in no mythical realm but are properties of things and concepts alike. To know some object or to comprehend the meaning of a proposition is to acquire its *form*. In the case of a physical body, it is to obtain its perceptual form; in the case of a proposition, it is to acquire its meaning, to understand it. *Hylo* and *morphê*, unlike in Plato's dualism, are not in any conflicting relation, nor is one reduced to the other. Aristotle's hylomorphism was meant to work equally well in his logic as in his physics; the concluding part of the later was psychology.

---

[11] Aristotle (1907: 5–6, 3–15); Polansky (2007: 67–70).

[12] Aristotle (1907: 5–6); Huffman (2009).

Hylomorphism together with *energheia/dunamis* was meant to explain life in living beings. A potentially alive body is actually alive when the *psuchê* in–forms it. Unlike Plato, Aristotle allows also for bottom–up relationships between body and *psuchê*; *psuchê* is the property of the body, but not just any body, a particular body for a particular *psuchê*. Not every matter can fit in to the form: a house cannot be made of air, nor a human being of stone.

The perceptive faculties of animals are also explained by the same method, but in this case Aristotle seems to allow for two interpretations. The first view, sometimes called 'literalist', speaks of the perceiving organ becoming *like* the object which it perceives. For instance, the eye jelly (Aristotle's example) becomes red when it is exposed to the color red; a hand becomes warm when it is exposed to the heat. In this case we do not only speak of the perceiving organ that acquires the form of the object or an event of perception, but also include some of its physical properties. Second, just as in the hylomorphic process of apprehension in logic and language where an acquisition of matter would be impossible, similarly in the case of sensory perception the perceiving organ acquires *only some* properties (forms) *without* any physical acquisition of them. For example, the nose smelling a rose does not become a rose itself, just as the understanding of a story about monsters does not transport any monsters to one's head (or, according to Aristotle, to one's heart).

In the contemporary philosophy of mind a literalist interpretation was proposed by Richard Sorabji first in 1974[13] with some later elaborations.[14] This interpretation is based on the text from *De anima* (II, 5, 418a3–4), where Aristotle considers an organ of perception to be able to perceive potentially that which the object of perception is actually. Some of its contemporary proponents suggest that this literalist account can in some way be accommodated within contemporary accounts of neurophysiology.[15]

This view has received substantial criticism for being incomplete, at least in explaining perception. If Aristotle had focused his theory of perception on the perceptual apparatus, then he would have had to explain what makes each organ react to the object of perception the way it does and why sometimes mere sharing in properties with the object does not suffice for perception to occur.

The second view is sometimes called intentionalist, first proposed by Franz Brentano.[16] In many ways it is based on his interpretative reading of Aquinas.[17] This interpretation underlines the formal aspect of perception when the object of perception and the organ share not the material but the formal features of an object. This interpretation relies especially on the text from *Metaphysics* Z, 1032a32, b5, b22. Among its contemporary proponents the most known is Burnyeat[18].

---

[13] Sorabji (1974).

[14] Sorabji (2001).

[15] Wilkes (1995: 124); Moravcsik (1995: 138); Ahn (1995: 363). But see contrary view in Freeland (1995: 231 n.184).

[16] Brentano (1867).

[17] Aquinas (1984: 430a–b).

[18] Burnyeat (1995).

It seems to be wrong to ascribe the literalist interpretation as the only one offered by Aristotle. There is textual evidence to support not so much that his theory was exclusively literalist or intentionalist, but that he held both of them to be his theory of perception. Furthermore, neuroscientists perhaps can adopt hylomorphism as an analogical tool in generalizing about cognition (see A. Damasio[19]). Yet, if it is used as an explanatory tool, it not only generates more questions than explanations, it fails to accommodate data according to the Aristotelian method. Hylomorphism was designed by Aristotle to provide a general explanation of how perception operates (among other aspects of living bodies). The eye is not an organ of perception, properly speaking, but the *medium* between light and the neuronal networks of the brain. The sense of touch (the most basic sense common to all animals in *De anima* II, 2, 413b4–7 and in III, 12, 434b8–24) is not located in the body of the perceiver. In fact, I can share many physical features of the object perceived yet fail to perceive anything at all if the neural communication or processing is somehow disrupted.

Attempts to picture Aristotle as a non–reductive materialist with similarities to functionalism or supervenience theories came short of success given the limited scope of Aristotelian psychology in the first place. As was mentioned before, his psychology concludes his physics, and its general scope was to explain life without falling into either materialism or substance dualism. In the contexts of perception and active intellect (*nôus poiêtikos*) this general description of Aristotle with the overall methodological goal of avoiding extremes failed.

According to Aristotle, perception does not occur because the eye captures the form of what is seen but because being seen, being perceived is in the nature of colors, and to perceive them is in the nature of the faculty of sight.[20] That nature (or essence, immaterial form) of what is perceived, itself does not require an explanation. The nature of colors and light, which is that of being seen, explains perception *simpliciter*. Burnyeat, in opposition to Sorabji, in this context speaks of perception as something static like "a state of affairs".[21]

Similarly, Aristotle's doctrine of potentiality and actuality refers to the notion of the nature of things. It is one thing to say that a stack of bricks becomes a house, it is quite another thing to say that an eye *actually* sees, a hand *actually* feels, since the subject of perception is not entirely situated in the organ itself but also in the object of perception.

Until *De anima* III, 5, Aristotle uses hylomorphism to explain all the topics of psychology and avoids the extremes of dualism and materialism. In that chapter, the shortest yet the most controversial of the book, he takes a whole different approach when he contrasts passive and active intellect. The active intellect has no physical organ to which it corresponds (although from *De anima* III, 5, 430a24–25 it seems that the active intellect relies on the data provided by the passive intellect). Finally, the active intellect is immortal and it survives the perishable body.

Two possible interpretations can be given here. The first, sometimes called the (same) literalist interpretation, claims that passive intellect is enough to account for

---

[19] Damasio (2000).

[20] Aristotle (1907: 427); Aristotle (1955: 1, 8); Aristotle (1937: 2, 13).

[21] Burnyeat (1995).

cognition and that here Aristotle is proposing some kind of substance–dualism. The active intellect, then, is a form, and form, according to Aristotelian metaphysics, is a substance (*Metaphysics* Z, 1032b1–14). This approach seems to fault on the grounds that it is the *psuchê* and not the intellect which is the form of the body, thus the *psuchê* must be considered as the substance not the intellect alone. Besides, the *psuchê* being the form *of* the composite body is a composite substance. The event of death is the event of, literally, dis(–in–)formation of the body from or by the *psuchê*: body is no more in–formed by the *psuchê* and *psuchê*, not being the form of the body ceases to exist. The body remains, but properly speaking, it is no more a body but a corpse. The nature of the *psuchê* is to in–form the body; the nature of the body is being in–formed by the *psuchê*. Death comes when the nature of these two is present no more.

But this critique obviously contradicts Aristotle's understanding of substance as a hylomorphic unity. If form is substance, then what the human being is in its full psycho–somatic actuality is determined not by hylomorphic union but by the form alone. This is the dualist element in Aristotle's psychology which sneaks into the discussions in *De anima* III, 5. When we die, we cease to be a substance anymore and that which survives is the active intellect, not a substance but a pure actuality.[22]

The second possibility, an intentionalist interpretation, would be to regard the active intellect as a property or as an aspect of the *psuchê* and then the *psuchê* as always the form of the body. Dualism remains but it is a sort of a property–, and not a substance–dualism. In this case Aristotle would retain his method of hylomorphism in explaining active intellect just as he explained the rest of the preceding psychology. Some readings of *De anima* III, 5 tend to accept this later reading. In this case the active intellect is unmixed with the body for it has no bodily organ of which it is a form. To retain this version of hylomorphism is to remain within some version of property–dualism nonetheless.

None of the above solutions seem to be satisfactory. The problem at heart is not only the fact that we have several conflicting texts even within *De anima*, but also the problem is with Aristotle's method in general. If one remains within the methodological framework of hylomorphism in relation to the active intellect in *De anima* III, 5, then the issue seems to be unsolvable. Aristotle's hylomorphism struggles at this point of investigation into the human *psuchê*. He could not have said that active intellect serves as a form to some part of the body. At the end of *De anima* III, 4 he points out that very problem: since *nôus* thinks itself, it is the object of its own intellectual activity.[23] It must be separated and, since it has no physical, material counterpart (except the passive in-

[22] Whether 'form' is considered to be a substance in Aristotle is the subject of much debate due to the conflicting textual evidence in the Aristotelian corpus in general and incomplete explanation in *Metaphysics* Z in particular. It would be safe to assume that in the case of a human being substance is the union between form and matter, thus form would not be considered a substance except synonymously. On Aristotle's inconsistency regarding his notion of substance in relation to form and universals, see Lesher (1971).

[23] Aristotle (1907: 429a10–13, b5–5). But there are other texts in Aristotle that point to the exceptional and separate character of the *nôus*, such as in *De anima* II, 1, 413a3–7, and in II, 2, 413b24–27, as well as in Aristotle (1970: 1070a 24–26).

tellect to which it stands as an actuality), it can indeed be considered immortal. Pure actuality, separated form, yet hardly a substance. The usual methodological candidate, hylomorphism, then, cannot solve this problem.

At the same time it cannot be considered a substance–dualism (similar to that of Plato's *Phaedo*) because the metaphysics of form, that can be interpreted as substance for the reasons given above, in *De anima* III, 5 is missing. Unfortunately, Aristotle is not consistent in his definitions of substance. In the *Categories*,[24] where a distinction is proposed between primary and secondary substances, what counts as a substance is the composition of matter and form, at least in the case of a human being. However, in *Metaphysics* Z 1028–1040 it is form alone that can be considered as substance, not the compound of form and matter.

Could active intellect in *De Anima* III, 5 be considered as a primary substance which belongs to genus (i.e., animals) and species (i.e., humans) while passive intellect to the secondary substances (i.e., Socrates) as a form of the matter, accidents (i.e., white)? Under this interpretation, then, *Metaphysics* Z would refer to the primary substances alone, but the text does not confirm such an interpretation. The problem is that there is no uniform interpretation of the Aristotelian metaphysics of form and substances which can be applied to his psychology. In this case it would be correct to say then that Aristotelian metaphysics is not capable of treating this issue of the active intellect with its methodological tools. It comes as an exception, an anomaly most likely embraced by Aristotle himself, provided that the text of *De Anima* III, 5 is authentic.

To solve this problem then, Aristotle would have to either give up his metaphysical system, or deal with this problem without it. *De Anima* III, 5 would seem to be the latter case. In this short chapter we see his metaphysics reaching limits and beyond. A possible philosophy of mind likewise would have to step out, so to speak, of his methodological constraints and either invent different tools or be as it is: the shortest chapter in the Aristotelian philosophical corpus.

Whether one adopts the literalist or intentionalist interpretation, for the sake of consistency there is no way out of adding some extra elements to Aristotle's texts. As a result, there is no strictly speaking 'literalist' interpretation because a true literal but comprehensive reading would be inconsistent and contradictory. Any coherent interpretation of *De Anima* III, 5 adds something to the text; 'literalist' interpretations are no exceptions.

In an attempt to respond to the epistemological and ontological questions with Aristotelian psychology it must be said that mind (active intellect) and the world are separate, indeed unmixed. Contemporary philosophers of mind have used Aristotelian psychology with a 'naturalized intellect' (viz. embodied intellect) assuming the roles of both passive and active intellect, thus bypassing *De Anima* III, 5 altogether. First, this interpretation is not of Aristotle, not what is offered in his texts. Second, his psychology had very different goals from that of the contemporary philosophy of mind. Active and passive intellects are explained in relation to their activities and in relation to the *psuchê*, as the principle of life in general. Thus, on the nature of the mind and on the

[24] Aristotle (1975: 2a 13–4b19).

nature of the world, what is asked here and what Aristotle asked are incompatibly different.

The solution proposed here is as follows. Aristotelian metaphysics, just as his epistemology, is foundationalist in nature. Naturalist interpretations of Aristotle, in fact, have no need for the active intellect in order to explain any cognitive activity of the mind. In the same way, Aristotelian hylomorphism in its naturalized version would gladly go from *De Anima* III, 4 to III, 6 bypassing III, 5 altogether. This matter/form, potentiality/actuality dynamics can be interpreted in a non–foundational manner that could be adopted by functionalist models of the mind without any need for foundational metaphysics.

Active intellect is that formalization of the passive intellect being the form of the lower psychological functions and the *psuchê* being the form of a body in general. One might interpret active intellect as something abstract and non–substantial without any reference to the body,[25] or as a divine principle of a complete understanding only in reference to which our incomplete human knowledge makes sense.[26] This is open to interpretation, but from Aristotle's texts at hand we can only say that active intellect plays an important foundational (and formational) role in the overall Aristotelian metaphysics of hylomorphism, which for contemporary metaphysicians and epistemologists appears rather redundant.

Aristotelian *nôus* is not fully naturalized, nor is his psychology completely dualism–free. All his doctrines, from logic and science to psychology and metaphysics were in–the–making rather than complete, dogmatically accomplished philosophical notions, where naturalized *psuchê* and dualism–free metaphysics was a goal, not an achievement.

This interpretation can be supported by the similarities one finds in the theories of knowledge in Plato and in Aristotle. As Hintikka emphasized, both considered that the only genuine knowledge that the mind can have is that of the immutable, eternal objects.[27] For Plato, Forms are what provide the mind with a true knowledge.[28] Likewise for Aristotle genuine knowledge is possible only of the unchanging realities.[29] Knowledge is genuine only when its objects are never changing which is when they are always true.[30] Thus, it is the active intellect that has an access to this kind of knowledge. Indeed, it is *only* the active intellect that *can* have such knowledge, for it is only the active intellect that can know the unchangeable and eternal. And because the active intellect has this unique access, it "makes all things" in the passive intellect and the *psuchê* in general.

The ontological question of existence in Aristotle's psychology stands as an alternative to the two extremes of the pre–Socratic forms of materialism and Platonic dualism. As we have seen, the method of approaching this divisive problem for Aristotle was

[25] Wedin (1995: 190).
[26] Caston (2006: 341).
[27] Hintikka (1967).
[28] Plato (1926a: 135 b–c; 439d–440c).
[29] Aristotle (1934: 1139b20–23).
[30] Aristotle (1933: 1039b27–1040a5).

rooted in his theory of hylomorphism and the four *aitia*. However, while successfully avoiding both reductive materialism and substance dualism, Aristotelian psychology seems to commit itself to some version of property dualism on the level of the explanation of perception. Although his notion of substance remains ambiguous, his theory on the intellectual powers remains dualistic of some sort.

However, these are metaphysical problems that remained not clarified by Aristotle. Nevertheless, Aristotle's clarifications of the relation between the different levels of the *psuchê* in Plato and his naturalized psychology on the levels of perception and cognition even now attract many philosophers of mind. One might doubt whether Aristotelian psychology solves the ontological problem but to say the least, it remains an attractive alternative to substance dualism, even for the post–Cartesian philosophers of mind. Thus, the question on existence is answered with the naturalized *psuchê*, while the question on the nature of that *psuchê* depends on the not–clarified Aristotelian metaphysics of substance.

Debate on the theory of perception is fundamental in answering the epistemological question. As I demonstrated above, the outcome of this debate will depend on explaining the metaphysical status of the medium of perception. Here the affinity between Aristotle's epistemology and semantics and their reliance on his metaphysics is clear. The relation between the perceiving organ and what is being perceived relies on his theory of matter and form. In turn, the matter and form are related to each other in terms of analogy and not in terms of some univocal identity. The correspondence between the organ and the perceived, knower and the known, word and object will rely on the notion of analogy. Not surprisingly, when in the philosophy of mind the question of identity between the brain and the mind will emerge, it will rely more on the analogical correspondence than on the univocal identity.

For Aristotle the question of identity will not have any part in his psychology. However, this is only true indirectly, for his psychology itself significantly relies upon his notion of analogical predication. Thus, the expressions of the meaning, the correspondence between perceived and the perceiver, are established on analogical terms. However, for the proponents of naturalism, analogy lacks in precision of explanation and reverses our questions, again, to the ontological status of the analogically related terms and objects.

1.2 From *Psuchê's* Substance to *Anima's* Subsistence: Aquinas' Subsistence–Dualism

For Aquinas, the soul (*anima*) is the incorporeal, subsistent principle of intellective activity[31]. He follows Aristotle's psychological hylomorphism by considering soul likewise the form of the body, which makes possible its activities. This form is the substance of the human person, but a human person is not identical with his/her soul. In-

[31] Aquinas (1889: 75. 2c).

stead, we are to speak of a human person as a composition of soul and body, contrary to Boethius' definition of a person as "*rationalis naturae individua substantia*".[32]

Soul in relation to the body is a subsistent entity and, at the same time, a substantial form. However, there can be no more than one substantial form in one being. Nutritive and sensory forms are only accidental to the totality of the human being; they do not define him/her. Thus, these accidental forms consider being in its actuality (I am actually seeing, I am actually touching) and in its certain respect (i.e., subject of perception), while the substantial form regards being in its potentiality and regards it absolutely. This allows Aquinas to regard substantial form as identical to the intellect which orders the whole body in its totality, in all its functions.[33]

The soul as a substantial entity is subsistent in relation to the body, and the body is in–formed by the soul although not in its parts but in its totality. Unlike the body, the soul can exist without its body. Yet, after the separation of the soul from the body we do not speak either of the body properly (instead the body is really a corpse)[34], or of the person in relation to the separated soul. The soul without its body loses its memories dependent on images.[35] In this Aquinas likewise follows Aristotle, for whom a dead body is similar to a statue of a body, which looks like a real body but is not.[36]

In commenting on this psychological theory in Aquinas, some authors consider it to be a substance–dualism.[37] Others speak of a non–materialistic realism.[38] Anthony Kenny acknowledges that Aquinas was able to avoid the extremes of reductive materialism and substance–dualism, but points out his insufficient explanation of the relation between soul and body.[39]

Just as in Plato and in Aristotle, Aquinas' philosophy depends upon his metaphysics of categories. For Plato and Aristotle the soul as a form is the substance of the body which is a uniform collection of accidents. Where Aquinas departs from Aristotelian metaphysics is his definition of being as such in the case of the soul: the soul is not a substance but subsistence. Just on this ground Aquinas cannot be considered a substance dualist. His notion of a soul as a subsistence is more dynamic than the Aristotelian substance, though it too employs the soul/body hylomorphism.

Aquinas avoids both substance–dualism and materialism, but he remains within a large Aristotelian philosophico–linguistic framework of a kind of property–dualism one might coin as a *subsistence*–dualism. For someone who looks to Aquinas to solve the Cartesian mind/body problem this subsistence–dualism will be of little if any help, just as in the case of the property–dualism of Aristotelian hylomorphism.

---

[32] "*Persona proprie dicitur rationalis naturae individua substantia*". "A person is property called an individual substance of a rational nature", in Boethius (1918: 37).

[33] Aquinas (1984: 412a15); Aquinas (1889: 76. 3, 220–223; 76. 4, 223–227; 77. 6, 245–247).

[34] Aquinas (1889: 76. 8c, 232–235).

[35] Aquinas (1889: 89, 370–384).

[36] Aquinas (1984: 412b25–26, 75); Aquinas (2000: I, ad 6).

[37] Hoffman (1990); Swinburne (1997: 306 n.9); Stump (1995).

[38] Pasnau (2003: 65–72); Kretzmann (1992).

[39] Kenny (2004: 145–159).

Aristotle's psychology was the conclusive treatise of his physics that implemented the key methods to be applied in metaphysics and in ethics. Its main goal was to explain life, differentiate between living and non–living beings, and to avoid both atomistic materialism and substance–dualism. *Psuchê* was the explanation and hylomorphism was the method. Aquinas' psychology seems to be even more fragmentary and intermediary between his metaphysics of subsistence and his subsequent theology of personhood (or perhaps introductory to his anthropological treatise). Neither of them had the Cartesian problem in mind; the philosophy of neither of them could even accommodate such problem. Some could see it as strength, others as weakness, but the Cartesian problem was absent from their philosophies. Finally, epistemological issues were never central to Plato, Aristotle or Aquinas as they were in Descartes.

In relation to Aristotle's naturalized psychology, Aquinas' unification of the soul as one and his insistence on the mind's governing abilities over the soul as much as the body, makes his doctrines more useful for theological anthropology (for which *Summa theologiae* I 75–79 were written as an introduction to his 'Treatise on Man') rather than for the contemporary philosophy of mind. His notion of the human person as subsistence is superior to that of predecessors' substance, but it can only be unproblematic for theologians for whom hylomorphism appears to be a solution for the substance–dualism of Plato or Descartes.

The nature of mind in Aquinas, as mentioned by Kenny, is problematic when considered from bottom–up in the biological relation of the body (but not just the brain) to the all–governing mind. Finally, clear dependence on the Aristotelian ontology of categories in his psychology puts Aquinas's philosophical system at odds with his superior anthropology in general and metaphysics of subsistence in particular. Again, perhaps from a theological perspective these problems are less perceptible and less problematic than when considered under the analysis of the contemporary philosophy of mind.

Aquinas' unification of the Aristotelian *psuchê* under its intellectual powers had more advantages for his theological anthropology than for any future philosophy of mind. How the *anima* would govern not only cognitive functions but also our physiology remained unclear. This was a significant departure from Aristotelian attempts to have a naturalized account of the *psuchê* that would not only be limited to the human being but was also meant to explain all life.

This turn from the living *psuchê* to the thinking mind or intellect *somehow* governing the rest of the body started with the works of Aquinas. The question of the animals seemingly sharing our perceptual powers was confronted by making more distinctions between sensory perception unique to humans (*vis cognitiva*) and those of the animals (*vis aestimativa*).[40] Again, the doctrine of participation resembling Aquinas' theory of subsistence, comes from an elaborated much more basic notion of analogy from Aristotle et al. However, similarity between perception in animals and humans cannot be explained by any participation *if* perception in humans depends on the intellect governing all which, obviously, in animals is absent. Before this ontological, most basic, problem is somehow resolved no '*mutatis mutandis*' would have any chance in the questions of epistemology and meaning.

---

[40] Aquinas (1889: 78. 4; 81.1).

But epistemological questions are part and parcel of Aquinas' ontology and metaphysics. Our knowledge of our own intellectual powers depends not on some basic consciousness–independent non–intellectual 'powers', but on the intellect itself in addition to the participation in the divine being by likeness and similarity to it. Cartesian circularity takes its roots from this anthropology (and from St. Augustine, as we will see). This circularity remains unsolved unless one assumes foundational structures originating from a theistic belief.

The semantics of analogical discourse was elaborated by Aquinas with a result being the referential theory of meaning. This theory will be explained in the following chapters at length. Here it will suffice to notice that such theory, where the object and designation stand in an immediate relation, originates from the ancient notions of analogy, historically rooted in mathematical analogy.

Aquinas' departure from the Aristotelian naturalization of the *psuchê* would be an attractive point for Descartes after his earlier philosophy of science would make space for his later metaphysically–minded one. There he will make the gap between human cognition (Spirits) and animals' perceptive functions (machines) even greater, finally resulting in the explanatory gap between mind and body.

2. The Second Psychology: from *Anima* to *Mind*

2.1 Substance Dualism Naturalized: Descartes

If metaphysics played a central role in Plato, Aristotle and Aquinas, and if notions of the substances and categories were essential to their inquiry into *psuchê* and the nature of human knowledge, for Descartes it was the other way around. Cartesian epistemology defined his metaphysics: our way of knowing the world determines our discourse on the world.

Euclidian geometry determined the Ancient Greek sentential, categorical logic and language by way of abstract forms resembling real objects in the physical world. Whether these abstract forms belonged to the domain of the human active intellect and depended on perception and learning, or they belonged to a separate realm and were learned through recollection, was to be determined.

In Descartes, on the contrary, geometry resembled human thought and subsequently it was projected upon reality. Thought and the external world were two different and distinct realities operating by two different sets of laws and rules. The major difference between Plato and Descartes is in their understanding of what was real and what qualifies as substances.[41] For Plato, the real substances are Forms. Forms were the only true substances in the world and through their likeness with the physical objects they emanate their true nature as if in some sort of revelation (hence the frequent use of myths in Plato's dialogues). The laws in the universe were essentially the same, those of the

[41] Broadie (2001).

Forms. Daniel Devereux speaks of an epistemological as well as an ontological separation between the Forms and their counterparts in the physical world.[42]

The distinct characteristic of Descartes' notion of the mind (*mens*) is that it was conceived in terms of consciousness in relation to the body. The body then was considered to be a machine, being a composition of elements like geometrical figures, atoms in the void. According to John Haldane and Edwin McCann, his dualistic 'philosophy of mind' was a consequence of his more general preoccupation with substituting ancient and medieval notions of hylomorphism with the vision of the mechanical philosophy of science according to which both, mind and body, are distinct entities, substances, on their own terms.[43] If the *psuchê* is no more the principle of life in all living systems from plants to humans, then the New Science would be in charge of the study of matter; philosophy and theology would be occupied with the study of the mind.

Although Descartes' epistemology ultimately determined his metaphysics, it was the scientific questions of the early 17th century that provoked metaphysical problems for which epistemology provided the theoretical basis of the explanations. Tom Sorell speaks of two philosophies of science in Descartes.[44] The first takes experience to be the source of knowledge and certainty about the observed world. From the observed effects we can prove the principles by which the world exists. This philosophy of the biologist and physicist can be found in his letters and in parts of the *Discours de la method* (1637), of the *Principia philosophiae* (published in 1644) and of the *Essais* (1637).

The second philosophy is that of a metaphysician that Descartes proposes after his scientific arguments in the *Discours de la méthode* were not recognized by the scientific community of his time. In this second period (after 1637) Descartes provides a "demonstration of the principles of physics by metaphysics"[45] and the unproved principles in the *Essais* and in the *Méthode* were supposed to be now proved in his *Meditationes de prima philosophia* (1641).

Philosophers of mind tend to miss this first period in Descartes' more scientifically informed period of philosophy and look into the *Meditationes de prima philosophia, cum obiectionibus et responsionibus,* and in *Les passions de l'âme* (1649) as the main source of his dualistic account of the world. However, anyone who looks into his earlier writings, such as *L'Homme* (1633), would find a study of human being even in his most complex psychological activities by studying his body alone.

In the pre–*Meditations* period Descartes elaborates his metaphysics of mind and body. In the *Regulae* Descartes makes two important distinctions between simple and complex and between intuition and deduction. Simple is a thing that is the most intelligible to the mind, such as the self and God are. The mind goes from simple to complex things and propositions in its process of understanding. In parallel, intuition is simple because it is instantaneous, while deduction is based on intuition, but it is a more complicated vision of the mind.[46]

---

[42] Devereux (2003).

[43] Haldane (1995); McCann (1995).

[44] Sorell (2000).

[45] Descartes (1996: X, 134).

[46] Descartes (1996: X, 366–370).

Meditationes provide the two metaphysical principles upon which observations can rely: that my thinking proves my existence and that God exists and is no deceiver. The second principle must provide the ground for trusting our senses and the first principle is supposed to prove God's existence. Evidently, an atheist would have no grounds to trust his or her senses on Descartes' account.

The First Meditation calls into doubt all beliefs, even the most evident. The Second Meditation begins by giving the first principle, that thinking implies personal existence. This principle is based on skeptical doubt: even if we accept the evil genius hypothesis that we are deceived about everything, the mere act of *accepting a hypothesis* (which is thinking) proves one's existence. In the rest of the Second Meditation Descartes asserts that since thoughts cannot exist without a thinker, the mind which thinks is a substance and that it must be a thinking thing, whose essential states are thinking and being conscious. At the end of the Second Meditation, by the force of the example of the wax, Descartes contrasts the two things, the two substances as: 'thinking', beyond space, whose essence is to think, to be conscious; and as 'extended', whose essence is extension in space.

Just as the same piece of wax can be hard and cold, liquid and hot, any physical body can go through changes of form, color, smell, etc. What remains unchanged in the physical body *qua* physical body is its extensiveness. That belongs to the essence of all physical reality. Likewise, since minds can be deceived by the physical senses, we might not distinguish dream from reality. What persists regardless of any changes is the act of thinking. This act belongs to the nature of the mind, being the thing that thinks. But, as we will see with Wittgenstein, it is rather the other way around: thoughts, ideas, concepts, feelings and emotions change quicker and more often than any physical object of our perception.

Meditations Three, Four and Five pave the way for Descartes to go from conceiving a mere possibility to actually affirm a real possibility. In these three Meditations he gives his ontological argument for God's existence, while in the last, the Sixth Meditation he makes the move from possible to actual. The second metaphysical principle of the existence of God serves as the proof that the mind and the body really exist, and with them the rest of the world.

Cartesian physiology is mechanistic, but not in the same way in which machines and clocks is mechanistic. The human body operates by infinitely many little streams, constantly in flux, through blood and other bodily liquids communicating through its members. Dennis Des Chene speaks of a "fluid mechanics" of the human body.[47] It is through this mechanistic explanation of the living system that Descartes reduces Aristotle's biological functions, explained by the presence of the *psuchê,* to the complexity of these "fluid mechanics". The operations of the sensitive *psuchê* are explained as mere instincts. The result of this reduction is the elimination of the basic vegetative part of the *psuchê* without which, it seems, we are left with a perfectly dualistic account of the person: mind and body.

For Descartes, physical objects are real substances but so is the human mind, and different laws rule both. The human subject is the only exception in which both worlds

[47] Des Chene (2001: 38–40).

come together and constitute some harmonious unity of this twofold composition of mind and body, of the *res cogitans* and of the *res extensa*.

As it was mentioned in connection with hylomorphism, Aristotle believed that how matter operates depends on its form. Matter is determined not by its own laws but by its form: the human being has that particular kind of a body because of its particular kind of form, which is its *psuchê*. Unlike the atomists, Aristotle did not think that matter is already determined by the way the atoms are arranged; its mere atomic composition would not explain its particular bodily composition.

Rejecting both Aristotelian hylomorphism and atomism, Descartes embraced mechanism for matter: bodies operate according to their own laws and these laws are mechanistic. 'Mindless bodies' (which included all animals except humans) operate on their own; 'mindful bodies' operate by the movement of the mind which pulls levers of the bodily parts. Somehow the two are connected and Descartes suggested that pineal gland is that place where mind and body interact. His experience with autopsies of the human brain helped him to determine that the pineal gland is the only part of the brain that is not duplicated in the two parts of the brain. It stands, so to speak, alone and in the slightly lower part of the middle of the brain. At the time, it seemed to be a suitable candidate for a physical explanation of the consciousness meeting bodily organ: it is unduplicated and it is in the middle. Yet, the mind is not equated with the brain: it lacks spatiality and it survives the death of the body.

By disavowing both Aristotelian hylomorphism and materialism, Descartes' only choice within substance ontology is dualism. And since Descartes was Christian, his notion of mind and body resembles very much an Augustinian version of dualism. Indeed, his contemporaries had indicated to him that the *cogito* argument and the consequent ontological argument for God's existence resembled a lot St. Augustine's assumption that even misconceived ideas prove one's existence,[48] or St. Augustine's insistence on the separation of body and mind.[49]

Cartesian interactionism answers the question about what connects the mind and the brain: mental states causally interact with the brain states. Whenever there is a causal interaction from the mental to the physical, the body acts in its members. Whenever there is an interaction from the physical to the mental, we have the process of perception.

If matter is seen to operate by mechanistic laws, the mind has its own laws of reason. In a way Descartes closes the problem of life in relation to matter by considering all matter as more or less a complex mechanism. He solves the problem of the mind by stating its laws as well and he indicates where the two interact in a human person. What comes clearly as the major problem in the *Obiecta et responsa* in the *Meditationes*, however, is the question of *how* the interaction occurs. His replies transform the question of *how* into the question of *where*. The Cartesian answer is that the interactions are brute facts about our bodies and minds.

The struggle with Aristotelianism and its consequent refutation in the 16th–17th centuries in the context of the scientific revolution resulted in the slow evolution from the

[48] "Si fallor sum". "If I err, I am", in Augustine (1955: XI, 26).

[49] Augustine (1968: X, 10); Augustine (1962: 39, 73).

Ancient and Medieval *psuchê/anima* to the Modern *mens/mind*. The slow process came also as a result of not seeing the soul as the principle of life, as that which distinguished living from non–living bodies. Life eventually began to be interpreted in mechanistic terms; the notion of mind seemed to be a better candidate than the soul as an explanation for intellectual activity. What was previously meant by the *psuchê*, now in part was attributed to the body, in part to the mind. As a result, a new kind of substance dualism with a new formulation of problems concerning mind and body emerged.

Descartes did not answer questions concerning the nature of the mind and the world. Their *substantial* differences would allow for some interaction, but without an essential similarity between the two, interaction in separation is all one can hope from Descartes. In the *Objections and Replies* however, he did not indicate any problem of lack. They aim, rather, at the specificity of the Cartesian metaphysics defined by his epistemology (the two principles) that itself would not allow for a genuine unity between mind and the world.

If Descartes had continued his first philosophy of science project, we might have had a whole different epistemology defined by a whole different kind of metaphysics and ontology. For historical (the Galileo case) and intellectual (unsuccessful scientific claims of 1637 writings) reasons, we do not have a Cartesian philosophy in which questions on the nature of the mind and questions on the nature of the world come in harmony. Instead, Cartesian dialectics comes with the price of a lack of answers to naturally asked questions of interaction, causality and harmony between the mind and the world.

The Cartesian definition of the mind–body problem settled the philosophical background for the next two centuries. In many ways, his dualistic language is in use even today. Most of the 17th – 18th century psychological philosophies can be regarded as a series of responses to his question of how to interpret the uneasy relationship between mind and body.

Among the main responses were the dualistic parallelisms of Leibniz (preestablished harmony)[50] and Malebranche (occasionalism),[51] and the monistic parallelism of Spinoza[52] and Berkeley[53]. While different in ontology and epistemology, these theories share the same foundationalism in the philosophical theology of Descartes: God as a guarantor of knowledge, perception and existence. For this reason these theories had been regarded as an historical elaboration of Cartesian philosophical psychology rather than new projects on their own. These theories have little, if any, application to the contemporary philosophy of mind.[54]

One exception is Thomas Hobbes' materialism, based on his mechanist ontology with important affinities to his semantics and political philosophy.[55] Many similarities

---

[50] Leibniz (1991: §17, §§78–79); Leibniz (2010: §60).

[51] Malebranche (1711: I, i–ii, 8–12).

[52] Spinoza (1846: I §§1–15).

[53] Berkeley (1837: 1.25,12; 27.288–38.289; III: 70–72).

[54] Davidson's 'anomalous monism' might be an exception. He did refer to Spinoza's notion of causality with the consequence for Davidson's denial that there are psychophysical laws. Davidson (1999b).

[55] Hobbes (1839: IV, ch. 25, 12, 406–408; III, ch. 1: 1–3).

can be brought between Hobbes and contemporary eliminative materialism, although a significant correction must be made in relation to the differences in the scientific world–view of Hobbes and today's science.[56]

Due to the lack of space and this work's limited purpose, an analysis of the above theories mentioned *en passant* cannot be given here. However, any such study would have to take into account the importance of the ontological background of these philosophers, as well as the significance of their historical affinity to Descartes and his philosophy of mind and body.

2.2 Locke: Many Bodies, One Identity

Locke, Hume, and Kant, notwithstanding great differences in their philosophical psychologies, had one question in common: What makes the person's identity? Their responses were based on their different ontological commitments which will be examined at some length in this last section of this chapter.

Locke's writings on the nature of mind and body come as a part of his theory of identity (What makes the same man?[57]) that was a theory, among other things, to facilitate some of the theological and political issues of his time. Locke postulated two theories of identity.

One is an identity determined by the unity of substance. The identity of matter is determined by the unity of atoms in each part of the matter. The same atoms make up the same substance, hence the same matter.

In the second theory, identity is determined according to what category of thing an individual belongs. This identity which is the "of the same kind" is solely the product of our intellect.[58] This idea of the *kind* determines that which used be called 'substance' by Aristotle. It is that which it is to be a horse or an oak tree. The material change in this second identity is gradual, which accounts for its different parts (i.e., horse's members or tree's branches, etc.).

The human being according to this second theory of identity fits into the category of 'person'. Locke goes on to define a person as: "A thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing, in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it".[59]

In this clearly post–Cartesian definition of personhood the main accent is on the consciousness of the "thinking thing". What about the body? How does the body relate to the conscious thinking thing? In Book II Chapter XXVII, §15, Locke provides a thought experiment in which a prince and a cobbler switch their bodies, in which case the prince in a cobbler's body would have his own consciousness. Yet, no one could say that it is still the same person, the prince:

---

[56] An impressive work on Hobbes' philosophy of mind was done by Pettit (2008).

[57] Locke (1975: II, 27 §16).

[58] Locke (1975: II, 27 §1).

[59] Locke (1975: II, 27 §9).

> For should the Soul of a Prince, carrying with it the consciousness of the Prince's past Life, enter and inform the Body of a Cobler as soon as deserted by his own Soul, every one sees, he would be the same Person with the Prince, accountable only for the Prince's Actions: But who would say it was the same Man? The Body too goes to the making the Man, and would, I guess, to every Body determine the Man in this case, wherein the Soul, with all its Princely Thoughts about it, would not make another Man: But he would be the same Cobler to every one besides himself.[60]

This personhood is the basis of Locke's criticism of Descartes' but also of Hobbes' ontologies. The identity of a human person does not coincide with his soul (or mind) or with his body. There is no immaterial substance that can account for the identity of a human being, no more than a physical substance can. Instead, Locke considers *consciousness* to be that unique mark which makes humans as such. Which body it takes, ultimately is not the issue: "But yet when we will enquire, what makes the same *Spirit*, *Man*, or *Person*, we must fix the *Ideas* of *Spirit*, *Man*, or *Person*, in our Minds; and having resolved with our selves what we mean by them, it will not be hard to determine, in either of them, or the like, when it is the *same*, and when not".[61]

The prince switching his body with the cobbler and Locke's similar examples serve another purpose. The issue of the resurrection of the body was one of the problems that Locke tried to elucidate with his theory of personal identity. God may choose a different body for someone's consciousness on the day of the resurrection and as long as the personal memories of that consciousness are preserved, we can speak of the same person:

> And thus we may be able without any difficulty to conceive, the same Person at the Resurrection, though in a Body not exactly in make or parts the same which he had here, the same consciousness going along with the Soul that inhabits it. But yet the Soul alone in the change of Bodies, would scarce to any one, but to him that makes the Soul the *Man*, be enough to make the same *Man*.[62]

There is another theory that Locke aimed to eliminate: Aristotelian hylomorphism and the Scholastic identification of human personhood with the soul as the substance. Perhaps, just as there is no immaterial substance or form in charge of the body to be identified with the human being, there is neither God's chosen monarch to which people are *naturally* subjected. Instead, we have consciousness as a collection of free, personal states and memories. A different psychology and moral philosophy presupposes likewise a quite different political theory.

Locke's theory of personal identity received many criticisms from the early critics, such as Leibniz, Berkeley, Hume, Butler and Reid, as well as from contemporary phi-

---

[60] Locke (1975: II, 27 §15).

[61] Locke (1975: II, 27 §15).

[62] Locke (1975: II, 27 §15).

losophers. For instance, in the words of Butler, "one should really think it Self–evident, that Consciousness of personal Identity presupposes, and therefore cannot constitute, personal Identity, any more than Knowledge in any other Case, can constitute Truth, which it presupposes".[63] This is just one of many earlier objections known as 'the circularity objection'.

Recently the theory has been criticized for undermining the physical aspect of consciousness,[64] but even more for identifying human beings with their states of consciousness as self–awareness, in which case, i.e. memory loss would have to imply at least some loss of personal identity.[65] Furthermore, the notion of 'personhood' extends beyond just human beings, i.e., to fictional characters and Trinitarian persons.[66]

Locke's psychology presents, however, one attractive point for the contemporary philosophy of mind. It is his disavowal of the classical notion of substance, whether physical or immaterial, in treating human consciousness which can be interpreted with modifications as a dismissal of centers of consciousness in neuroscience.

For Locke the human mind is much more than the collective personal consciousness, but it is this particular mark which makes the relationship between the mind and the world personal. Consciousness is an awareness of one's own mental contents: "the perception of what passes in a Man's own mind".[67] Although mind was not considered in the Cartesian terms of a substance, Locke's abandonment of the classical notion of substance did not help him to explain in what ways the conscious mind is related to the world. If consciousness is that which makes the mind to know the world (but then what about animals' knowledge without self–consciousness?), how is perception of the physical possible for the immaterial?

Locke's explanation is related to his notion of ideas as phenomenological entities, sense–data that can at times even have shapes and colors. Locke was the first to suggest that color spectrum inversion be applied as an argument for the qualia in the contemporary philosophy of mind.[68] His argument is that people, who are born with reversed color spectrum, when seeing violet perceive the same color and are acquainted with the same sense–datum as people without reversed color spectrum see marigold.[69] What is responsible for perception is not the body, just as it is not the body that decides on one's personal identity, but the collection of ideas, the mind.

Locke's metaphysics, and hence epistemology, remained dualistic, in many ways similar to that of Descartes. Or rather, it was consciousness' job to have that relationship but consciousness itself remained defined in purely mental terms.

---

[63] Butler (1852: 301).

[64] McCann (1999); Olson (2010).

[65] Noonan (2003: 23–44); Shoemaker (2012).

[66] Flew (1951).

[67] Locke (1975: II, 1 §19).

[68] I.e., Shoemaker (1984: 357–381).

[69] Locke (1975: II, 32 §15).

2.3 Hume: Patterns of Necessary Connections

Hume is credited with having a bundle theory of mind according to which the mind is a bundle of impressions and ideas. (A similar bundle theory is already present in Locke's *Essay*, together with his general criticism of the notion of substance on the basis that it lacks empirical content.[70]) This bundle theory together with discarding the notion of substance permeated Hume's thought not only in his treatment of the mind and ideas but generally in his metaphysics. In fact, his treatment of the mind and ideas comes under his general treatment of causation in the world.[71]

An important difference between Locke and Hume comes in their interpretations of ideas. Locke did not make a clear distinction between impressions that are formed by perceptual experience and ideas which are formed on the basis of the impressions. Hume's *Treatise* begins with this methodological distinction and further analysis of both.

Hume insists that prior to any thinking there must be some material for thoughts to arise. This material is supplied by perceptual experience, which comes either as an impression or as an idea. Hence, both impressions and ideas for Hume fall into the category which Locke called 'ideas'.

The difference between impressions and ideas depend on the force by which they enter the mind. Impressions have the greatest force and they include sensations, passions and emotions. Ideas are said to be "faint images" of thinking and reasoning.

There are two types of impressions and ideas: simple and complex. Simple are those that cannot be analyzed any further; complex are composites of simple ones. Simple impressions are color, taste and smell that are attributes of, i.e., an apple. An impression and an idea of an apple is complex, containing several simple impressions and ideas.[72]

Impressions and ideas correlate with each other. Every simple idea resembles a simple impression and vice versa. This, according to Hume, can be made known through a direct inspection of the mind. Yet, this correlation cannot be proved because one can only be acquainted with one's objects of the mind by inspecting his or her own mind.

The order of correspondence is always from simple impressions to simple ideas. This would prove that impressions cause ideas in the mind. Thus, everything that is qualified as the mind's content originates from (simple) impressions, which leaves no room for innate ideas. In this Hume is in agreement with Locke again. As a result, all mental activity must be because of perception in the form of an impression or an idea.

---

[70] Locke (1975: II, 32 §6, III, 2 §20, III, 5 §14, III, 12 §12).

[71] In line with McCann (1995) and Broughton (2005: 43), but contrary to Biro, for whom "For Hume, understanding the workings of the mind is the key to understanding everything else. There is a sense, therefore, in which to write about Hume's philosophy of mind is to write about all of his philosophy". Biro (2006).

[72] Biro (2006).

What comes as knowledge is a collection of the complex ideas that mind, by habit, unites in bundles.[73] Hume's controversial theory of causality explains the mind's role in assembling ideas together.

Section VII of D. Hume's *An Enquiry Concerning Human Understanding* treats the idea of necessary connection. The text proceeds with an apparent *crescendo*: from a comparison between the clarity of mathematics in contrast to the obscurity of the moral and metaphysical sciences regarding the basic ideas of reasoning and of forming necessary connections, to the accounts of causality, and to his instruction on how one must construct the discourse of the necessary connection of cause and effect. This account of the necessary connection will provide the right method of philosophical inquiry, which will be treated in the subsequent sections of the *Enquiry*.

This textual construction begins from the most basic, if not common, observations of how mathematical reasoning differs from that of philosophy. Hume immediately finds mathematical reasoning insufficient when it comes to explaining how necessary connections come about. Geometrical axiomatisation affirms consequences on arithmetical bases but gives no general explanation of how these connections occur in the real world.

Hume turns to what Barry Stroud calls "the negative phase" of his account of causation and inductive inference. His negative argument shows that our causal expectations are never formed on the basis of reason alone. Hume asserts that by mere observations of bodies we are not able to "discover any power or necessary connexion".[74] But our reason cannot provide such a connection either. Those philosophers who have recourse to such explanations see every movement of the will as an occasion for divine action,[75] which in Hume's analysis is no explanation at all. For Hume, the mind's causal reconstruction of events in the world, viz. from perception to imagination, cannot be explained in reference to the relation of ideas but must be a matter of fact.[76]

All we are left with is to acknowledge that causation is an associative relation in our constant observation of objects. But that too is an inexact and obscure metaphysical discourse, in need of being specified and purified. In our examination of the single cases of events, we regard them as causally related. In fact, all our impressions are built on their conjunctions. But when we take a single case in isolation from all others, we see no such connection. Hence, the mere repetition of these conjunctions provides the connections for us. An explanation of this odd process brings Hume to the positive phase of his account for causation.[77]

The mind acquires the habit of perceiving single events and expecting the usual consequences. The *feeling* of a transition from one event to another determines the facts as our own interpretations of the events. But there is nothing further in the case: these connections are no more than causal inferences of our mind. These experiences of mak-

---

[73] Locke (1975: I, 1 §4).

[74] Hume (1975: VII, 1 §63); Stroud (1977: ch. 3).

[75] Hume refers to Malebranche's occasionalism. Hume (1975: VII, 1 §70).

[76] For Kant, if the causal principle is not analytic, then it must be synthetic (A189/B232). Following the interpretation of Bayne (2004: 26–34).

[77] Hume (1975: VII, 2 §74).

ing necessary connections occur only because of the sensations (the external impressions)[78] of the perceived situations, together with our background knowledge and memories.

This brings Hume to provide two definitions of cause. The first definition specifies cause as "*an object, followed by another, and where all objects similar to the first are followed by objects similar to the second*".[79] This definition accounts for all the *external impressions* indispensable for our knowledge acquisition. The second definition indicates cause as "an object followed by another, and whose appearance always conveys the thought to that other".[80] This second definition focuses on the *internal sensations* in our understanding the perceived as an event. These two definitions, seen as separate, provoked much controversy in later literature.[81] The confusion perhaps comes from the tendency to label Hume as an empiricist. But Section VII of his *Enquiry* shows his attempt to integrate the external impressions with the internal sensations that is usually ascribed to Kant.

Which of the two definitions is truly Humean? The answer is: both. It is at this point that we reach the highpoint in Hume's account of causality. This highpoint is the conjunction of the two definitions, for only together can they rightly account for how the mind is capable of bridging sensations and impressions in causal connections.

The rest of the *Enquiry* is built upon the program spelled out in the first sections. The right method of philosophy consists in always providing the reason for the experience, the same experience that "teaches us the nature and bounds of cause and effect, and enables us to infer the existence of one object from that of another".[82] Hume's skepticism is a cautious attitude toward all *a priori* statements, since these are always formed upon limited experience.

Finally, Hume's conception of causes tells us that we always observe adjacent and successive events, which results in his theory of meaning according to which all concepts must be grounded in experience. Hume's rejection of metaphysics (especially in sections 1 and 12 of the *Enquiry*) seems to be incompatible with this notion of meaning grounded in the human mind always in need of the confirmation by experience. This Humean naturalism, combined with his acceptance of causality (as an essentially metaphysical principle) can only be justified by his naturalistic rationalism. In other words, the mind accepting the principle of causality must be regarded itself as an empirical fact.

Hume and Locke are credited with, what is called, "the bundle dualism": mental states without a subject. If we take a closer look at their metaphysics and philosophy of science first and then reflect on their notions of the soul and mind second, and only as the consequence of their metaphysics, we see Hume's and Locke's endeavor against

---

[78] The distinction is borrowed from J. Locke: distinction between *sensation* and *reflection*. The first are the outward senses (sight, etc.) and the second are the inward senses (self–awareness). Millican and Beebee (2007: 163–199).

[79] Hume (1975: VII, 2 §76).

[80] Hume (1975: VII, 2 §77).

[81] Dicker (2001: 110–116).

[82] Hume (1978: XII, 3 §164).

dualism as an explanation, striving for a unified theory of ideas and the mind's operations.

Bundle dualism goes in hand with associationism, a view that causality in the world is explained by building associations between perceptual evidence on the one hand, and thoughts and ideas on the other. This will be used by metaphysical behaviorism in explaining human behavior, although without any reference to the mental contents of one's experience.

2.4 Kant: The Ineffable Soul

Kant had approached the theme of mind and body at different times with different methods and results. One common theme that goes through his philosophy of mind in every period is, just as for all of his predecessors, the centrality of metaphysics and the importance of the philosophy of mind for epistemology in general.

Kant's writings on the mind can be easily divided into before and after his critical philosophy.[83] In the pre–critical period, two topics were central to Kantian metaphysics: freedom and self–consciousness. According to Kant himself, this period gave way to his rational psychology. The ability to recognize one's own self, one's 'I', is an ability to affirm certainty about oneself and certainty about one's freedom. The self–consciousness is that of a free subject and that subject coincides with one's soul.

Kant in *Lecture 1* (L_1) of his *Vorlesungen über Metaphysik* lists four topics as the major themes of his Rational Psychology:

1. the soul is a substance;
2. the soul is simple;
3. the soul is a single substance;
4. the soul is a spontaneous agent.[84]

These are his "transcendental concepts" by which he interpreted the soul in general, but in effect it was an outline for his metaphysics. (1) To say that the soul is a substance is to imply that "I am a substance". It also means that I experience myself through my soul and its powers. (2) When I say "I think", I am expressing a simple representation that occurs in one subject. It is not a composite of several representations but of one, hence the soul, or the 'I' must be simple. (3) My consciousness is that of a single substance, not of a composite, hence I am conscious of myself as a single subject, one substance. (4) The soul is free, it is spontaneous subject. I am conscious of my actions and determinations and this being conscious of them renders me free.[85]

These theses are always read in comparison with 'The Paralogisms of Pure Reason', which is Chapter 1 of the Book II ('On the Dialectical Inferences of Pure Reason') of the

[83] Karl Ameriks further divides the pre–critical period into three: empiricist period until 1755, rationalist until 1763 and skeptical until 1768. Ameriks (2000: 3).

[84] Kant (1968: XXVIII/1, 265).

[85] Kant (1968: XXVIII/1, 269).

'Transcendental Dialectic' of the *Critique of Pure Reason* (A341–405, B399–432). Of the entire book, this part is the most altered in the second edition.[86]

The first three themes of the lectures appear as the three Paralogisms in the *Critique* as the arguments of the rational psychology now refuted by Kant. The last point on freedom is defended and persists throughout Kant's critical period.

A careful reading of both versions of the *Critique* shows no clear indication why Kant rejected the first three claims of his earlier rational psychology. What we get from the text, especially from the second edition of the *Critique*, is that through the use of empirical methods of investigation one cannot reach the conclusions rational psychology claims to make, not that its arguments are false. Too often we read Kant in the context of his closest predecessors (Locke and Hume in particular) for whom rejection of the self as a substance was part of their rejection of Cartesian metaphysics. Matters are more complicated with Kant: in neither version of the *Critique* do we actually read a clear refutation of the above points of the rational psychology from his lecture, nor an argument against them. Instead, Kant clearly insists that there is no demonstration of the spiritual and eternal status of the self.

The same is true for simplicity and substantiality. Kant rejects materialism in favor of the immaterial self, but immateriality does not imply simplicity and vice–versa.[87] Likewise, apperception does not prove one's existence, nor, contrary to Descartes, does thinking prove one's self.[88]

Kant's theory of mind must be read in the context of his two major doctrines in the *Critique*: transcendental deduction and transcendental idealism. Transcendental deduction implies that we can come to know certain a priori principles governing our experience and transcendental idealism says that these principles are valid only subjectively. The last point also implies subjective certainty of our knowledge of the world.

Kant's transcendental deduction is an attempt to prove the objective validity of our knowledge by indicating that the objects of our experience *are* the categories. But if our knowledge is essentially a composition of ideas and concepts (as it was for Locke and Hume), and if ideas and concepts do not indeed correspond to the objects (things-in-themselves) but to our intuition and to our prior concepts,[89] no objective certainty can be proven. For Kant's critics, the goal of his transcendental logic to "determine the origin, the range, and the objective validity of such rational cognitions"[90] is problematic in being successful.

In this context there simply is no possibility of treating the mind (or the soul) as a philosophical object of investigation. Even what *can* be said of mind cannot be proven by an *a priori* argumentation; even less so by an empirical investigation. Hence, this later Kantian skepticism about the philosophy of mind and the *generis sui* refutation of rational psychology.

---

[86] Buroker (2006: 213).

[87] Kant (1956: A 346/B 404).

[88] Kant (1956: A 343/B 401).

[89] Kant (1956: B xvi-xvii).

[90] Kant (1956: A 57/B 81).

McCann speaks of one condition for the philosophy of mind to be meaningful and objective, viz. to have affirmed causal relations that connect perceptions and consciousness to the external world in space and time. The self must be "an empirical self" but if the physical world itself relies on the 'I think', then the empirical self is an illusion. There is no base for the proof of the self outside of the self.[91]

Kant's insistence on the inexpressible of the physical in mentalistic language has similarities with Davidson's 'anomalous monism' and his stress on the necessity and at the same time inability of using physicalistic language for mental events.[92] Finally, Kant had great influence on the philosophies of mind of the two major authors for this work: C. S. Peirce and L. Wittgenstein. Peirce's revision of the Kantian categories brought him to the reconsideration of the notion of continuum and understanding of the mind neither as a substance nor as a property but as a semeiotic continuum. For Wittgenstein the ineffability question and the question of phenomenological vs. physical language would occupy most of his philosophical endeavors.

Conclusion

With Kant we have for the first time in the history of philosophy a paradoxical affirmation that the ontological and epistemological questions on the nature of the soul/mind and the world are a necessary consequence of our human nature, the product of human reason, but that such questions cannot be answered by that same reason; they become semantically ineffable. This paradox was already emerging in Aristotle who saw the solution in a foundationalist approach in his philosophy. In a different time when Aristotelianism was largely dismissed, Descartes and his successors proposed the same solution within a framework of a foundationalist epistemology. Kant's struggle and then criticism of a rational psychology was his criticism of philosophical foundationalism. Kant did not refute rational psychology on the grounds that its quest belongs to our nature to inquire about the mind. The quest is valid but cannot be carried out in any conclusive way because it would presuppose stepping out of the inquiring mind. This ineffability of mind will be clearly adopted by the adherents of the universality of language and ineffability of semantics in the 19th and 20th centuries.

This Kantian ineffability toward the mental is a stumbling block to many whose philosophy of mind appears more optimistic. New accounts of Aristotle, Hume and Kant demand a naturalized version of the mind and result in new ('naturalized') versions of these authors. Fairness to the history of philosophy shows significant inconsistencies in these readings.

This chapter showed the philosophical evolution from *psuchê* to the modern notion of the mind. Difference in terminology implies differences in the understanding of human nature and its relation to the world. These questions were asked because it was assumed that they logically follow from ontologies and theories of knowledge of the selected philosophers. Nonetheless, I was not able to provide a single satisfactory an-

[91] McCann (1995: 345–346).

[92] Davidson (1987b).

swer by considering different metaphysics of *psuchê*, mind and body, having as the starting point the metaphysics of these philosophers. Certain manipulations of the history of philosophy by which one places the philosophy of mind at the center of every philosopher's work probably would accomplish the task of having the answers. Since I do not think this method is warranted, it was not used in this chapter.

The above-unanswered questions were meant to be an exercise in the history of philosophy to test the validity of these questions and at the same time to demonstrate that metaphysics of the past fails to answer them. The mere fact that these questions were not answered does not make them invalid. Paradox does not mean contradiction; unanswered questions do not imply their falsehood. Because the contemporary philosophy of mind is situated within modern metaphysics and epistemology, these questions will be carried into the next chapter.

CHAPTER II

Mind and Body: The Many Senses of Identity

Introduction

In 1903 Charles S. Peirce gave his famous "Lectures on Pragmatism", where he made this remark on modern philosophy's treatment of the human mind:

> A subtle and almost ineradicable narrowness in the conception of Normative Science runs through almost all modern philosophy in making it relate exclusively to the human mind. The beautiful is conceived to be relative to human taste, right and wrong concern human conduct alone, logic deals with human reasoning. Now in the truest sense these sciences certainly are indeed sciences of mind. Only, modern philosophy has never been able quite to shake off the Cartesian idea of the mind, as something that "resides" – such is the term – in the pineal gland. Everybody laughs at this nowadays, and yet everybody continues to think of mind in this same general way, as something within this person or that, belonging to him and correlative to the real world. A whole course of lectures would be required to expose this error. I can only hint that if you reflect upon it, without being dominated by preconceived ideas, you will soon begin to perceive that it is a very narrow view of mind. I should think it must appear so to anybody who was sufficiently soaked in the *Critic of the Pure Reason*.[1]

Over one hundred years later, despite the developments of the 'philosophy of mind', 'cognitive science', and the great progress of the neurosciences, Peirce's remark is as relevant as it ever was. Philosophical language remains persistently Cartesian even though no philosopher today would claim to be a substance dualist.

The contemporary philosophy of mind begins by refuting Cartesian substance dualism and by searching for some kind of identity between the mind and the body in the context of the new scientific developments in neuroscience and medicine. As separation involves dualism, identity involves some kind of monism. Monism comes in two types: idealism and physicalism, but in the philosophy of mind of the twentieth–century idealism does not figure. Physicalism implies an identity between mind and brain; the

[1] Peirce (1934: 81).

question is, which identity is it? This chapter will examine most of the contemporary theories of the philosophy of mind and their precursors (emergentism and behaviorism) from the view of identity, in addition to the four previous questions of ontology, metaphysics, epistemology and semantics.

The notion of identity must not be considered as a radically separate topic from ontology and metaphysics. Depending on how identity is defined, the ontological question of existence will be answered as well. The notion of identity will also imply the metaphysical question of the nature of mind and body, since identity implies *some* nature. Finally, it will pave the way to answer the epistemological question by indication of what it is that can be known.

Paul Gilbert speaks of identity in terms of analogy and, in examining passages from Aristotle, individualizes five kinds of analogy.[2] I will follow his pattern but will turn the idea around and suggest regarding all kinds of analogy *as identity*, and adding tautology to the top of the list. The treatment of tautology partially will follow the historical study by B. Dreben and J. Floyd,[3] however the theme of tautology in relation to identity will be made more explicit in the next chapter, especially in relation to Frege, Russell and Wittgenstein.

The first kind of identity is linked to the law of identity, expressed as 'identical to itself'. Of all other kinds of identity presented here, without doubt, this is the strongest. The controversy begins when we ask if the law of identity expresses any knowledge, conveys any information. Those who answer negatively, consider the law of identity to be tautologous, thus expressing nothing. Those who answer positively, consider the law of identity to determine the limits of our language and to be essential in determining truth.

To the first group belonged Locke, who claimed that knowledge is characterized by four propositional relations between ideas: identity, relation, necessary connection and existence. Identity was especially important for formal logic, but in the context of knowledge as relation between ideas, it is the most basic, primitive and less expressive. Therefore, since all formal logic (based on the three laws of thought: the law of identity, excluded middle and non–contradiction) is occupied with the most primitive propositions and their relations, it is essentially pointless. Identity is nothing more than *just* a tautology such as 'Gold is gold' and 'Red is not blue,' being no information of any kind. It is no more than an intuitive knowledge of the simple discernment of ideas.[4]

Similar to Locke was the position of Kant, for whom analytical propositions are those "whose certainty rests on *identity* of concepts" and are of two kinds: 'explicit' and 'implicit'. 'Explicit' are tautologies: "Tautological propositions are *virtualiter* empty or *void of consequences*, for they are of no avail or use. Such is, for example, the tautological proposition, *Man is man*. For if I know nothing else of man than that he is man, I know nothing else of him at all".[5]

---

[2] Gilbert (1995: 75–107); Gilbert (1991: 263–266).

[3] Dreben and Floyd (1991).

[4] Locke (1975: I, 2 §§19–20; IV, 1 §4).

[5] „Tautologische Sätze sind *virtualiter* leer oder *folgeleer*; denn sie sind ohne Nutzen und Gebrauch. Dergleichen ist z. B. der tautologische Satz: *der Mensch ist Mensch.* Denn wenn ich vom

Unlike the explicit propositions, implicitly identical ones "clarify the predicate" ("*sie machen das Prädikat*") and "are not void of consequences" ("*sind dagegen nicht folge– oder fruchtleer*").[6] But unlike for Locke, for Kant propositions of formal logic and mathematics are not analytic but synthetic a priori, thus *not* tautologous.

Hegel elaborated the Lockean/Kantian idea of tautology being empty of information, speaking of the law of identity: "in its positive expression $A = A$ is, in the first instance, nothing more than the expression of an empty *tautology*. It has therefore been rightly remarked that this law of thought has *no content* and leads no further".[7]

Poincaré strongly builds on Locke's and Hegel's ideas of formal logic being uninformative and tautologous. However he is much closer to Kant in asserting that mathematics is based on mathematical induction, thus being informative.[8]

To the same tradition of tautology being uninformative belonged Russell and Wittgenstein. For the author of the *Tractatus*, all propositions of logic and mathematics are analytic and tautologous propositions and say nothing of truth. Truth depends only on the relation of the proposition itself and reality. Hence, the picture theory of meaning was meant to describe such relations. Burton Dreben and Juliet Floyd thus assess the earlier Wittgenstein's notion of tautology: "If truth can be discerned from the propositional sign alone, then no claim is made upon reality, there is nothing corresponding to the proposition that makes it true or false, and hence, intuitively, the proposition is uninformative, superfluous, empty, perhaps not even a genuine proposition: in short, *tautologous*".[9]

Indeed, TLP claims that logical propositions are lacking in sense, meaningless, they are *sinnlos* (see TLP 4.461, 5.132, 5.1362, 5.5351: i.e., tautology, logic, mathematics), while propositions of the TLP themselves are nonsensical, *unsinnig* (see TLP 3.24, 4.003, 4.124, 4.1274, 4.4611, 5.473, 5.5303, 5.5351, 5.5422, 5.5571, 6.45, 6.51, 6.54: i.e., metaphysics). In the next chapter I will elaborate on the view of identity in Russell and Wittgenstein.

Leibniz, Frege and C. I. Lewis gave the opposite, a positive view on the law of identity, different from tautology. Here their views are classified as the second type.

For Leibniz, the law of identity is the basis for defining truth and it is *not* tautologous:

> From the fact that A is A, or for example, that three–legged is three–legged, it is obvious that anything is as much as it is or is equal to itself. Hence (to show how useful identities are by an example) philosophers have long ago demonstrated that a part is less than the whole by assuming only this definition: that is less which is equal to a part of another (the greater).[10]

Menschen nichts weiter zu sagen weiß, als daß Mensch ist, so weiß ich gar weiter nichts von ihm". Kant (1923: I §37); Kant (1988: 117).

[6] Kant (1988: 118).

[7] „[Dieser Sätz] in seinem positive Ausdrucke A = A, ist zunächst nichts weiter, als der Ausdruck der leeren *Tautologie.* Es ist daher richtig bemerkt worden, daß dieses Denkgesetz *ohne Inhalt* sey und nicht weiter führe". Hegel (1978: I, 2, ii, A, 262); Hegel (1989: I, 2, 2, A, 411).

[8] Poincaré (1935: 1–2).

[9] Dreben and Floyd (1991: 33).

[10] "Ex eo quod A est A, seu quod tripedale verbi gratia est tripedale, manifestum est unumquodque tantum (nunc) esse quantum est, seu esse sibi ipsi aequale. Unde (ut exemplo usum

Unlike for Kant, for Leibniz all truth is analytic and, just as the law of identity, is never tautologous.

Frege's views on the nature of arithmetic in his *Grundlagen* are notoriously anti–Kantian and are largely based on Leibniz's ideas. Analytic propositions are those that are derived from the laws of logic, but they are not tautologies:

> Our next aim must be to show that the Number which belongs to the concept *F* is identical [*gleich*] with the Number which belongs to the concept *G* if the concept *F* is equal [*gleichzahlig*] to the concept *G*. This sounds, of course, like a tautology. But it is not; the meaning of the word "equal" is not to be inferred from its etymology, but taken to be as I defined it above.[11]

Thus, unlike for Kant, arithmetic just as formal logic was analytic. Logical truth was analytic but not empty of content, or tautologous.

Finally, C. I. Lewis' position is somewhat intermediate between the two types presented above. For him, 'tautologous' meant "exhaust of all possibilities" and "any logical principle (and, in fact, any other truth which can be certified by logic alone) is tautological in the sense that it is an analytic proposition".[12] But Lewis is in clear disagreement with the first group of philosophers, for he did not believe that logical (analytical) propositions were lacking in sense. Although not as significant as synthetic propositions, they are 'significant' in demarcating the limits of possibilities and necessities in our thoughts and language.

In relation to the contemporary philosophy of mind, the only set of theories that would use the law of identity in describing the relations between mind and brain is eliminative reductionism. In this context, the identity is tautologous and lacking in any sense. The only true reality is that of the physical. All predication of the mind in relation to the brain is superfluous.

Although eliminative reductionism comes in many forms and versions, I will limit my presentation in this introduction to that of Francis Crick's by now classical formulation of his 'Astonishing Hypothesis': "'You,' your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules. As Lewis Carroll's Alice might have phrased it: 'You're nothing but a pack of neurons.'"[13]

To be clear, what is senseless or tautologous here in the eyes of a reductive eliminativist, of course, are not 'the behavior of a vast assembly of nerve cells and their associated

---

ostendam identicarum) demonstratum est jam dudum a philosophis, partem esse Minorem Toto, posita hac definitione: Minus est quod parti alterius (majoris) aequale est". Leibniz (1890: 299–300); Leibniz (1976: 226).

[11] „Wir wollen nun zunächst zeigen, dass die Anzahl, welche dem Begriffe F zukommt, gleich der Anzahl ist, welche dem Begriffe G zukommt, wenn der Begriff F dem Begriffe G gleichzahlig ist. Dies klingt freilich wie eine Tautologie, ist es aber nicht, da die Bedeutung des Wortes „gleichzahlig" nicht aus der Zusammensetzung, sondern aus der eben gegebenen Erklärung hervorgeht". Frege (1950: §73).

[12] Lewis and Langford (1932: 211).

[13] Crick (1995: 3).

molecules,' but 'joys', 'sorrows', 'memories', etc. These later ones are, or meant to be, translatable into the 'language' of 'a vast assembly of nerve cells.' The characterization of the "no more than" makes this kind of identity, indeed, tautologous, the strongest but at the same time least informative, or not informative at all. Eliminative physicalists will try to demonstrate that the propositions of folk psychology about mind and consciousness are meaningless, just as tautological propositions are meaningless and uninformative.

A weaker kind of identity is expressed in univocal and equivocal predications. Aristotle's *Categories*[14] begin by defining both in terms of identity to the same name. For univocal predication, the name and the definition in relation to the name is the same, while for equivocal predication the name and the definition are different. Aristotle gives several examples for each predication. For instance, in univocal predication a man and an ox are both 'animal' in name as well as in definition. On the other hand, in equivocation, a real man and a drawn man can both be called an 'animal', but each is defined differently. Thus, the main difference between equivocal and univocal predications is in how the terms are defined or whether the terms have the same sense (univocal) or different (equivocal).

In both cases there is a need to stress where the identity statement comes in. Both univocal and equivocal predication implies the identity of terms, i.e. the 'animal', and the difference, again, being only in whether the sense of the word 'animal' is one or two. There is no identity of the reality of the two animals: a man or an ox, a real man or an image of a man. Hence, it is for this reason that I shall consider the type–identity theories ('pain is a C–fiber stimulation') and behaviorism ('mental states are manifested in external behavior') as univocal predication and functionalism ('mental states are physical states and are functions of the brain'); the original computationalism of A. Church and A. Turing ('mental states are computational states and are mechanical rule–following'); and cognitivist theories ('mental states are computational states of the brain') as implying equivocal predication between the mental and the physical. None of these theories, in fact, ever claimed having introduced two distinct realities, of which one would be 'the mind' and the other some realization of 'the physical' (the brain, computer's hardware, etc.). All the theories within these two classifications, as well as the remaining ones in this chapter, are monistic and thus predicate only one reality, albeit without elimination (as in tautology).

In univocal predication, a man and an ox are two distinct objects, but when we define each of them as 'animal', a univocal identity is established between them. A man and an ox can be considered as the two distinct manifestations of being an animal. Here, however, we should not take too literally Aristotle's examples of the two distinct objects for univocal predication. If not, univocal would apply only to substance dualism. Instead univocal can also apply to two distinct concepts sharing the same definition as well as the same name. For instance, Frege's distinction between number (*Anzahl*) and numeral (*Zahl*), or C. S. Peirce's types and tokens are examples that do not involve any distinct physical objects while do share similarity in name and definition.

---

[14] Aristotle (1975: 1a).

Some will consider univocal predication as a sortal one, in which case the identity of the defined objects is shared. As Steven T. Kuhn states: "Sortal predicate, roughly, a predicate whose application to an object says what kind of object it is and implies conditions for objects of that kind to be identical".[15] Therefore sortal predicates do imply relation of identity.

However, notice that while univocal predicates accommodate sortal predicate, equivocal do not. Furthermore, univocal identity is stronger than equivocal, since it shares names *and* definitions. One of the suggestions of this chapter, in fact, will be that the theories which are considered to imply univocal prediction implement a stronger notion of identity than those that imply equivocal predication, while certainly having a weaker notion of identity than those theories implying tautology.

Two further notions of identity are related to the concept of analogy. Here too we can refer to Aristotle's writings on analogy, found in the same text of the *Categories,* but also in *Sophistical Refutations*[16]*, Metaphysics*[17] and *Nicomachean Ethics*[18]. In these texts Aristotle defines analogy as the use with related senses. Unfortunately, as with many other topics in Aristotle, his notion of analogy was ambiguous and from Boethius to Cajetan had raised many interpretations and further elaborations. Without going into a detailed presentation of analogy, I will only indicate two senses that, in my view, are relevant to the notion of identity of mind and body.

First is analogy of proportions or relations. For instance, the word 'healthy' can be used *analogically* of food as well as of walking. Notice that here, just as in the cases of univocal and equivocal predications, the identity is that of the term concerned and not of the objects of analogy. The notion of analogy in language, logic, metaphysics or mathematics (in which the concept originated) is used as a methodological tool for understanding proportions and relations. The notion of analogy was also used in the Middle Ages in determining the same predications of God and creatures.

In the philosophy of mind such a tool can be also applied in understanding the relations of mind and body. Being a weaker notion of identity, analogy of proportions was used in British Emergentism ('mind is an emergent property of the brain') and token–theories of mind ('mental states are tokens of the brain states,' or types), such as supervenience and 'anomalous monism' (mind as 'a supervenient property of the brain').

In the above example taken from Aristotle, the word 'healthy' is referred to two distinct subjects: food and walking. In the chosen theories of the philosophy of mind, particularly those that are prone to property dualism, mental states are (emergent or supervenient) properties of the physical states. The relation of the identity is convened in the notion of 'property' while both physical and mental states stand in different ways in such relation to each other. To the same category of analogy of proportion belong T. Kuhn's 'paradigms', and R. Harré's 'scientific models' used in the philosophy of science.

Analogy of proportion is a relation of the three elements, while analogy of proportionality, close to the notion of metaphor, is a relation of the four elements. Aristotle's

---

[15] Kuhn (1999: 865).

[16] Aristotle (1955: 165b25–167a10).

[17] Aristotle (1933: 1003a33–35).

[18] Aristotle (1934: 1131a).

example is the following: 'The cup is to Dionysus as the shield to Ares,' and vice–versa 'The shield is to Ares as the cup to Dionysus.' Notice that in the former analogy of proportion such reversal is not possible, for 'brain being property of the mind' is wrong.

In this analogy of proportionality, just as in the previous kinds of identity, the identity is in the relations of the components. In this four–element relation, the identity stands of the first pair ('the cup to Dionysus') to the second ('the shield to Ares'). To see the contrast between the identity by tautology, let us recall that the connective between the two relata was 'no more than,' while here the connective is 'is.' Frege would insist on the further differentiation of which 'is' is used in these examples: the 'is' of identity ('Phosphorus is Hesperus'), the 'is' of predication ('Aristotle is a philosopher'), the 'is' of existence ('God is'; 'There is at least one letter on this page'), or the 'is' of class ('A horse is a four–legged animal').[19] However, as many commentators noted, the ambiguity lies not with the verb 'is' but within that to which the verb is referred.[20] Likewise identity is not stressed within any connective between the relata but in a special relationship of the relata stressed by the statements of analogy. To the analogy of proportionality belongs the strong AI Thesis ('Mind to the brain as software to the hardware').

The final and weakest notion of identity is that of a metaphor.[21] In Aristotle two passages are of the most importance, from *Poetics*[22] and from *Rhetoric*[23]. Similar to the analogy of proportionality, metaphor also uses four terms, two terms related to the other two. The difference between analogy of proportionality and metaphor consists in that in the former there is some established relationship not only within each compared pair but also between them. Thus, in the analogy of proportionality the relationship is 1 to 2 as 3 to 4, while in metaphor the relationship can also occur in the way that 2 to 4 as 1 to 3. Hence, both are correct: 'Old age to life as the evening to day,' and 'Life to day as old age to the evening.' This second relation cannot be established in the case of analogy of proportionality, since 'Dionysus to Ares as the cup to the shield' makes little sense. Since the relation of identity is weaker in metaphors than in any analogy, metaphors have more linguistic plasticity and are open to more improvisations, indeed, 'transfers' from one pair to another.

In the context of the philosophy of mind I shall discuss only one theory that uses such weak identity, J. Searle's 'biological naturalism': 'Minds to the lower–level features of the brain, as the higher–level features of a system to the lower–level elements of that system.' Notice that here the first element can be easily and meaningfully related to the third, and the second to the fourth. However, Searle's 'biological naturalism' sometimes can also be considered as an analogy of proportionality, given his analogy of liquidity: 'mind to brain as liquidity to water,' in which case the (metaphorical) reversal of the 1 to 3 and 2 to 4 is impossible. Yet, given the fact that his 'biological naturalism' is expressed

[19] Frege's so called ambiguity thesis can be found in Frege (1967: 167–178).

[20] For historical introduction and systematic treatment of this topic, see Knuuttila and Hintikka (1986).

[21] "*Μεταφορά*, f. *μεταφέρειν* to transfer, f. *μετα–* meta–1 + *φέρειν* (root *φερ–* : *φορ–*) to bear, carry", Oxford English Dictionary (2009).

[22] Aristotle (1995: 1457b).

[23] Aristotle (1926: 1406b–1412a).

more precisely in metaphorical rather than in analogical form, it will be treated here as a metaphorical identity.

For many the notion that identity is expressed by analogy and metaphor might seem strange if not misleading. We must keep in mind, however, that the notion of analogy itself comes from mathematics (from Greek *ἀναλογία* equality of ratios, proportion).[24] Given three terms in a proportion, it is possible to determine the fourth term with the following: '$x = \frac{bc}{a}$, in the case that $a : b = c : x$.'[25] Again, unlike in metaphor, the second and the fourth, the first and the third terms cannot be exchanged, and metaphors, being applied mostly in rhetoric and poetics, cannot be written in mathematical symbolism.[26]

This, however, does not imply that in metaphors, although less than in analogies, there is no statement of identity. I. A. Richards[27] and M. C. Beardsley[28] had spoken of 'an identity statement (*X* is *Y*),' of 'a predication or membership statement (*X* is a *G*),' or of 'a statement of inclusion (*F*s are *G*s)' as indispensable parts of any metaphor. Indeed, if the pairs of propositions in metaphors have no identity of any kind, it would be no metaphor at all.[29]

In a final note it should be mentioned again that all of the above–mentioned theories of mind and body are monistic. Thus, they all make some sort of identity between the mind and the body. My purpose in this introduction, as it shall be for the rest of this chapter, is to point out what kind of identity it is. The major theories of mind in this chapter will be treated exclusively from the perspective of identity of mind and body, in addition to the previous four questions defined in the preceding chapter. However, my first thesis of this work shall be the conclusion for this chapter and for the first part, that although the theories treated here are monistic, those that rely on equivocal, analogical and metaphorical predications had failed to state the mind/body identity and are, essentially, dualistic theories. Those theories that are implying the law of identity and univocal predication, on the other hand, either contribute nothing in elucidating the relationship between the mind and the body, or their statements of identity result in contradiction. Finally, what unites all of these theories is the paradox of inconsistency between monistic ontology and dualistic language. The solution of this paradox more often resulted in either elimination of the mind or in recognition of inefficiency of language.

I conclude this introduction with, perhaps, if not the clearest definition of identity (such to my knowledge has not been provided yet), at least the most widely spread in the English language from the 4th edition of *The Oxford English Dictionary*. See, however,

---

[24] "Analogy, Inference of the Truth of an unknown result obtained by noting its similarity to a result already known to be True. In the hands of a skilled mathematician, analogy can be a very powerful tool for suggesting new and extending old results. However, subtleties can render results obtained by analogy incorrect, so rigorous Proof is still needed". Weisstein (1999: 42).

[25] The above example is from Guzzo, Mathieu and Lia (2006: 402). For a proof and analysis, see Lardner (1828: 255), Crelle (1834), Cortazar (1847).

[26] On the notions of proportion and proportionality in ancient Greek mathematics and philosophy see Mueller (2006).

[27] Richards (1936).

[28] Beardsley (1962).

[29] For historical and systematic overview on metaphor, see Hills (2012).

the note for the etymology of the term from the same source. There identity is defined as "The quality or condition of being the same in substance, composition, nature, properties, or in particular qualities under consideration; absolute or essential sameness; oneness".[30]

1. Behaviorism and Identity Theories

1.1 Metaphysical, Methodological and Logical Behaviorism

Behaviorism came as a series of responses to modern philosophy's dualism of mind and body, its insurmountable approach to human nature, and to psychology and physics. It was a 'series' because there were three kinds of behaviorism: metaphysical or psychological, methodological, and logical. All three were developed more or less at the same time, in the first half of the twentieth century. Each had different authors but all three can be found in the works of Francis Skinner, who was considered to be its major representative. Despite significant differences among diverse types of behaviorism, it can be generalized as the view according to which any difference between two mental states can be fully accounted by a clear difference in the behavior associated with each state.[31]

The earliest type of behaviorism was developed by Pavlov and Thorndike; it goes by the name of *metaphysical* or *psychological* behaviorism. For this kind of behaviorism, behavior is explained by reference only to the external stimuli as the sources in the environment. In its explanation reduced to external stimuli, this behaviorism eliminates any reference to mental or internal events.

For example, one can train a dog to respond to the sound of a bell prior to feeding. A dog is trained by the repetition of this mechanism to respond to the sound of the bell, but this response is purely physiological (i.e., salivation). Thus the theory must be also explained in terms of external stimuli (i.e., the sound of a bell prior to feeding).

Metaphysical behaviorism was inspired by the associationism of Locke and Hume. Associationism says that behavior is learned through associations and learning techniques that are considered to be methods of association. It is based on the deeper epistemic premise that perceptions (stimuli) should be associated with thoughts and ideas. The result of this association is knowledge of actions and learned behavior. These built associations were supposed to explain causal ('necessary' for Hume) relations in the

---

[30] "Various suggestions have been offered as to the formation. Need was evidently felt of a noun of condition or quality from idem to express the notion of 'sameness', side by side with those of 'likeness' and 'oneness' expressed by *similitās* and *ūnitās*: hence the form of the suffix. But idem had no combining stem. Some have thought that ident(i)– was taken from the L. adv. *identidem* 'over and over again, repeatedly', connexion with which appears to be suggested by Du Cange's explanation of *identitās* as *'quævis actio repetita'*. Meyer–Lübke suggests that in the formation there was present some association between *idem* and *id ens* 'that being', whence *identitās* like *entitās*. But assimilation to *entitās* may have been merely to avoid the solecism of **idemitās* or **idemtās*. However originated, ident(i)– became the combining stem of *idem*, and the series *ūnitās*, *ūnicus*, *ūnificus*, *ūnificāre*, was paralleled by *identitās*, *identicus*, *identificus*, *identificāre*: see identic, identific, identify above".

[31] Graham (2011).

world. Thus, 'to behave' in a certain way is a demonstration of one's knowledge about how one event is connected with another and how one perceptual experience is associated with one's ideas and thoughts. An important difference between British Empiricists' associationism and metaphysical behaviorism is that the behaviorists eliminated any reference to mental events in explaining human behavior; what remained were stimuli that were supposed to explain everything.

The second type of behaviorism made claims about psychology being the natural science of behavior and not of the mind. Therefore it goes by the name 'methodological'. Its major proponent was Watson, in whose writings, unlike in metaphysical behaviorism, one does not find eliminativist suggestions toward mental states. Instead, he claimed that mental states are private and as such they cannot be the object of an empirical science.

Methodological behaviorism was strongly influenced by logical positivism whose major influence was to insure scientific foundations for psychology. Watson expressed this effect by stating psychology's method of prediction of and control over an animal's or a human's behavior: "To predict, given the stimulus, what reaction will take place; or, given the reaction, state what the situation or stimulus is that has caused the reaction".[32]

The third kind of behaviorism, logical behaviorism, whose major proponent was Gilbert Ryle,[33] had even stronger roots in logical positivism, sharing its major premise of verificationism proposed by Schlick and Waismann, members of the Vienna Circle. The method of verification stated that only through verification in experience one can state the truth, meaning and justification of one's knowledge. It had particular strength in logic and the philosophy of language in the service of experimental science either to produce verifiable–by–experience propositions (synthetic), or to test propositions which are true by their own definition (analytical). The above definition of verificationism falls neither within synthetic nor within analytical propositions; a fact that contributed to its modification by Carnap[34] but later was abandoned especially under the criticisms of its major opponent, Karl Popper in his 1934 *Logik der Forschung*.[35]

Logical behaviorism was an attempt to respond to Peirce's challenge of dualism, quoted at the beginning of this chapter. If substance dualism considers mental and physical events separately, logical behaviorism interprets them as one: my belief that it rains and my walking under an umbrella in the rain are not two experiences but one. The mental state of the belief that it rains is inseparable from my behavior as a response to the belief that I have.

---

[32] Watson (1930: 11).

[33] Sometimes Wittgenstein is considered as a logical behaviorist but that it is a false assumption has been concluded by many scholars. His letters alone dismiss such interpretation in Wittgenstein (1995: 294). Wittgenstein's association with logical behaviorism is linked to his acquaintance with Ryle, of whom he allegedly said that that he is one of only two philosophers who understood his work. The name of the second philosopher is not known. See Monk (1990: 436). Against the perception of Wittgenstein as behaviorist or dualist, see Overgaard (2004: 263–286).

[34] Carnap (1936); Carnap (1937).

[35] Popper (1976).

As it was noticed earlier, all three types of behaviorism were prominent in the works of Skinner. Unless the language of mental events is eliminated, it is translated into the language of behavior and is interpreted as such, as a behavior of the subject.[36]

Graham underlines three reasons why behaviorism was rejected by the subsequent development in psychology, philosophy of mind and cognitive science. First, particularly in the case of Skinner, is the behaviorists' assumption that the explanation of behavior must be done in reference to the external–to–the–organism stimuli. This approach gives very limited space for cognitive neuroscience to study the information–processing mechanisms in the brain. It gives little, if any, credit to neuroscience and its methods.[37]

The second reason is the behaviorists' insistence on the too close relationship between the subjects' behaviors and their mental and/or physical states. Graham's 'zombie' analogy sums up this objection in the argument for qualia (qualitative private mental states), which are experienced not just in the behavior but foremost in sensing these qualitative states (i.e., 'pleasureness', 'painfulness', etc.). Thus, a zombie might act as if s/he is having a painful or pleasant experience but actually is not having anything of the sort at all. A study of her/his behavior tells us nothing about her/his mental state.

The third objection was given by N. Chomsky. Behaviorist theories are unable to explain the language learning process, especially children's rapid acquisition of grammar. Associative learning methods do not explain children's ability to understand the meaning of words by simply repeating them.[38] Instead, Chomsky argued that linguistic behavior is explained by the innate ability for language which can and must be studied by neuroscientific methods. This last objection was also against the main idea of behaviorism that language must be studied by observing the linguistic behavior and not, in Chomskean terms, studying language as an organ.[39]

All three versions of behaviorism were the first attempts to provide an account of physicalism in contemporary philosophy of mind: reductive eliminative physicalism in the case of metaphysical behaviorism and reductive non-eliminative physicalism in the cases of methodological and logical forms of behaviorism. Due to their incomplete ontology, their epistemic (the possibility of studying and knowing human behavior) and their semantic (the capacity of explaining the link between language, the world and knowledge) premises were incomplete and inconsistent. By 'ontology' here I mean their understanding of what is actually studied in the subject and by 'incomplete' I mean their underestimation of the brain as an indispensable source for the explanation of the organisms' behavior. The phrase 'incomplete ontology' also implies their eliminativist account of mental events in explaining human behavior: behavior is as such not because of the stimuli but because of the mental states that are caused by the stimuli. If behaviorism would be applied to animal cognition only (i.e., I. Pavlov's version of psychological behaviorism), it might be true, at least to some extent.

The history of behaviorism became a hard–learned lesson of excluding neuroscience in the study of the mind. It also showed the importance of language in the study of

---

[36] Skinner (1974: 18).

[37] Stich (1984: 647–649).

[38] Chomsky (1959).

[39] Skinner (1977).

human behavior. What remained of behaviorism was its eliminativist thesis, but elimination presupposes reduction into something. Behaviorists reduced mental states to behavior as a response to stimuli; other reductionist theories took the opposite route and proposed reduction to the physical, whereby the physical meant the brain, even on the whole type–level.

What was considered to be behaviorism's greatest accomplishment, viz. bringing together ontological, epistemological and semantic questions together, turned out to be its major difficulty and the cause of its ultimate decline. What this bringing together means is the consideration of one's behavior as the manifestation of one's state of the mind to the point of dismissing the importance of empirical study by neuroscience, according to some behaviorists. Meaning and knowledge of one's behavior was supposed to mean the same as one's intending and willing. In this context, the metaphysical question of the nature of the mind would seem to be redundant.

However, it was not dispensing with the metaphysics of the mind that brought behaviorism to irrelevance in the present philosophy of mind. It was rather taking for granted the identity of one's actions and one's mental states. Ironically, it was the logic, the language itself, that did not add between the statements and their actualizations in reality, not an overestimation of the logical and linguistic importance. It was precisely the question of language turning into the question of identity that was the central issue for the next, here presented, theory of mind.

1.2. Type–Identity Theories of Mind/Body

The identity theory of mind/body holds that mental events, states, processes are identical to the events, states, processes of the brain. In the philosophy of mind it comes in two versions: type theories and token theories. A type theory identifies types of mental events with types of brain states. They are not just correlated with brain states, they *are* these states. The most celebrated example of type–identity theories is that pain is a C–fiber stimulation.

Such identification might, but does not have to, by definition eliminate types of mental states. Mental states are the manifestation, in some sense, of the brain states and in this way both types are identical but not exclusive one of another.

A token theory says that every token of a mental event is identical with tokens of brain states. Type identity theories imply that under all types also tokens are identical, while token identity theories deny the identity between the types. Each theory has several variations and within each theory notions of 'type', 'token' and what is considered to be an 'identity' are disputed. In this part I will focus only on type–identity theories developed mostly by Smart and Place, while token–identity theories will be treated in connection with supervenience and anomalous monism.

Type–identity theories developed within and as a consequence of logical behaviorism and thus it make sense to treat them in this work as a theoretical consequence. The type–identity theory was first proposed in 1956 by U. T. Place, in his article "Is Consciousness a Brain Process?" and further developed by H. Feigl in his 1958 paper "The 'Mental' and the 'Physical'". J. J. C. Smart acknowledges that these papers came after discussions with

him and C. B. Martin at the University of Adelaide, and at that time were elaborations of G. Ryle's account of the first person behaviors as 'avowals'.[40] For Ryle, to have a certain mental state was to have a behavior corresponding to it, i.e., to have a toothache is to rub one's cheek with a hand, etc. Smart's suggestion was that these 'avowals' are to be explained one day in purely physicalistic terms by neuroscience, but he also realized that this thesis in fact goes against behaviorism itself.

Smart accepted Place's explanation of consciousness as a process in the brain to be as consistent as identifying lightning with "a motion of electric charges".[41] But one must keep in mind that 'sensation', which can be part of conscious experience, is the same as a brain process. Smart introduced the Fregean distinction between sense and reference where 'sensation' and 'brain process' are different senses of the same reference, just as 'Morning Star' and 'Evening Star' are different *Sinnen* of the same (one) *Bedeutung*, viz. planet 'Venus'.

Place usually referred to the relationship between mind and brain in terms of a 'constitution': lightning is *constituted* by an electrical discharge but it does not mean the same as 'motion of electric charge'. Smart, on the other hand, spoke of an identity: a professor of anatomy is identical with the dean of the medical school. But the property of being a professor of anatomy is not identical with the property of being a dean of the medical school.[42]

To avoid the objection of confusing sensations with the states, Smart introduced the notion of properties being 'topic neutral'. Just as his notion of identity was that of logical, or mathematical identity ('a = b'), 'topic neutrality' was the same as that of logical constants: 'and', 'not', 'or', 'if', 'if, then', and 'iff'. These constants bear no meaning on their own and without notions on their right and left for which they act as connectors, their use by themselves generates no meaning. In this sense, to say that some sensation is caused by some perceptual experience is insufficient for stating that the sensation's properties are physical or mental (non–physical).

Place introduced a further important distinction in our usual accounting of the perceptual experience. Whenever I report my perceptual experience, I do not describe the way the things are but the way they appear to me, the way my sensation tells me how the things are.[43]

These descriptions, at best, can be identified with the way the brain works but not the way the world is. It is in this sense that our sensations are 'topic neutral': they can be both or either mental or physical, just as it is in cases of logic or arithmetic.[44]

An important critique of and contribution to the type–identity theory was given by David Lewis when he insisted that an essential feature of any experience is its causal role: causality understood in terms of cause–effect dynamics. He linked this causal role of experience to physical states and physical states were interpreted as experiences.[45]

---

[40] Smart (2011).

[41] Place (1954: 255).

[42] Smart (2011).

[43] Place (1956: 49–50).

[44] Smart (1959).

[45] Lewis (1980).

Another contribution to the type–theory of mind and brain was done by Armstrong with his idea of an identity theory as 'Central State Materialism'. He identified beliefs and desires as brain states and sensations as topic neutral expressions in the public language. In this way we have no privileged access to our mental states.[46]

These contributions made two very important clarifications within the type–identity theory. First, on the ontological level it brought the identified mind with the brain on the same causal level with the rest of the world. As a result, mind can and should be studied by physics (here physics is understood in a very broad sense of 'physical sciences') just as is the rest of the physical world. Second, it alienated language from the studies of mind, even if topic neutrality does play an important role of an analogy between logic, linguistic and mind.

Concerning the second point Smart notices that the way in which the brain represents the world might not be like a language. The representation might be like a map. A map relates every feature on it to every other feature. Nevertheless maps contain a finite amount of information. They have not infinitely many parts. We can think of beliefs as expressing the different bits of information that can be extracted from the map. Considered in this way beliefs would correspond nearly enough to the individualist beliefs characteristic of folk psychology.[47]

Language is considered as an abstraction from reality just as map is an abstract and incomplete description of topological space. That there could not be any perfect representation between the world and language is obvious, and if the analogy of the map would hold for language it would make sense. But in the above explanation it is not language but the brain that is compared to a map, which raises serious doubts about the realist–idealist nature and undermines the entire identity notion of a type–identity theory. What the brain (or a subject in effect) perceives is not the world itself but sense–data or some other kind of perceptual content whose relationship to both, brain and reality as it is, is yet to be specified.

Functionalism came from this theoretical context and it made many improvements for the type–identity theory. The major improvement consisted in adopting the multiple realizability thesis: the same mental states can be realized in many ways of which the physical states of the brain is just one of them. But functionlists deny type–identity as unsustainable and scarcely defendable. The next section will examine these points more.

Type–identity theories received many criticisms of which the following two must be mentioned here. First, and the most common objection, is type–theory's treatment of consciousness. Second is the question of identity in general.

David Chalmers was long time critic of the type–identity theory in its treatment of consciousness as a brain process. He objects that neuronal states are not capable of providing an explanation for the subjective qualitative states which constitute much of what we call consciousness.[48]

Smart's reply is to equate consciousness with awareness and insist that this awareness is none other than proprioception of the brain by the brain. That might be very well

---

[46] Armstrong (1968).

[47] Smart (2011).

[48] Chalmers (1996: 146–148).

studied by neuroscience. It is common mistake to consider consciousness as awareness especially in motor action. B. Libet's famous experiment was meant to show that consciousness (here too considered as awareness) is not required in performing even the most basic motor actions. What the experiment shows, instead, is that *awareness* plays no role in performing motor actions because it requires a considerably longer time to become aware even of one's own actions than to perform them. Consciousness comes as something extra to the awareness and it requires even more time than awareness and more 'work' by the neuronal connections for it to be acquired. Consciousness requires decision and, in the case of Libet's experiment, it requires the decision to take part in that experiment.

Even if Smart is correct in attributing awareness to proprioception of the brain by the brain, it still solves nothing of the problem of qualia which must be the content of consciousness, not awareness. Awareness is closer to perception in general than consciousness but awareness does not include consciousness. Awareness also requires no language; consciousness is expressed in language whenever language is used.

The second objection was put forward by S. Kripke but it was an objection for any identity theory, type and token alike.[49] The objection in relation to the type–identity theory stands on the fact that, by Smart's own acknowledgement, the type–identity is not necessary, it is contingent. For Kripke, only the necessarily true identity is true at all. If the term 'pain' is identical with some brain states corresponding to it, determining it, then the term is true necessarily and it cannot be imagined otherwise. Contingent *relationship* implies no *identity* at all.

But the same criticism applies to the token–identity theories as well (i.e., Davidson's 'anomalous monism'). If particular mental events are identical with some brain states then the identity, if it is an identity, must apply rigidly, viz., in all possible worlds, in all circumstances. But they cannot be identical necessarily, i.e., because they are multiply realizable or because there are no psychophysical laws which determine the nomic nature of their identity, etc. Hence, they cannot be identical at all.

What would count for Kripke as a necessary identity? From thermodynamics we know that heat is identical with its molecular kinetic energy in all circumstances, and there is no single situation in which this identity does not hold, so it cannot be contingent.[50] We say that we feel heat, have a sensation of heat, but the relationship here is contingent between our sensation and the molecular kinetic energy but not the heat itself.

The same analogy cannot hold with the sensation of pain because we cannot feel pain any differently from pain. If pain is a stimulation of a C–fiber then it cannot be a C–fiber stimulation without being felt as pain. If the identity theorists would say that a physical state (a stimulation of a C–fiber) produces the mental state (pain) then one might speak of some emergent property of pain on the physical state of a C–fiber stimulation. But the theory insists on an identity and a necessary co–occurrence.[51]

---

[49] Kripke (1980).

[50] But see Smart's objection to this example in Smart (2011): "Actually the proposition is not quite true, for what about radiant heat? What about heat as defined in classical thermodynamics which is 'topic neutral' compared with statistical thermodynamics?"

[51] Kripke (1980: 150–151).

Lewis objected to this strong understanding of identity as necessity and suggested that his notion of causality makes it possible to consider the identity between physical and mental states as contingent.[52] As to the identity between pain and our sensation of it, Rosental insists that this identity can be contingent, i.e., when we are not aware of our pain due to some distraction or some other co–occurring sensation.[53]

The main objection against type–identity theories remains qualitative experiences and the consciousness connected with them. The type–identity proponents' response, equating consciousness with awareness and awareness with the brain's proprioception, does not seem to be plausible for the reasons given above. The weakest point of the type–identity theory turned out to be that which was proposed as its main thesis by their authors: the ontological identity between mind and brain as an identity of types.

The type–identity theory stems from logical behaviorism and inherits some of its problems, such as the identity between mental states of a phenomenal character and linguistic descriptions in terms of 'topic neutrality'. It weakened theory's explanatory power leaving little possibility for an epistemic explanation of the identity. Causal relations introduced by Lewis and Armstrong provided some improvement to the theory but were not able, it seems, to enhance its epistemic strength.

The type–identity theory started with clarifying the behaviorists' problem of language and logic regarding the notion of identity. Preserving the physicalist commitments of behaviorism, the type–identity theory relied on univocal predication using reduction of the mental states to the brain states without the elimination of the former. The question of identity expressed in terms of the correspondence of types, however, failed to clarify the dynamics of the interaction between mind and brain. The type–identity theorists, in a way, had chosen the opposite way from modern philosophers: instead of proceeding from the ontological and metaphysical questions to the questions of knowledge and meaning, the type–identity theorists started with the question of identity and meaning with results for knowledge and existence. The question of metaphysics, of course, did not matter, for the nature of the mental experience was supposed to be 'taken care of' by the mere affirmation of which the physical states correspond to the particular mental states. This, in turn, resulted in the problems of qualia and charges with epiphenomenalism (lack of top–down causation) to which the type–identity theory did not respond adequately.

The most obvious failure of the type–identity theory, however, seems to be concerning the epistemological question strictly related to the univocal identity postulated in terms of type correspondence: how can we know that a particular type of the brain state corresponds to the particular type of the mental state if the brain state is 'measured', so to speak, in neural activity, while mental states are defined in exclusively linguistic terms? This mismatch between what can be defined empirically and what can be only described in language is obvious. Finally, how can this categorical mismatch be meaningfully expressed, if differences between the one and the other seem to be of the grammar and not a purely semantic issue? One way of solving this problem would be to look closer at the

[52] Lewis (1980).

[53] Rosental (1995: 354).

grammatical similarities and differences between physical and mental states. Computational theories of mind suggested precisely this venue.

2. Computationalism and Functionalism

2.1 Computational Theories of Mind

The term 'Computational Theory of Mind' was first proposed by Hilary Putnam in 1960 with significant developments by Jerry Fodor, but computational models of mind were developed much earlier and were the consequence of a new rethinking of mathematics in its theoretical and philosophical foundations. The idea of computation was particularly the result of the crisis in mathematics brought by the development of the new, non–Euclidian geometries, which had undermined intuition as the foundation of geometrical knowledge, consequently the foundation of mathematics as well.

Kantian intuition and Millian psychologisms were challenged by the objectivism and realism of Frege, Peano and Hilbert, often named as a 'formalist program'. The main premise of the formalist program was to insure that all mathematical reasoning is based on axioms to which intuitions were to be either reduced or eliminated.

Frege's 'logicist' project of settling the rules of language on the secure foundation of Cantor's set–theory based logic was part of that formalist program but it was doomed by its own inconsistency or, rather, inability to accommodate semantics within exclusively synthetic logico–mathematical constraints. Russell's famous paradox in 1902 and his theory of types as its solution had shown a need for greater complexity in dealing with semantics. Likewise, in 1931–2 Gödel's incompleteness theorems had demonstrated further constraints on what could be formalized.

A consequence of the formalist program was the mathematical quest for what class of functions could be considered computable or decidable by an algorithm. An important distinction between computable and non–computable algorithms was proposed by Alan Turing in 1936: the notion of a computing machine. Algorithm was defined as a function of computing by such a machine. For the first time an analogy between human computation and machine computation was proposed and then extended to the general analogy between mind and machine/computer.

Much of cognitive science's idea is based upon this initial analogy further transformed into reduction of the mind into computational functions of a machine. The two most important developments in cognitive science at that time were Chomsky's substitution of the behaviorist conception of language–learning with generative grammar,[54] later translated into the language of thought (the 'Mentalese') by Fodor,[55] and Marr's theory of vision that saw the human mind as an algorithmic symbol processor[56]. Since then none of the original ideas of the computational theory of mind remained unchanged.

[54] Chomsky (1959).
[55] Fodor (1975).
[56] Marr (1983).

Chomsky's project evolved into the minimalist project, while Marr's theory of vision has been challenged and today is largely overcome by developments in neuroscience.[57]

In the last sixty years cognitive science has developed significantly. Yet, the major problems within the computational theories of mind remained unchanged. In this presentation I will focus on what I believe are the absolutely essential problems of the computational models of mind: the problem of reasoning and the problem of continuity. It would not be an overgeneralization to say that the rest of the computational philosophy of mind and cognitive science is, in one way or another, an interpretation of these two original topics, which, to this day, remain largely unresolved.

2.1.1 *Thinking as Calculation*

It has been argued that Turing was the first to advocate the development of computers as a way of ontogeny for Artificial Intelligence (AI). It has also been argued that Wittgenstein was the first one to repudiate this theory, using arguments from his philosophy of language and epistemology.[58] We know as a fact that both Turing and Wittgenstein were Fellows of Cambridge and in 1939 were teaching two distinct courses under the same title: 'Foundations of Mathematics.'[59]

However, in the notes collected from Wittgenstein's *Lectures on the Foundations of Mathematics, Cambridge 1939*, there is literally not one remark on AI or Turing's article on the issue from 1936. The first explicit remark of Wittgenstein was made in his *Nachlass*, but a more important reference can be found in the *Remarks on the Philosophy of Psychology* (from 30.07.1947):

> Turing 'Machines'. These machines are indeed *people* who calculate. And what he said can be expressed in the form of *games.* And this would be an interesting game in which one is brought through certain rules to nonsensical instructions. I think of games like a 'racing game'. Someone gives a command "Go on in the same way", when this makes no sense, i.e., because the person can only run in the circle. Order has sense only in certain positions.[60]

This passage states a difference between what people normally do when they consciously follow the rules of some game (i.e., the 'racing game'), and when they follow the rules mechanically, without reflection. It is the second kind of rule–following to

---

[57] Hess (2004); Searle (1992: 197–226).

[58] Proudfoot (2004: 359).

[59] Monk (1990: 417).

[60] „Turing 'Maschinen'. Diese Maschinen sind ja die *Menschen,* welche kalkulieren. Und man könnte, was er sagt, auch in Form von *Spielen* ausdrücken. Und zwar wären die interessanten Spiele solche, bei denen man gewissen Regeln gemäß zu unsinnigen Anweisungen gelangt. Ich denke an Spiele ähnlich dem "Wettrennspiel". Man erhielte etwa den Befehl "Setze auf die gleiche Art fort", wenn dies keinen Sinn ergibt, etwa, weil man in einen Zirkel gerät; denn jener Befehl hat eben nur an gewissen Stellen Sinn". Wittgenstein (1980: §1096). See also Wittgenstein (2000: MS 135: 117–118, MS 229: 448).

which Turing Machines are compared. Here Wittgenstein makes an important assertion that Turing Machine can be only partially applied to our thinking.

A passage from *The Blue Book* demonstrates this point further:

> The problem here arises which could be expressed by the question: "Is it possible for a machine to think?" (whether the action of this machine can be described and predicated by the laws of physics or, probably, only by laws of a different kind applying to the behavior of organisms). And the trouble which is expressed in this question is not really that we don't yet know a machine which could do the job. The question is not analogous to that which someone might have asked a hundred years ago: "Can a machine liquefy a gas?" The trouble is rather that the sentence, "A machine thinks (perceives, wishes)": seems somehow nonsensical. It is as though we had asked "Has the number 3 a colour?"[61]

Unlike the empirical question whether a machine can liquefy gas, the question whether a machine can think cannot be answered: it is logically absurd, as is the question whether 3 has a color. They are meaningless because they violate rules of logical reasoning. However, if one cannot give an answer to the question 'Has the number 3 a color?,' Turing has argued that the same criteria of reasoning we apply to humans can as well be applied to machines.

The notion of mechanism was already at the center of interest of many mathematicians in Cambridge at least ten years before Turing began serious work on it.[62] In Wittgenstein's writing from that time there are some hunches of this theme. Thus, in 1933 Wittgenstein writes:

> When you think of an idea as something specifically human, organic, one might ask: "could there be a thought–prosthesis, an inorganic substitute for thought?" But now, if thinking is in writing and speaking, why can't it be done by a machine? "Yes, but a machine knows nothing" – of course, talk of a prosthesis of seeing and hearing makes no sense. Although we speak of an artificial foot, we don't speak of an artificial foot–pain.
>
> "But can a machine think?" – Could it have pain? It depends on what you mean by "something *has* pain".[63]

The nature of thinking seems to be at stake here for Wittgenstein. Alan Turing's paper from 1936 "On Computable Numbers" appeared to Wittgenstein as an attempt to integrate thought as being the exclusive prerogative of human *esse* with a machine's

---

[61] Wittgenstein (1958: 47).

[62] Hodges (1983: 90–91).

[63] „Wenn man an den Gedanken als etwas spezifisch Menschliches, Organisches denkt, möchte man fragen: ‚könnte es denn eine Gedankenprothese geben, einen anorganischen Ersatz für den Gedanken?' Aber wenn das Denken nun im Schreiben oder Sprechen besteht, warum soll dies nicht eine Maschine tun? – ‚Ja, aber die Maschine weiß von nichts!' – Freilich von einer Prothese des Sehens und Hörens zu reden hat keinen Sinn. Man redet zwar von einem künstlichen Fuß, aber nicht von künstlichen Fußschmerzen.

‚Aber könnte eine Maschine denken?' – Könnte sie Schmerzen haben? Hier kommt es darauf an, was man darunter versteht: ‚etwas *habe* Schmerze.'" Wittgenstein (1974b: 105).

functions. Shanker suggests that Turing proposed to integrate what for Wittgenstein seemed to be necessarily separated: the independent issues in mathematical logic and the philosophy of mind.[64]

Thus, in *Lectures on the Foundations of Mathematics*, Wittgenstein's preoccupation was precisely with this integrative approach of Turing and at that time Wittgenstein had not considered the Mechanist Thesis in connection with Turing's argument. Similarly, in the *Remarks on the Foundations of Mathematics* all that we encounter is Wittgenstein's critique of the philosophical argument given by Turing in his "On Computable Numbers".

Turing presents his argument in two places of this paper. At the very beginning of the paper (§1) he defines 'computing machines' and indicates their function:

> We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions q_1, q_2, ..., q_R which will be called "*m*–configurations". The machine is supplied with a "tape" (the analogue of paper) running through it, and divided into sections (called "squares") each capable of bearing a "symbol". At any moment there is just one square, say the *r*–th, bearing the symbol $\mathfrak{C}(r)$ which is "in the machine". We may call this square the "scanned square". The symbol on the scanned square may be called the "scanned symbol". The "scanned symbol" is the only one of which the machine is, so to speak, "directly aware".[65]

At the core of this thesis is a postulate of the possibility to transform a recursive function into binary terms. Suppose there is a machine that can compute analogues of those functions, given that there is a system of encoded '0s' and '1s'. Hence, the function and the argument, the table of instructions and the tape input must be encoded in binary terms and then converted into an analogue of a binary system. The temptation here is to view Turing machine as a blueprint for a primitive computer. However, according to Shanker's interpretation, what Turing is after is a logical design, which he will actualize five years later by developing the binary code through the use of electrical signals.[66]

In §9 he turns to defend this argument in epistemic terms. It is because of the §9 that the reader (e.g., Wittgenstein) is left with an impression that Turing shifts from mathematics to philosophy while he looks for validation of his thesis. Shanker speaks of the one–to–one correspondence between the computer's behavior and the observed symbols. This is the 'state of mind' of the computer. So, the state of the system consists of the computer itself and the tape.[67]

But that was precisely the object of Wittgenstein's criticism. Wittgenstein argued against bounded mathematical and philosophical postulates: the philosophical argument misinterprets the mathematical one. In the passage above, Turing defines human calculation in purely mechanical terms and then proceeds to speak of the machine in cognitive terms. Wittgenstein, in a way, proceeds in reverse. First, he asks whether a machine calculates, then he turns to the question of the nature of human reasoning. Thus in book V of the *Remarks on the Foundations of Mathematics* he writes: "Does a calculating

---

[64] Shanker (1987: 617).
[65] Turing (1965: 117).
[66] Shanker (1987: 618).
[67] Turing (1965: 136).

machine *calculate*? Imagine that a computing machine had developed by chance, and now someone by chance presses its buttons (or an animal runs over it) and it calculates the product 25 × 20 –".[68]

Shanker makes an important assertion that while in empirical counting we can distinguish the process of counting itself and knowledge of doing the counting or having counted, in the mathematical concept of calculation these two are the same under the notion of 'mathematical normativity'.[69]

In fact, the theme of mathematical normativity seems to stand at the center of the *Remarks on the Foundations of Mathematics* and the *Lectures on the Foundations of Mathematics*, with which Turing's machines argument can be easily integrated. In Book V of the *Remarks on the Foundations of Mathematics* Wittgenstein treats this problem directly, but only after reading Book VI, where Wittgenstein presents his remarks on rule–following, can we comprehensively grasp his objection to Turing's thesis. In discussing the nature of calculation Wittgenstein turns to the argument of rule–following:

> There might be a cave–man who had made for himself *regular* strings of signs. He entertained himself, i.e., by drawing on the wall of the cave
>
> – · – – · – – · – – ·
>
> or – · – · · – · · · – · · · · · –
>
> But he does not follow the general expression of a rule. And we are not saying he is acting regularly just because we can form such an expression.[70]

The question is, therefore, under what circumstances can we say that someone follows the rule? Even though we can construct a rule to describe a certain pattern in someone's behavior, it does not necessarily imply that that individual was following *this* rule, or *any* rule for that matter. In the above example one cannot make mistake in following a rule. According to L. Caruana, "If I cannot in principle be ever mistaken in following a rule beyond its exemplary cases, then there can be no rule-following".[71]

In the following passage Wittgenstein postulates that in order to construct a rule–following behavior, it must be normative:

> If one of two chimpanzees once scratched a figure | – – | in the clay, and another the series | – – | | – – | etc., the first would not have given a rule nor would the second have been following it, no matter what was going on in their minds.

---

[68] „*Rechnet* die Rechenmaschine?
Denk dir, eine Rechenmaschine ware durch Zufall entstanden; nun druckt Einer durch Zufall auf ihre Knöpfe (oder ein Tier läuft über sie) und sie rechnet das Produkt 25 × 20.– „ Wittgenstein (1983: V, §2).

[69] Shanker (1987: 619).

[70] „Es könnte doch einen Höhlenmenschen geben, der für sich selbst *regelmäßige* Zeichenfolgen hervorbrächte. Er unterhielte sich z.B. damit, an die Wand der Höhle zu zeichnen

– · – – · – – · – – ·

oder – · – · · – · · · – · · · · · –

Aber er folgt nicht dem allgemeinen Ausdruck einer Regel. Und wir sagen nicht, er handle regelmäßig, weil wir so einen Ausdruck bilden können". Wittgenstein (1983: VI, §41).

[71] Caruana (2003: 144).

> But if, i.e., we observed the phenomenon of a kind of instruction, and of example and imitation, success and failure of a test, of reward and punishment and the like, and one would have at the end trained a way to put figures that the chimpanzee had not seen before in a sequence, the way it was in the first example, then we could say, that one chimpanzee was writing down rules and the other was following them.[72]

What constitutes normativity is not rule–following itself, which would result in circularity, but any genuine rule–following must be rooted "In the community as it engages in a given practice".[73] However, in the above quoted passage such practice comes only by way of an example in the second paragraph, and not in the description of an accidental pattern production, in the first paragraph.

In his reference to Turing machines, Wittgenstein denies even a possibility that such machines can possess a flock of normative concepts which is the calculation itself in Wittgenstein's account. In the case of chimpanzees as in the case of the 'calculating machines', the question 'How did *x* arrive at the answer?' is not a question of an a priori possibility or impossibility of the calculative function, but rather it is a question about the rules in use. The agent is capable of giving an account of the rule–following in the function of calculating.

Turing's response to this objection was given precisely by indicating that because of the machine's ability to follow the sub–rules of a given program, he was capable of constructing the picture of mechanical calculation. Thus, if chimpanzees are capable of following some simple rules, why should not we ascribe the same capability to machines' activity, viz., of following the rules? To this Wittgenstein responded: "'This calculation is purely mechanical; a machine can do it'". What kind of machine? One that is made of ordinary materials – or a super–machine? Are you not confusing the hardness of a rule with the hardness of a material?"[74]

These quotations could be read in parallel with the anti–computationalist thought experiments of Searle and Jackson. The Chinese Room Argument, the Chinese Nation Argument and the Mary Argument are missing names from Wittgenstein's accounts of these thought experiments. What they had shown, in fact, was not that human thought cannot be compared to calculation, but that this metaphor cannot be the whole story in accounting for human reasoning. Thus, we have here the relation of analogy and/or metaphor, and not that of identity as the AI Thesis' proponents argue for. However, the

[72] „Wenn von zwei Schimpansen der eine einmal die Figur | – – | in den Lehmboden ritzte und ein anderer darauf die Reihe | – – | | – – | etc., so hätte der erste nicht eine Regel gegeben und der zweite ihr gefolgt, was immer auch dabei in der Seele der beiden vorginge.

Beobachtete man aber z. B. das Phänomen einer Art von Unterricht, eines Vormachens und Nachahmens geglückter und mißgeglückter Versuche, von Belohnung und Strafe und dergleichen; würde am Ende der so Abgerichtete Figuren, die er bis dahin nicht gesehen hatte, wie im ersten Beispiel aneinander reihen, so würden wir wohl sagen, der eine Schimpanse schreibe Regeln hin, der andere befolge sie". Wittgenstein (1983: VI, §42).

[73] Caruana (2003: 145).

[74] „'Dieser Kalkül ist rein mechanish; eine Machine könnte ihn ausführen.' Was für eine Machine? Eine, die aus gewöhnlichen Materialen hergestellt ist – oder eine Über–Maschine? Verwechselst du nicht die Härte einer Regel mit der Härte eines Materials?" Wittgenstein (1983: III, §87).

point here is that these ideas were already proposed by Wittgenstein in the 1930's and the 1940's at the rise of computationalism but without the fancy names of the well–known much later thought experiments in the philosophy of mind.

Turing's thesis proceeds by implying that the whole description of calculus can be given as purely mechanical, viz., the rules of calculation are divided into a series of meaningless sub–rules, which escape our conscious awareness and the description of them in the process of calculating. Thus, the final response of Wittgenstein, as we have seen above is as follows: "Turing's machines are really *humans* who calculate".[75]

It would be wrong to assume that Wittgenstein was ascribing to Turing's understanding of the machine's calculating as a human activity. At this point, the notion of effective calculation (algorithms) must be introduced in order to give a comprehensive account of the Turing's thesis. For it was through Turing's clarifications of the algorithms, which he inherited from Hilbert, that he came to the conclusion that the notion of computability can be explained in terms of mechanical procedures. Thus, the nature of effective calculation stands at the core of Turing's thesis in "On Computable Numbers".

As was stressed before, the reason for Wittgenstein's involvement with Turing's Thesis was Wittgenstein's understanding that it was a distortion of its mathematical content. Shanker notices that, "The most basic principles of his approach to the philosophy of mathematics demanded that the epistemological thread in Turing's argument be severed from the mathematical".[76]

It needs to be mentioned, however, that the success of Turing's thesis in part lies in the long preceding discussion of the *Entscheidungsproblem* before the publication of "On Computable Numbers". It was Hilbert's attempt to reduce transfinitary mathematical truths to finitary. Hilbert's postulate was the basic epistemological premise that the human mind is restricted by its limitations. The creation of a computing machine must then anticipate this limitation by fixing it in advance.

The effort of Alonzo Church was to construct analogues of the integers and the algorithms to be performed on the integers by using λ–calculus. He also came to the definition that λ–definable functions are calculable functions: "in the case of any λ–definable function of positive integers, the process of reduction of formulas to normal form provides an algorithm for the effective calculation of particular values of the function".[77]

This, of course, is far from conducting any analysis of the epistemological concepts in question. Church's Thesis, indeed, was limited to defining the range of number–theoretic functions for which there are algorithms. The notion of recursivness was employed for the sake of 'effectively calculable functions'. Shanker asserts that if some new method would demonstrate the existence of non–recursive functions they could not be effectively defined.[78]

---

[75] "If calculation appears to us as a mechanical activity, then *the person* who calculates is the machine". „Wenn uns das Rechnen als maschinelle Tätigkeit erscheint, so ist *der Mensch,* der die Rechnung ausführt, die Maschine". Wittgenstein (1983: IV, §20).

[76] Shanker (1987: 623).

[77] Church (1936: 349).

[78] Shanker (1987: 624).

For Wittgenstein, the fact that there cannot be a nonrecursive effectively calculable function is a logico–grammatical fact. For only in use can this thesis be appreciated as certain, and not by the way of induction. Church Thesis cannot be refuted not because of its irrefutable premises, but because any doubt must be logically excluded. This theme resembles Wittgenstein's latter epistemology: "With the word 'certain' we express the utter conviction the absence of any doubt, and by doing this we seek to convince others. That's *subjective* certainty. But when is something objectively certain? – If an error is not possible. But what kind possibility is it? Mustn't the mistake be *logically* excluded?"[79]

Despite all the clarity with which Church has presented his thesis, his postulates represent a significant gap, viz., he offers no proofs for the claim that effectively calculable functions are effective procedures, in the way it was presented by Hilbert. Instead, Church only affirmed that effective procedures are algorithms. Turing's "On Computable Numbers" offered a new approach that filed the gap and offered new perspectives and solutions to the problem.

In large, building upon the conclusions of Hilbert and Church, Turing has given an analysis of algorithmically calculable functions (initiated by Hilbert but by and large ignored in Church). Next, and most importantly, Turing redefined the notion of effective procedures with much wider epistemological implications.

But it would be wrong to assume that Turing has postulated any kind of intelligence to the machines that would perform such functions. On the contrary, any cognitive abilities of a machine emerge only as *an effect* of the programs:

> Let us suppose we have set up a machine with certain initial instruction tables, so constructed that these tables might on occasion, if good reason arose, modify those tables. One can imagine that after the machine had been operating for some time, the instructions would have altered out of all recognition, but nevertheless still be such that one would have to admit that the machine was still doing very worthwhile calculations. Possibly it might still be getting results of the type desired when the machine was first set up, but in a much more efficient manner.[80]

This implies that the complexity of the program and its following of small sub–rules represents something of an intelligent character. Turing continues: "In such a case one would have to admit that the progress of the machine had not been foreseen when its original instructions were put in. It would be like a pupil who had learnt much from his master, but had added much more by his own work. When this happens I feel that one is obliged to regard the machine as showing intelligence".[81]

From Hilbert's thesis that all number–theoretic functions are recursively calculable, Turing proceeded to affirm that all effective number–theoretic functions are mechanically calculable. Turing has shown that his computable machine was as powerful as a

---

[79] „Mit dem Wort 'gewiß' drücken wir die völlige Überzeugung, die Abwesenheit jedes Zweifels aus, und wir suchen damit den Andern zu überzeugen. Das ist *subjektive* Gewißheit.

Wann aber ist etwas objektiv gewiß? – Wenn ein Irrtum nicht möglich ist. Aber was für eine Möglichkeit ist das? Muß der Irrtum nicht *logisch* ausgeschlossen sein?" Wittgenstein (1969: §194).

[80] Turing (1986: 122–123).

[81] Turing (1986: 123).

mathematical computer. What Turing has achieved in "On Computable Numbers" was a transition from recursion theory to computer science.

According to Turing's thesis, his machines are the only automatic formal systems we need. Turing has proven that we need only one, universal machine, which can do the work of all possible automatic formal systems. Interestingly enough, such a machine does not have to be complicated. According to Haufeland's explanation, all that the Turing machine does is to transcribe some set of elementary rules into a well–organized form.[82] All that such a machine needs is some basic abilities in order to encode and to follow arbitrary rules in a form specified by the program.

Among the mathematicians who rejected this later Turing postulate and whose views often paralleled those of Wittgenstein, was Gödel. The notion of the analysis of algorithmically calculable functions was accepted by Gödel, as well as Wittgenstein. Gödel went even further in giving credit to Turing for redefining some important epistemological notions, viz. that Turing has offered some significant epistemological interpretations and perspectives to the points lacking in Church's thesis concerning the mathematical characterization of a class of functions. But when it comes to the description of the nature of thought and the nature of calculation being mechanical, Gödel refuses to see much common ground.

The difference between mechanically and humanly effective procedures is *qualitative.* Gödel insists that human procedures transcend the mechanical ones. For Gödel, Turing's thesis has a certain contradiction: either the human mind can decide more theoretical questions than any given machine, or there is a certain number of theoretical questions that are undecidable for human mind. According to Gödel (and Hilbert), the second variant must be rejected:

> So the following disjunctive conclusion is inevitable: *Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exists absolutely unsolvable diophantine problems of the type specified* (where the case that both terms of the disjunction are true is not excluded, so that there are, strictly speaking, three alternatives).[83]

Wang interprets this by saying that in this case it would mean the irrationality of the mind asking questions it cannot answer, yet at the same time insisting that only the human mind can answer them.[84]

However, from the point of view of the parallels between effective and mechanical procedures, it is irrelevant whether or not we might know *that* they will terminate. Webb concludes that the execution of a procedure does not depend on whether it can or cannot terminate, but rather on whether the execution of the 'atomic tasks' of that procedure is possible.[85]

---

[82] Haufeland (1985: 139–140).
[83] Gödel (1995: 310).
[84] Wang (1974: 325).
[85] Webb (1980: 224).

Indeed, Turing's undecidable sentences are the "guardian angels of computability theory" that actually saved the Mechanistic Thesis.[86] The whole point of learning programs is that we cannot always predict how they will develop. This seems to be in accord with what Turing affirmed in his lecture of 1947 quoted above.

But the question which Gödel answers negatively, whether mechanical thought and thought can be co–extensive (although Gödel has accepted partial co–extensivity) stands at the center of Wittgenstein's concern with the thesis of AI. The question for Wittgenstein thus, is whether epistemology has anything to do with the difference between mechanical and effective procedures.

Turing's theory of AI consists in the complexity of the program which the machine follows, and not the individual steps of the algorithm. Turing postulated the possibility of moving from fixed to self–modifying algorithms. In other words, instead of repeating the same basic steps, a machine can alter its program and increase the level of sophistication and the performance of its tasks.

In a mechanistic interpretation of the AI, this postulate of Turing has been employed in arguing that a 'human computer' in its rule–following needs intelligence. The computer, viz. a computing person, must understand the language in which the rules are formulated. Turing's Machine was designed to bypass this 'intelligence problem' by reducing all rules to these simple three: to scan, to print and to erase symbols on a tape while it moves. These rules and their following presume no intelligence and that, for Turing, implied the possibility of a consistent psychology of computation.

This was precisely the problem with which Wittgenstein was so much concerned, viz. the assumption that by scanning, printing, and erasing symbols on the tape, the machine shows its ability to follow 'meaningless sub–rules', or 'atomic rules'. Shanker speaks of two aspects of the concept of algorithms that were under careful investigation by Wittgenstein: first, that it makes sense to speak of types of meaningless rule–following and, second, that the thesis of mechanically following a rule is correct.[87]

In *Philosophical Investigations* Part 1, §§185–242 Wittgenstein argues that a pupil can learn the Euclidian algorithm correctly without knowing why the answer is such and such.[88] But the whole point of the rule–following is much more complex. To say that one has mastered the rule means to assume that such a person is capable of explaining and justifying the rule. Thus, to master an algorithm means much more than just learning each of the sub–rules, without understanding the general pattern of the functions of its atomic tasks.[89]

To understand that q follows from p is to grasp the nature of the conceptual relationship between the meaning of q and the meaning of p. It is to know that p implies q. If someone accidentally presses the knobs '25', '×', and '20' of a calculating machine and gets the result '500' we cannot say that this someone has calculated that '25 × 20 = 500'.[90]

---

[86] Webb (1980: 202).

[87] Shanker (1987: 634).

[88] Wittgenstein (2001: §185).

[89] Wittgenstein (2001: §194).

[90] Wittgenstein (1983: VII, §61).

Turing proceeds by giving an odd interpretation of the nature of calculation by reducing it to a stage where mechanically following a rule and following a mechanical rule result in being the same. Following simple rules in his interpretation becomes the same as performing mechanical operations.

An ability to calculate involves the mechanical application of sub–rules, whether they are meaningless or not. The sub–rules of the algorithm can be interpreted as meaningful in some categorical sense. This would allow us to speak of symbolic systems as possessing their own original meanings and therefore, of the brain having its original meaning. Haufeland speaks of this faculty as 'artificial intelligence' proper.[91]

In this respect, Wittgenstein did not deny that someone can perform certain calculating steps mechanically, but for him it did not involve the issue of doing it consciously or unconsciously: "One follows the rule 'mechanically'. Thus, one compares himself to a 'mechanism.' Mechanically means: without thinking. But *absolutely* without thinking? Without *thinking about* [it]".[92]

Wittgenstein has nothing against following rules mechanically, but for him it is not the same as speaking of such behavior as a rule–following. Someone's actions can be considered as rule following *only* if this person is capable of giving intelligible reasons for her/his behavior. Thus, when we calculate we perform many rules unreflectingly. But we are capable of giving an account of our behavior if someone would ask us to do so. Shanker notices that for Wittgenstein the agent must be capable of justifying her/his actions by reference to the rule, and that rule must be accounted for by the one who follows it.[93]

Since Turing missed this important difference, Wittgenstein insists that Turing's basic fallacy consists in considering the simplicity of the elementary rules of algorithm to be the same as following them mechanically. In other words, from the fact that we might follow such rules unreflectingly it cannot be concluded that they are non–cognitive and that a machine can follow them. This is why Wittgenstein insists that Turing machines are actually humans who calculate. In his later article, Turing met this challenge by reasoning that: "It is possible to produce the effect of a computing machine by writing down a set of rules of procedure and asking a man to carry them out. Such a combination of a man with written instructions will be called a 'Paper Machine'. A man provided with paper, pencil, and rubber, and subject to strict discipline, is in effect a universal machine".[94]

Turing insists that the calculation of such a 'Paper Machine' is indistinguishable from the basic calculation techniques performed by an artificial computer. Humans and machines would share at least that much (or that little) in common. Shanker stresses the historical context in which Turing was compelled to extend his mechanistic thesis to an epistemological premise.[95]

---

[91] Haufeland (1985: 119).

[92] „Man folgt der Regel ›*mechanisch*‹. Aber *ganz* ohne zu denken? Ohne *nachzudenken*". Wittgenstein (1983: VII, §60).

[93] Shanker (1998: 160).

[94] Turing (1969: 9).

[95] Shanker (1987: 642).

Without the epistemological arguments offered by Turing, Church's Thesis lacks strength. There is common agreement in the literature on the importance of Turing's contribution to the Mechanist Thesis. Based on this framework, Turing proceeded with his argument in §9 in "On Computable Numbers", and his analysis stands as a definition of an important epistemological notion.

In conclusion, in terms of the analogy between calculation by a human mind and a machine, Turing's and Church's theories are correct to point out the identity of mechanical behavior. But that does not account for reasoning and thinking which are much more than just calculation. But that does not account for reasoning and thinking, which are much more than just calculation. Furthermore, calculation cannot be identified with rule-following. Turing's use of algorithms had simplified mechanical calculations by formalizing them. Bringing in the epistemological premise was counterproductive to the whole notion of the mechanist thesis, if mechanism here means eliminating thinking from algorithm creation for the sets of rules. After Turing, the AI thesis had taken almost for granted that mechanistic and epistemological thesis would go together, or that one would imply the other, thus mechanism would include intelligence and intelligence would imply mechanism.

It is important in the study of AI to have a more thoughtful investigation of the epistemological premises, similar to what is offered by Turing in "On Computable Numbers". For the study shows that language–use and reasoning do require rule–following. Differences between a machine's rule–*processing* and an human's rule–*following* is not only quantitative but qualitative as well.

2.1.2 *Mechanism*

As was mentioned above, the computational theories of mind have their theoretical roots in the mathematical crisis of the 19th century, starting with the emergence of non–Euclidian geometries and Cantor's theory of sets. In 1918, Hilbert addressed this crisis as the crisis in the foundations of mathematics and made two important postulates. First, the language of mathematics (axioms, definitions, etc.) must be formalized, viz. they must represent a formal system composed of symbols which then should be defined in provable formulas. Second, these formal systems must be consistent, or true in the sense of not being in violation of the Principle of Non–Contradiction ($A \cdot {\sim}A$).

Gödel's Completeness and (two) Incompleteness Theorems should be regarded as responses to Hilbert's demands known as 'Hilbert's Program'. First, in 1930 Gödel published his Completeness Theorem which states that the system of the first–order logic (of Russell's *Principia Mathematica*) is complete. In what sense is it 'complete'? It lists a number of logically valid (thus, logically true) sentences according to purely mechanical rules of inference. It is here that the idea of a Turing machine can be justified: such a formal system can be organized into a sequence of rules followed mechanically.[96]

[96] That computable numbers are also enumerable goes in accord with Gödel's two incompleteness theorems and that Hilbert's second problem has no solution Turing mentions in the introduction of his 1936 paper in Turing (1965: 116–153).

Gödel's Completeness Theorem comes as an application of Hilbert's first postulate. However, in 1931 Gödel proposed two Incompleteness Theorems which demonstrate the impossibility of fulfilling Hilbert's second postulate. The First Incompleteness Theorem states (in continuation with his previous Completeness Theorem) that in any first–order system of arithmetic (true and consistent), there are arithmetical propositions which are true but cannot be proven logically in such a system. If some system *P* is consistent, then there are sentences in *P* which cannot be proved in *P*.[97] What 'consistent' means here is that there can be no proven contradiction by means of logical proofs. In other words, if *P* is a logically consistent system, then it contains a sentence which is neither provable nor refutable from *P*. But that does not mean that it is neither provable nor refutable *in general*. Therefore, Gödel continues with the Second Incompleteness Theorem by saying that if *P* is consistent, then its consistency is not provable from *P*. Theorem XI states that: 'The consistency of *P* is not provable in *P*".[98]

Has Gödel demonstrated that there are some truths that cannot be proved in any absolute sense? If so, his First Incompleteness Theorem would run something like: 'If some system *P* is consistent, then there are sentences in *P* which cannot be proved,' omitting 'in *P*.' And it is often omitted in the interpretations of this theory. Several formulations of the theory by Gödel himself show that this small addition is very important: "The proposition that is undecidable *in the system PM* [Principia Mathematica] still was decided by metamathematical considerations".[99]

Jaakko Hintikka stresses this point by saying that the very same proposition which is not provable in one axiom system of arithmetic can be proved in another formal mathematical system. There are no absolutely unprovable propositions in arithmetic in general; a proposition is unprovable only in the same system in which it was formulated. In mathematics, old axioms are constantly giving way to new proofs found through new theorems.[100]

Torkel Franzén makes further remarks on the demarcation of the Gödel's incompleteness theorems. What the theorems claim (especially the First Theorem) is that sentences in a consistent formal system *P* cannot be proved or decided in that system, but they can be successfully proved in some other system.[101]

Gödel's incompleteness theorems made many important philosophical contributions, of which the notion of computability is, perhaps, the greatest. Alan Turing's work in this matter should be seen as a continuation of Gödel's own proposals. Gödel himself acknowledged it in a note to his "On Formally Undecidable Propositions" in 1963. He mentions Turing's work in relation to his incompleteness theorems:

> In consequence of later advances, in particular of the fact that due to A. M. Turing's work a precise and unquestionably adequate definition of the general notion of a formal system can now be given, a completely general version of Theorems VI and XI is now possible. That is, it can be proved rigorously that in *every* consistent formal system

[97] Theorem VI in Gödel (1986: 172).

[98] Theorem XI in Gödel (1986: 192). Translation is by S. C. Kleene.

[99] Gödel (1986: 151). Translation is by S. C. Kleene.

[100] Hintikka (2000: 37).

[101] Franzén (2005: ch. 2).

> that contains a certain amount of finitary number theory, there exist undecidable arithmetic propositions and that, moreover, the consistency of any such system cannot be proven in the system.[102]

As was presented earlier, Turing's main thesis in his 1936 paper can be reduced to the answer to the question 'what is computable?' For Turing, computability means mechanistic rule–following by an ideal machine (Turing Machines). If Gödel's incompleteness theorems are applied to the Church–Turing Thesis, it would necessarily mean that computers (in Turing's sense) are limited to just a mechanical rule–following and that the human mind is incompatibly more powerful. In his 1951 paper, quoted in the preceding section, Gödel himself states this problem.[103]

Could the above quote commit Gödel to saying that human mind is incompatible with computers? Authors like Lucas, Penrose[104] and Hofstadter[105] seem to imply that. Due to the lack of space, I will only briefly present here the proposals of Lucas, since he was the first to draw attention to Gödel's incompleteness theorems in proving the minds' consistency.

Lucas has argued that Gödel's theorems refute mechanism and allow us to distinguish between conscious and unconscious beings. Mechanism is considered by Lucas to make no differentiation between a self–conscious mind and a non–conscious machine. Against mechanism Lucas uses Gödel's theorems in showing that the human mind can "consider itself and its performance and yet not be other than that which did the performance".[106] This is the major difference between the human mind and a machine. Unlike machines, the human mind can assert its own consistency.

The above interpretation, perhaps, would make sense if we were to adopt an 'incomplete' version of the First Incompleteness Theorem considered earlier and accept an absolute undecidability in mathematics. For example, there could be a true but *absolutely* unprovable arithmetical proposition that was generated by the human mind (but never by a computer) from any of the defined rules of inference. Yet, Gödel never proved the existence of such propositions or the consistency of mathematics.

In a critical paper, Judson Webb notices that Lucas is reading too much into what Gödel's theorems were supposed to do. They were never meant to distinguish between conscious and mechanical operations. The issue of the consistency of the human mind cannot be resolved by Gödel's theorem and it cannot be *shown*, because a demonstration would always require some kind of reference, or logic by which rules it would be consistent. But there is no one logic which would declare it consistent without there being another one by which it is not. This goes in hand with what Gödel himself had shown in his 1951 lecture, quoted above, where the issue of the consistency of the human mind is not the subject of any proof.

---

[102] This is a much later addition to his 1931 paper and appears only in English: "Note added 28 August 1963", in Gödel (1986: 195).

[103] Gödel (1995: 310).

[104] Penrose (2002); Penrose (1996).

[105] Hofstadter (1979).

[106] Lucas (1961: 124–125). With many critical replies, such as mentioned already, see Webb (1968); Franzén (2005); Hintikka (2000).

A similar argument was developed later by Putnam in his 1960 paper,[107] although in that very same text he concludes that the mind is after all a machine, both inconsistent, and that Gödel's theorems cannot be applied to humans. While we know that a machine is inconsistent, we cannot in principle prove consistency by the Turing machine.

Lucas's notion of consistency that comes from Turing and Gödel is not enough to prove that the human mind is in some way superior to calculating machinery. As Stewart Shapiro explains, if by mechanism we mean mechanical rule–following then both, mind and machines are mechanistic under this view. If, on the other hand, by mechanism we mean lack of consistency, then we have no way of proving that the mind is anymore consistent than a Turing Machine, and Gödel's incompleteness theorems are of no help to us (or to Lucas).[108]

We can pass from one system to another, from one language-game to another without being restricted by any formal system. In this sense the mind is "incompatibly more powerful" than any machine, and not that the mind is incompatible with any machine in principle. The issue is not that minds and machines are incompatible in principle, but that their compatibility is rather very limited and somewhat superficial.[109]

In the debate between Wittgenstein and Turing on the notion of what thinking is, Wittgenstein insisted that thinking cannot be reduced to mechanically following rules of inference. As Hintikka has shown, by extending such rule–following to rule–following in a language–game, even a mere following of the rules of inference would require strategies (of that game) in order to *know* which rule to apply first.[110]

In sum, the problem of consistency for the AI thesis states that while we know (thanks to Gödel's theorems) that machines are inconsistent on the basis of them being closed, true systems, we not only do not know if the human mind is consistent or not, we have no possibility of proving either assertion. Without that, the identity between a human and the machine's even basic calculation remain no more than a metaphor.

2.1.3 *Connectionism*

Computational theories of mind aimed at a double reduction: the reduction of epistemology to mechanism and the reduction of meaning to grammar, semantics to syntax. The second was an obvious consequence of the first, and the first was historically situated within the larger philosophical and scientific context of logical positivism in its many forms. Ultimately, philosophy would be entirely excluded from the laboratory of the explanation of mind. 'Mind are computers' was not meant to be an analogy but an eliminative identity statement. However, without elaborating a satisfactory epistemology and consistent semantics, it remained no more than a working metaphor.

---

[107] Putnam (1960).

[108] Shapiro (1998).

[109] Dreyfus (1972); Dreyfus (1992); Dreyfus (1986); Harel (2000).

[110] Hintikka (2000: 70).

Starting with[111] the works of Rosenblatt,[112] but most importantly with the works of Rumelhart and McClelland,[113] a different computational theory of mind was proposed. It is based on a non–serial distribution of the levels of neuronal networks. This model, called connectionism, was meant to explain especially language acquisition and use by modeling the neural networks distributions which occur on sub–symbolic levels. It would seem not to require that the language of thought and its explanations be from the bottom up.[114]

Whether connectionism can dispense with the language of thought is contentious and it was noticed that at least some structures of the language of thought was still required even at a sub–symbolic levels.[115] But if connectionism is meant to be more than just a comparative model projected on how the brain works, it must explain the relationships between the model itself and the brain networks. Connectionist models are representations of some particular brain functions but so far they were not able to provide a global account for the mind. In Armstrong's argument, at least intuition tells us that the mind is not a collection of bundles somehow brought together, but a unity working together.[116]

Connectionism tries to explain cognitive abilities using artificial neural networks. These neural networks, or nets, are simplified models of the brain: just as the brain can be viewed as a collection of interconnected neurons, these models are a number of interconnected units. Connections between these units are measured by weights indicating the strength of the connections. While neurons in the brain are connected via synapses, the weights of the connections between the units indicate the strength of the units' communicability. Connectionism is especially successful in explaining face recognition and identification of basic grammatical structures. While the classical computational theory of mind is strongly based on the idea of rule–following according to formal logic, connectionism seems to avoid such conceptualization by its explanatory model of networks connections. However, the idea of analogy between the human mind and the computer remains the same in both classical computationalism and connectionism. The major difference between the two is that connectionism does not claim that the human mind is a processing device of symbolic language, or that symbolic language is a necessary component in the cognitive abilities of the brain.

The central idea of connectionism is that connections are patterns within neural networks. These networks are seen as units themself consisting of inputs, outputs and hidden units. The usual way of interpreting inputs is the analogy with sensory neurons and that of outputs is motor neurons. The network receives the information through the

---

[111] Justin Leiber claims that Alan Turing in 1948 was the first one to suggest that the cortex is a universal machine and that the machine resembles the architecture of interconnected neural networks. He referred to the infant's cortex as "an unorganized machine, which can be organised by suitable interfering training". Leiber (1991: 16); Turing (1969); Copeland and Proudfoot (1996).

[112] Rosenblatt (1958).

[113] McClelland and Rumelhart (1986).

[114] Smolensky (1987).

[115] Fodor and McLaughlin in Macdonald and Macdonald (1991).

[116] Armstrong (1968); Armstrong (1997).

input and processes it in the hidden units with the resulting processed information in the outputs.

The most common kind of networks is a feed forward net, where the activation of information flows directly from the inputs through hidden units to the output units. There are, however, more realistic models in which the hidden units, represented as many layers, and connections are considered to be sending recurrent information from higher to lower levels. These realistic models are employed in explaining short–term memory, while feed forward net models in processing and learning basic grammar.

Connectionists believe that these models can explain to some degree the cognitive functions of the brain, seen as a giant information–processing center. Since information is processed in the same way, what can serve as the major explanation is the difference in the weights of the connections between the units. Weights can be positive and negative. The weight is positive if the sending unit excites the receiving unit, and it is negative if the receiving unit inhibits the activity of the sending unit.

Thus, the major goal in the connectionist program is to find the right number of weights in performing each cognitive task. This is accomplished by the adjustment of the algorithms and by designing special training sets. One of the most common applications of such techniques is backpropagation, especially successful in face–recognition skills. For example, different weights are attached to the different features of faces, such as gender, age, expressions, etc. The weights are adjusted in order to receive the desired values of the net's output, and after many repetitions (often as many as hundreds of thousands of rounds of weight adjustments[117]), the network can learn correct face recognition.

Some criticism was directed toward connectionism in terms of a mismatch of the connectionist paradigm of net learning and the way in which humans and non–human primates learn. Sometimes learning occurs from single events, while the connectionist model of learning, especially in backpropagations, requires much training and repetition.

One of the best known models was a trained network by Rumelhart and McClelland that was able to predict the correct past tense of English irregular verbs. The net was trained to predict correct forms of past tense irregular verbs, such as 'come / came', 'go / went', and correctly predict forms of the verbs not on the list.[118]

Perhaps the most significant contribution by connectionism to cognitive science was the idea that single concepts are not localized in some single parts of the brain but are rather distributed across large areas. The idea of a Parallel Distributed Processing (PDP) is quite different from original connectionism. These distributed representations, chiefly in the hidden units, were the result of many connectionist models' training. This was especially applied as an alternative account of meaning to that of the classical computationalism. According to the later, all semantic representations are composed of symbolic atoms, the same way language is composed of words, words or letters and letters of symbols.[119] In such a view, meaning is strictly dependent on the syntax of language, but syntax itself, in this atomic semantics view, has no meaning.

---

[117] Garson (2010).

[118] Rumelhart and McClelland (1986).

[119] Garson (1997).

Connectionism avoids this problem by eliminating atoms altogether. Distributed representations are patterns of a dynamic activity across all the units. The information transmitted does not need to be interpreted itself. This view was suggested by Horgan and Tieson under the name of 'representations without rules'.[120] The main idea is to view representations not as steps or rules of a computer program, but rather as regularities, or 'soft laws'.[121]

This later point comes as a part of a much larger debate on two major topics in relation to connectionism: the necessity/dispensability of rules in higher cognition and the semantic similarity problem. Often this debate, comprising the above mentioned issues, is known as the 'systematicity debate'. The core issue here is the claim that the connectionist model can only be applied to association–processing, while language and reasoning in higher–cognitive abilities need something more than association. The response was that connectionist models *can*, in fact account for higher–cognitive abilities if it implies some of the methods of classical computationalism.

Connectionism is often seen as the alternative to the classical computational theory of mind which claimed that information is processed as a set of symbols. Connectionism challenges this theory by considering information as connection strengths between the units of neural networks, represented in different weights. Connectionists view cognition as the dynamic evolution of network activities which depend on the strengths of connections within each unit and its 'neighbors'.

Despite these obvious differences, some attempts of bridging the two views have been made. Implementational connectionism, for example, suggests viewing cognition as the brain's implementations of a symbolic processor. Neural networks are implemented precisely in conducting symbolic processing.

Notwithstanding this notion of connectionism, Fodor and Pylyshyn speak of human intelligence in terms of systematicity, which connectionism cannot explain.[122] The notion of systematicity refers to the ability of human intelligence to make connections between one idea and another. In this view, the connectionist model cannot make the connection between the sentence 'John loves Mary' and 'Mary loves John'. In the classical computational view, however, such a string of thought can be put into easily discernible rules of symbolic logic. Fodor and McLaughlin further conclude that, by connectionist model standards, human intelligence seems to lack systematicity.[123] In a rebuttal, David Chalmers and Ken Aizawa point out that what Fodor, Pylyshyn and McLaughlin demand is a nomic necessity, such as that postulated by Saul Kripke,[124] which neither connectionist nor classical computational models can accommodate.[125]

The second problem of the semantic similarity emerges in relation to the connectionist model's interpretation of different meanings of brain states. Different patterns in neural activities share some similarities which in turn can be interpreted as intrinsic properties

[120] Horgan and Tienson (1989).
[121] Horgan and Tienson (1990).
[122] Fodor and Pylyshyn (1988).
[123] Fodor and McLaughlin (1991).
[124] Kripke (1971).
[125] Chalmers (1993); Aizawa (1997).

in determining semantic information. Fodor and Lepore indicate that human brains are too diverse in their architecture and in the way neural networks process information.[126] Learning, acquiring and using concepts vary from person to person and depend on personal knowledge, culture, history, etc. As neural network activities cannot account for even basic concept formation in a single person, they are not at all suitable for a general theory of meaning, whichever theory one may decide on.

Most of these philosophical issues concerning connectionism and computationalism essentially touch a more basic philosophical problem related to the notion of concept. The classical or essentialist notion of concept looks into the common features of all instances of a given concept and considers it to be the reason why such a concept is used. In this view, under the concept 'tiger' would fall a large, black and orange feline. Thus, 'a large black and orange feline' is a necessary and sufficient analytical condition for the concept 'tiger' to exist. Do not albino tigers fall under that concept? Apparently not, but that is a contradiction.

Wittgenstein's notion of a *'Familienähnlichkeit'*, 'family resemblance', can serve as an alternative notion of concept and possibly fit into the general understanding of connectionist distributed processes in cognition. With this notion of the 'family resemblance' Wittgenstein suggests comparing different language–games in order to see what sort of similarities there are between them. When we do that, we notice that several language–games share the same complex network of related similarities. Just as in one human family each member does not share the same feature (i.e., everyone's noses have similar features), yet they resemble each other in one feature or another. Similarly with concepts: we should not look for the same feature underlining all instances of its use, but rather commonalities shared by the same uses in which that concept occurs, viz. in similar language–games.[127]

Wittgenstein's notion of family resemblance interpreted as language–games avoids the comparison of concepts to types and of representations to tokens (in the Peircean type–token distinction). The notion of concept as type would fall under Wittgenstein's criticism of "craving for generality".[128]The family resemblance of language–games, on the other hand, refuses generalizations and demands an analysis of games rather than searching for one abstract feature in all of them.

It is important *not to* consider 'family resemblance' as an on–its–own term of Wittgenstein philosophy, something equally significant as his language–games or picture theory, but rather as a qualifying notion of language–games. Thus, concepts are acquired as a result of playing (*in*–acting) different language–games of a family resemblance similarity.

Wittgenstein's notion of language–games complicates the 'story' of connectionism or of the PDP, but it also makes this paradigm to resemble more the reality of the neural nets of the brain. Some of the major problems of connectionist model are philosophical in nature; hence the solutions should also be philosophical.

Furthermore, according to the criticism of classical computationalism, if the mind does follow some nomic based rules in language and higher forms of reasoning, then

---

[126] Fodor and Lepore (1999).

[127] Wittgenstein (1974b: §§66–71).

[128] Wittgenstein (1958: 18).

connectionism focusing on the neural networks' operations so far did not explain what these rules were.[129] Connectionists (proponents of what is called, 'radical connectionism') sometimes deny that such rules exist at all or that there is any symbolic representation in the mind, and often the affirmation of such rules is considered to be part of a dispensable folk psychology. However, without proving the nomic characteristic of the bottom–up networks' connections, eliminating symbolic representation does not seem to be an option. Hence, connectionism, as a philosophical theory of mind, risks being no more than a descriptive paradigm with an identity of metaphor in use. If classical computationalism implies that mind is a computer, then (non–eliminative) connectionism implies that mind is dynamically distributed connections between neural nets, decided by the differences in these connections' weights. As it was mentioned before, at least for higher cognition this metaphor is not enough.

2.2 Functionalism and the Multiple Realizability Thesis

Mental states are formed by causal relations between each other, perceptual inputs and behavioral outputs. Mental states are functional relations between these three. 'Mind', in short, is a function of the brain. This view is called 'functionalism' and it has amassed several theoretical sources and many variations since it was proposed by H. Putnam, D. Armstrong and D. Lewis.

Ned Block lists three sources of functionalism. First is Putnam's and Fodor's computational theory of mind. Second is Smart's topic neutrality which, as was mentioned in the preceding section, was interpreted by Armstrong and Lewis as a functionalist conception of the mental states. Third is later–Wittgenstein's notion of semantics at the core of which meaning is defined as use, later developed by Sellars and Harman.[130]

The first two came as a correction to the type–identity theory forwarded by Smart and Place, but the corrections were quite different. Since Armstrong's and Lewis' contribution to the discussion was already mentioned before, the following presentation will focus only on the first source and sort of functionalism, which became not only the most influential functionalist theory in the philosophy of mind but also remains the most prominent interpretation of the mind–body relation up to day.

Fodor and Putnam indicated that type–identity theory focuses on the identity relationship between mental and neuronal states, such as pain and C–fiber stimulation. But that would limit the identity theory only to those organisms (and living organisms only) which have neuronal structures that can accommodate such analogies. That would imply a 'species chauvinism' that would not accommodate, i.e., computing machines, AI, extraterrestrials and those organisms that might have pain without having any C–fibers.[131]

---

[129] Smolensky (1991).

[130] Block (1995: 323).

[131] It is interesting to note, that in his 2011 article, J. C. C. Smart, while defining type–identity theory, does not use the phrase 'C–fiber stimulation'. It is not because some C–fiber deprived creatures would feel excluded, but rather because these creatures would include human species,

What really matters is not biological constitution or lack of such but what being in pain or being in pleasure contribute to the whole of the given organism, what function they play in the organism's overall state. 'Pain' is a type identified with general functional roles such as inputs, outputs, psychological states, etc., which are the tokens.

Under functionalism, this core analogy of the type–token identity theory was applied within the larger context of the computational theory of mind. Putnam spoke of mental states as the functional states of a computer. Whether realizations of the same type–states are in various hardware configurations or in different biological and chemical inputs/outputs of some given organism, we can speak of multiple realizations of the same types of mental states.

This type–token identity theory with the application of its core idea of a multiple realizability does affirm the distinctiveness of mental and physical states. Mental states are types of their physical states/tokens, even if they do not need to be realized just in an animal's brain. The theory does not equate mental to physical but it also does not affirm an existence of some non–physical substances, thus avoiding substance dualism.

Functionalism paved the way for two major disciplines: cognitivism and the revised version of the AI thesis of Church–Turing. Cognitivism applies much of the functionalist approach by regarding humans as "information–processing systems".[132] The AI project made some progress in making computers perform many tasks for which, generally, a human intelligence is required. What made their project possible was their way of interpreting human intelligence as a complex information–processing device.[133]

This later form of functionalism sometimes is called 'machine functionalism'. In the course of time it was criticized for being too liberal (as opposed to 'conservative' species–chauvinism) and thus too elusive in explaining the causal mechanisms of mind and brain. An alternative form of functionalism was proposed, 'teleological functionalism', according to which there must be a goal–like, a purpose element in the notion of realization. The physical realization must have a true 'organic integrity' with the organism for which it is counted as a function.[134]

But teleological functionalism is incompatible with machine functionalism. The main point of machine functionalism and its main theoretical achievement is to make sure that *regardless* of the organism, or the hardware, the mental state is realizable anyway. Teleological functionalism, on the other hand, insists that the mental state must be realized in virtue of the biological purpose of the organism. To be in pain belongs to the *overall* structure and nature of the organism.[135]

Teleological functionalism raised another issue with machine functionalism in that the later does not account for intentionality, for the aboutness of the mental states, directness

---

whose neurobiology of pain extends beyond C–fiber stimulations. Of course, this identity phrase is situated in the knowledge of neurobiology of the 1950's and lacks the contemporary under-standing of the problem. See Smart (1959); Smart (2011).

[132] Lycan (1995: 319).

[133] Haugeland (2008).

[134] Sober (2008).

[135] Gulik (2008).

of the mind. Teleological functionalism suggests regarding intentionality in terms of reference to the mental states' psychological and biological functions.[136]

Some obvious criticism emerged against this account of intentionality. If biological states are accounted for intentionality or for aboutness of the mental states, then these biological states are teleological themselves. It is one thing to say that every living organism has its purpose in its functioning, it is another thing to insist that this purpose is somehow related to the intentionality of the mental states.[137]

Another problem with functionalism is its inability to account for qualitative mental experiences, called 'qualia'. A series of objections had been proposed in the way that both, functionalism and cognitivism, cannot accommodate this essential feature of human consciousness.

The first of these objections was Thomas Nagel's celebrated insistence on the fact that even if one knows all there is know about some organism's perceptual, biological, neurological and the like functions, we would not be able to know *what it feels like* to be that organism.[138]

Frank Jackson's 'Mary's Room',[139] Lawrence Davis and Ned Block's 'Chinese Nation/Room Argument'[140] thought experiments and the like in different ways approach diverse problems regarding qualia. The literature for and against these thought experiments is enormous, and space does not permit going into analysis of all these. What seems to be the central difficulty is not that these problems emerge but the most fundamental distinction between ontology and metaphysics, which is not clarified in either form of functionalism. Following Ned Block's distinction, if the ontological question is concerned with the *existence* of each type, the metaphysical question is concerned with the *nature* of each type.[141] To be sure, behaviorism and type–identity theories did raise the ontological question. The former eliminates any mental states and explains behavior in terms of the stimuli. The problem, as I have mentioned earlier, is that we do not act just on the basis of the stimuli but on the basis of our beliefs and desires which are mental. The latter, without eliminating mental states, identifies mental and physical states in terms of their type unity. Although both theories did refuse Cartesian dualism, they were neither able to explain the old Cartesian problem of interaction, nor to account for language and qualia.

According to Ned Block, functionalism tells us what different mental states have in common, viz. their function within the overall given organism. But it does not tell us whether the organism is living or a machine. By ignoring the ontological question, functionalism cannot confront the epistemological question of our access to the mental content.

The problem with the qualitative experiences cannot be explained without first answering the question: how do we come to know what we know? The answer cannot be

---

[136] Dretske (1988).

[137] Searle (1979).

[138] Nagel (1974).

[139] Jackson (1982); Jackson (1986).

[140] Block (1978).

[141] Block (1995: 326).

the same for all organisms and it cannot be explained in terms of their functional architecture in the context of the overall organism.

In order to answer the question 'how one comes to know what one knows?', the metaphysical question must be answered first: 'what is it that knows?' The functionalist metaphysical diversion from ontology precludes it from confronting the question about one's knowledge and that makes the problem of qualia unanswerable.

Later functionalism generally adopted the externalist theory of meaning which would reduce the question of intentionality to causality in the physical world. But causality is not enough to account for aboutness of the mental states and their linguistic expressions. Questions of intentionality and meaning are closely related. Despite the fact that functionalism remains the main theory within philosophy of mind and cognitive psychology, it struggles to elucidate the problem of intentionality from its externalist theoretical semantic basis.

A charge against functionalism that it does not explain problems of intentionality in depth was posed by Daniel Dennett, who suggested an amended version of teleological functionalism. His suggestion was to decompose intentionality into hierarchically structured intentional systems and sub–systems. The hierarchy would run according to the degree of intelligence ascribed to neuronal networks, sub–networks, neurons, etc. The lower one gets, the less intelligent neurons are.[142] Lycan called this version of teleological functionalism 'homuncular functionalism'.[143]

Lycan suggested two kinds of interpretation.[144] First, Dennett's interpretation of neurons as intentional intelligent agents must be taken at its face value and then philosophers who dismiss it as a neo–vitalism are right. Second, intelligence and intentionality are not *really* ascribed to particular neurons, networks, brain areas, etc., but they are ascribed analogically or paradigmatically. But then in this later case the explanation of intentionality is no different from teleological functionalism.

Putnam's own disavowal of functionalism begun with his externalistic semantics according to which the meaning of the words does not depend upon how the organism functions but on its relation with the external world.[145] He later called functionalism as a 'Utopian' theory guilty of 'scientism' by suggesting an unrealistic and ideal psychological theory.[146]

What is counted as the major achievement of early functionalism, the multiple realizability thesis, has also been challenged in the last two decades by philosophers of mind, science and neuroscience research. When neuroscience studies the chimpanzee's visual system, it projects the results on humans presupposing a continuity across the species. This continuity is unwarranted if one is to adopt a multiple realizability thesis. John Bickle also indicates that in the cases of neuronal damages following strokes in patients and further recuperation, functions return with time but never to the same level of performance. It is true that in some cases different brain areas will 'step in' to compensate

---

[142] Dennett (1971); Dennett (1978).

[143] Lycan (1981).

[144] Lycan (2008).

[145] Putnam (1975).

[146] Putnam (1995).

for damaged areas, but the person will not have the same plasticity in, i.e., speech or motor action as before.[147]

Among the many philosophical objections against the multiple realizability thesis is Lawrence Shapiro's claim that the thesis was never proved. Take the example of two different corkscrews, one made of steel and the other of aluminum, of different colors, etc. Their physical qualities do not make their functions as corkscrews any different. Their property differences in the way they are realized make no difference in their performance.[148]

The same function is realized in different physical features, but these physical features are often irrelevant for the purpose to be realized. In case of psychological states we can speak of a 'reverse multiple realization', where the same physical states can account for different mental states.[149]

Old myths die hard; Putnam's distancing in the 1970's from functionalism and the objections against multiple realizability thesis did little against functionalism remaining as the main theory in the philosophy of mind. It seems that its continuous success might also be explained by its close affinity to computationalism and the constant interest in the fascinating successes of AI.

The major problems with computationalism and functionalism seem to be concerning logic and language. This is why Wittgenstein's and Peirce's criticisms of even the earliest versions of computationalist theories of mind are important. For them the questions of logic and language were central to their philosophical methods, and, against their philosophies, comparisons between the brain's and the machine's operations (software in the larger meaning) had clear limitations. H. Putnam's disillusion with his own 1960s functionalism in the 1980s is followed by his increasing interest and writing on Wittgenstein's and Peirce's philosophies of language and mathematics from the 1990s on. This, of course, is no coincidence.

The computationalist model of the mind can serve well as an analogy but remembering the limitations any analogy has. Mental states are only partially rule–following, while brain states represent multi–dimensional complexity for which analogies with programming does justice only in part. This also results in an incomplete understanding on the methodological level of how brain states cause mental states.

3. Biological Naturalism

John R. Searle's 'biological naturalism' is a *seul en scène* of the philosophy of mind. It developed as a consequence to his critical response to, what he calls, 'the strong AI thesis': the mind is to the brain as the software is to the hardware. But his philosophy of mind has important links to his internalist theory of meaning in the larger context of his philosophy of mind, especially to his elaboration of J. Austin's speech acts.[150]

---

147 Bickle (2006).

148 Shapiro (2000).

149 On 'reverse multiple realizability', see Hull (1972); Hull (1974); Bechtel and Hamilton (2007).

150 Searle (2007).

Searle points out an important linguistic failure to which most philosophers of mind are still committed, viz. their language of 'mind' and 'body' is profoundly Cartesian even if no one would ascribe today to the Cartesian dualism. The exclusive terms 'physical' and 'mental' go contrary to our commitments to the physicalism of contemporary science and philosophy. While the world is physical there are nonetheless irreducible mental phenomena but these claims should be neither dualistic nor inconsistent.[151]

The answer to this philosophical and linguistic puzzle is his 'biological naturalism'. Mental states are caused by the neurobiology of the brain, but they do not enjoy some sort of separate existence on their own, instead, they are higher–level features of the brain.[152]

Yet, Searle does not operate within an emergentist nor functionalist terminology. Mind is not a mere function of the brain, nor does it supervene on the brain. The analogy used here is that of the liquidity of water or solidity of wood: just as liquidity is an essential feature of water, mind is an essential feature of the brain. Liquidity or solidity is explained by the molecular and cellular features of water or wood. Consciousness is that higher–level brain's feature that is *caused* by the brain's lower–level operations. These are the operations of neurons, synapses, etc.

However, while we can safely assume that the liquidity of water is fully explained by the study of water's molecular structure, consciousness cannot be explained by the study of the brain's lower–level elements and their operations. Searle is aware of this disanalogy: consciousness is a first–person realm, entirely subjective and cannot be reduced to some third–person reality. But this first–person acquaintance with my own subjective states of consciousness makes me infer about consciousness in general.

Knowledge of other minds, however, is not inferred from people's behavior (behaviorism), but rather from the causality of my own to the causality of others'. I have first–person acquaintance with the causal structure of my own states of consciousness and that allows me to infer that others' mental causation must be like mine.

Descartes and Hume had failed to account for mental causation due to their dualistic account of the world. Searle proposes an atemporal account of causation: it is not cause followed by an effect causation but one which is linked essentially to the nature of that which causes and that which is caused. Liquidity of water is causally essential to its property of quenching thirst; consciousness is entirely formed by the neurons but consciousness is essentially important to the functioning of an organism. Both substance dualism and epiphenomenalism are avoided because consciousness is explained only through biological operations of the lower–level brain functions and at the same time causality is explained as an operation of these lower–levels as consciousness operations.

This biological naturalism, as was mentioned before, came as a consequence of Searle's critical response to the AI thesis of considering mind in purely syntactical terms. Searle's response came in the form of the Chinese Room Argument, his most celebrated thought experiment.

A non–Chinese speaker is locked in a room filled with boxes, each marked with a Chinese character. The person is given instructions (i.e., in English) on how to match Chinese

---

[151] Searle (1995); Searle (1992: ch. 2).

[152] Searle (2000).

characters together, and following these instructions, that person combines characters into sentences which are questions, then a new set of rules are given to match more characters into more sentences which are answers to the questions composed. All the person does is to follow the instructions and make the sentences, questions and answers, the meaning of which, nor the reading of which, that person knows. Searle claims that this person would pass the Turing test (human/machine responses will be indistinguishable to an observer/examiner) and should be claimed to be proficient in Chinese. However, since he does not know a single reading of a Chinese character, Searle claims, such a claim is absurd.[153]

Searle refutes the strong AI thesis on the assumption that if programs are syntactical, the mental contents are semantic, hence syntax is not sufficient for semantics, just as an ability to follow rules is not sufficient for knowing their meaning, and programs are not minds.

As I had argued earlier, this argument is very similar to Wittgenstein's criticisms of Turing Machines that are allegedly a perfect metaphor of human mind. Just as Wittgenstein, Searle is not arguing that no comparison between mind and computer can be made ('computer', '*Rechner*' means simply someone that calculates, and '*Rechner*' can also mean 'calculator', whether machine or human, in German). What both did argue, however, is that not all mental states can be explained by the reference to computation.

Searle stresses this point when he affirms that cognitivism is immune to his Chinese Room Argument. Cognitivism affirms that every mental state has a computational structure which makes every mental state a computational process. But Searle refutes cognitivism as well on the grounds that a process can be considered to be computational only relative to some observer and not intrinsically as such. Computation can be assigned to anything but it is not an intrinsic feature of the world.

Computationalism and cognitivism presuppose that conscious processes are blind rule–following which cannot be made conscious in principle. Searle's notion of consciousness refutes this general assumption.[154] His 'Connection Principle' states that every unconscious state can be made conscious in principle and that there is a logical connection between consciousness and unconsciousness. It is explained by an assumption that all mental activity consists of these two features: consciousness and neurobiological processes. These later processes are unconscious *de facto* but not *in principium*.[155]

Searle introduces an important distinction between prior intentions and intentions in action, and insists that consciousness, being a causal agency, directly and dynamically influences both.[156] Top–down causation in cognitive processes requires that some goal would precede the action in order to achieve the desired goal. Experiments performed by Nielsen, Fourneret and Jeannerod had shown that subjects were aware of a goal but not of their own motor performances.[157]

---

[153] Searle (1980); Searle (1984: 31–41).

[154] Searle (1992: ch. 7).

[155] Searle (1992: ch. 3).

[156] Searle (1983: ch. 3).

[157] Fourneret and Jeannerod (1998).

The reason being is that more time is needed for an action to become conscious than for performing that very same action. Benjamin Libet has shown that consciousness of an action emerges after 500 msec, while only 50 msec are necessary for the primary motor cortex to activate the spinal motor nerve cells.[158] Whenever the desired goal is not obtained, due to this discrepancy between the will of the agent and actual performance, fully conscious prior intentions would correct the intentions in action.[159]

One of the most important contributions of Searle in philosophy of mind and, especially in his philosophy of language, is his discussion on intentionality. His notion of intentionality goes in hand with his work on speech acts and it brings him to the two–fold distinction between 'networks' and 'backgrounds', and the mind's directness toward mind–dependent and mind–independent objects of acquaintance.[160]

Intentional states are as such only in reference to other intentional states. Perceiving that something is an *x* comes only within the network of other perceptual experiences. But how can one distinguish between seeing an *x* from seeing that it is an *x*? Searle introduces the notion of background against which we can test out perceptual contents and define them against already acquired knowledge.

That knowledge requires some social background is clear, but why should intentionality of perception require it too? Seeing *a* yellow station wagon does not require seeing *that* it is a yellow station wagon. The reason for Searle to consider these two together rests on his insistence that since experiences have the condition of satisfaction they must also have a propositional content.[161] A condition of satisfaction must be obtained for the mental state to be true. Fred Dretske notices this awkward logic by which 'seeing *x*' becomes equivalent to 'seeing that *x*'.[162]

Against this notion that intentionality must be the satisfaction of one's desires or the correspondence of one's beliefs with the facts of the world, Wittgenstein once remarked that verification of one's beliefs does not always coincide with recognition of its objective. When I wanted to satisfy my hunger by eating an apple but someone punched me in the stomach thus taking away my appetite, it would imply that my hunger was satisfied by the punch and it was the punch in the stomach that I desired, not an apple.[163]

The confusion in question is between intentionality of perception and intentionality of knowledge. Intentionality of perception, or reference, does not include one's knowledge about what one senses or feels; i.e. the mere fact of seeing a yellow station wagon, or sense of hunger, or any other feeling. Some animals and young children have this kind of intentionality: they can perceive a yellow station wagon without knowing what they perceive. On the other hand, intentionality of knowledge would require *knowing that* a yellow station wagon is in front of me and that kind of intentionality would require language and, additionally, acquired semantics.

---

[158] Libet (1985).

[159] Libet (1985: 529).

[160] Searle (1983: ch. 5); Searle (1992: ch. 8).

[161] Searle (1992: ch. 2).

[162] Dretske (2003).

[163] Wittgenstein (1975: III, §64).

Intentionality as a network without the notion of a background can be seen as a circular argument. The background in intentionality is also introduced to safeguard intentionality from an endless regress of explanations of one's intentionality of desires and beliefs. But in this way intentionality seems to be an overly robust ontological notion, perhaps also exaggerated in defending against computational reductionism.[164]

Searle further insists that our use of concepts are of the two kinds: socially–imposed and world–structured. 'President', 'student', 'money' are all socially–imposed concepts, they depend on the social construction and the use by society. It comes from his internalist notion of meaning: the explanation of the relation between mind and reality must be based on our cognitive resources *alone*. The mind sets the conditions of satisfaction and it is up to the world to meet these conditions. On the contrary, such a concept as gravitation is not some social construction which I can disregard as something subjective and walk out of a twenty–store window unharmed.[165]

If intentionality is the mind's being directed at something or about in the world, then it can be entirely expressed in language and, in fact, be reduced to the way language operates. It makes no difference whether the intentional object is existent or not; language does not make distinctions based on existence or on the lack of it. This kind of intentionality (call it 'linguistic intentionality') is distinct from the reductive account of intentionality by 'computer functionalism', in that the later treats intentionality as being in a state of computation and the basic rules of computation are *always* followed unconsciously. This account of intentionality makes no distinction between basic rules of thought on the neural levels of the brain and the higher cognitive functions of which we are aware and which are the only ones that do have intentional contents.

Linguistic intentionality accommodates both, the intentionality of perception and the intentionality of knowledge but explains them by a twofold reference to the syntax (rules of language which must be followed in order for an intentional state to have meaning to its owner) and to a person's relation to the states in the world toward which the intentional act is directed (language–games). Here 'language' means not just syntax, it also includes semantics and pragmatics, as well as the many uses of linguistic terms. Here, the notion of language goes beyond its usual definitions in linguistics.

The problem of distinguishing between the socially–imposed and the world–structured contents of intentional acts does not appear for linguistic intentionality. Intentional acts can be completely wrong about how the world really is because they are formed in language, by the language and depend not on the world but on the relationship between world and language, viz. *semantically*. My description of the world can be completely wrong (i.e. my disbelief in gravitation), and my interpretation of perceptual experience is totally misguided (i.e. seeing sunrise on the east and sunset on the west affirms my belief in geocentrism), but this is the problem of my social background network that, for some reason, is wrong. Language and intentionality has nothing to do with it. Or better still, my wrong belief–system cannot be fixed by language nor by perception alone, it needs, perhaps, correct education.

---

[164] Searle (1983: ch. 5).

[165] Searle (1983: ch. 8).

The veracity of my linguistic expressions is not checked out against my epistemological backgrounds or intentions but against my use of them in the world and in a given community. The truthfulness of my statements will not depend on social agreement (which can be misguided as was geocentrism for centuries) or on accepted definitions (which are constantly revised) but on the active use of language. The correctness of use will be determined by winning in a language–game. Likewise, the falsity of such use is determined by the loss. The truth-value of linguistic intentionality will depend on how grammar (rules) will be used in semantics (strategies) in our actions.

The major issue with Searle's philosophy of mind, however, does not come with his semantics or epistemology but with his ontology. It has been argued that biological naturalism is yet another version of the property dualism,[166] and Searle had argued against it. The problem is this. If consciousness is a higher–level property of the brain's lower–level functions, then it appears that biological naturalism is committed to a property dualism on the grounds of not drawing a clear identity between consciousness and the neural activity of the brain. Searle, indeed, resists reduction of the first–person conscious experience to the third–person neural activity.

Searle makes an important distinction which, according to him, puts him at odds with property dualism. He is advocating *causal reduction*: consciousness has no distinct causal power distinctly from the brain. But this causal reduction does not imply *ontological reduction*, viz. to the third–person neural activities: "The difference is that consciousness has a first–person ontology; that is, it only exists *as experienced* by some human or animal, and therefore, it cannot be reduced to something that has a third–person ontology, something that exists independently of experiences. It is as simple as that".[167]

Why is the neural activity of the brain distinct from consciousness if consciousness is causally dependent on the neural activity? Why is this activity a third–person ontology? He insists that consciousness is not something 'over and above' the neural processes of the brain and that, for him, is enough to be distinct from property dualism. But 'over and above' is not really a definition of property but rather substance dualism. Property dualism refutes an ontologically independent status of consciousness or of any other mental property, claiming that these are *properties* of the same ontology.

Searle introduces the term 'feature' to describe consciousness, a feature of the brain but recognizes himself that 'feature' is also in vocabulary of property dualists. He insists that property dualists believe that this feature is something additional to the brain, but this is, again, more in line with substance rather than property dualism. At best Searle has defended himself from substance dualism but the issue of regarding consciousness as some property of the brain even if it has no 'life on its own' remains.

The main problem of property dualism is not that it considers the mind as something extra to the brain, but that it offers a separate, irreducible explanation of the mind's nature and causality from that of the brain's nature and causality. This compromises significantly the theory's explanatory power. It also begs the question of the connection between the two and the question of how 'minds are caused by brains'. Searle insists that

[166] Manson (2003).

[167] Searle (2002: 60).

counting at all is wrong, but his language remains persistently dualistic throughout his writings. Thus, the ontological problem remains unsolved.

Biological naturalism started as a consequence of Searle's Chinese Room thought experiment which was meant to demonstrate the failure of a strong AI–thesis to properly account for the analogy between a computer and a human person. Instead, Searle's own experiment of biological naturalism resulted in suggesting even a broader metaphor, far from identifying mental and brain states. I suggested that biological naturalism itself comes as a result of Searle's earlier ordinary language philosophy with its fundamental belief that meaning can and is expressible in various linguistic forms. In philosophy of mind it often takes the form of internalism at the basis of which stands the presumption that 'meaning in the head' and 'meaning in language' *somehow* correspond to each other, or that the first is the cause of the second. Searle's dictum of "brains cause minds" must be read in the context of his internalism and direct realism. Epistemological and semantic questions are intrinsically bound to these methodological foundations.

The major accomplishment of biological naturalism is the acceptance of the metaphysics of the mind resulting in being able to accommodate the problem of qualitative experiences. On the ontological level, however, with metaphor as the only identity statement, biological naturalism remains problematic. Even those who embrace internalism and direct realism, see that biological naturalism fails to state the mind/body problem adequately.

In the following chapter, in the context of Wittgenstein's philosophy of mind and language, I will return to the problems of internalism and ordinary language philosophy. That presentation will come as an alternative treatment of meaning and mind. In terms of the metaphorical and analogical uses, for some, reduction of all mental states to the complexity of the brain states with the elimination of the former is the most plausible solution that accepts no analogies but the strongest kind of identity statement.

4. Eliminative Materialism

Eliminative materialism is a negative theory which is founded on the refutation of a common–sense understanding of mind. This common-sense understanding, sometimes called 'folk psychology', is wrong for a simple reason: it presupposes the existence of mental states which, in effect, do not exist. Eliminativists often refer to the Cartesian certainty at the beginning of the Second Meditation in respect to *res cogitans*;[168] eliminativists refuse that which for Descartes was the foundation of his epistemological and metaphysical system. There is, however, a common thesis that eliminativists share with substance dualists: our mental states cannot be reduced to neuronal states of the brain. For eliminativists the reason why reduction is impossible is, again, because that which is supposed to be reduced does not exist *simpliciter*: mental states are non-existent and any discourse about them are mere common-sense postulates.

William Sellars' semantics is one of the major inspirations for eliminativism. Sellars speaks of a linguistic and theoretical framework in which we predicate about someone's

[168] Ramsey (2011).

or our own behavior. Mental states have names and these names constitute a certain theory. Thus, we have 'theory–theory': generalizations and conjectures about everyday experience in our language.[169] This theory–theory makes common–sense assumptions about causal properties (beliefs and desires are caused by some experiences) and the intentionality of our beliefs and desires.

These two assumptions are criticized as projections of identity from our common–sense conceptions of our own experiences of objective reality. The strength of folk psychology is in subjective experience. One's experience of pain or pleasure is real, but our beliefs, the self–consciousness of one's experience, and our desire for pleasure and not pain is theory–laden. Paul Churchland speaks of introspection being determined by common–sense theories. The solution would be to put folk psychology to the test of empirical investigations by neuroscience and translating (but thus eliminating)[170] common–sense preconception into an empirical description of brain states.[171]

Patricia Churchland insists on the need for testing the folk psychology terms against empirical evidence especially in the case of intentionality. Which brain states can point out one's mental directness or aboutness of a desire or a belief? But even the basic sentence structures which are fundamental for any meaning are without any empirical evidence. At no level of scientific analysis does science share *any* properties we predicate of our beliefs.[172]

This point is supported by some connectionist models of the brain. According to these models, information is highly distributed so much so that no causally distinct data can indicate anything resembling the way our language works. Although language is often the central topic of investigation by the connectionists, no other cognitive features have been matched to the analysis of the neuronal networks.[173]

Dennett evaluated the connectionist results by saying that our philosophical concepts are so inadequate that they fail to pick out anything real. Our concepts of pain would never correspond to anything in computational systems which are much more complex than our talk of qualia and intentionality.[174]

There has been much criticism of eliminative materialism. The criticism usually falls into two camps: those that defend folk–psychology,[175] and those that accept the possibility that folk–psychology may not be right, but, even so, say this possibility would not be enough reason to eliminate mental phenomena.[176] Against the connectionist analysis eliminating any mental properties it has been argued even by the supporters of eliminativism that the connectionist view is too demanding and that 'stupid' neurons and their

---

[169] Sellars (1956).

[170] 'Translation' such as Latin prefix '*trans*–' "across, to or on the farther side of, beyond, over", and '*locus*' "place" (OED). Thus, 'translation by elimination' is 'throwing away'.

[171] Churchland (1988).

[172] Churchland (1986).

[173] Ramsey, Stich and Garon (1990).

[174] Dennett (1978); Dennett (1988).

[175] Fodor (1987).

[176] Rosenthal (1995).

slightly more 'intelligent' networks would not be able to demonstrate intelligence and intentionality at such basic levels.[177]

These and many other objections and replies have a very impressive literature, but arguments *pro* and *con* eliminativism appear to be neither conclusive nor targeting the core issue of eliminative materialism, which is its interpretation of identity. Eliminative materialism is not an identity theory, unless tautology is interpreted as such, but eliminativism interprets folk–psychology as an identity theory. The main point of eliminativism is its refutation of folk–psychology on the basis that it fails to identify its ontology and its own theory. Stephen Stich objects to the eliminativist assumption about folk psychology because it is not able to account for the mental states due to its ignorance of the brain processes. For Stich it is not clear when the theory in question should be dismissed as scientifically irrelevant based on its incomplete account of reality.[178]

But there are deeper problems regarding eliminativist ontology and its notion of identity. Any identity, except tautology, would imply some kind of folk–psychology. If one accepts Quine's thesis of the indeterminacy of radical translation, then there are two possibilities for treating identity: either all our discourse on mind–body becomes just a metaphor without much further understanding of how far we reach the truth, or it becomes tautology in relation to the brain states. Any true identity theory based on dualistic language would necessarily fall into the irrelevant realm of folk–psychology.

Eliminativist discourse accepts propositional dualism by which one refers to the mind and body without accepting the ontological duality of substances or properties. While stating that folk psychology is false, eliminativism not only never suggested any alternative language but continues to implement folk psychology in its own narrative. This contradiction becomes even more acute when we ask the epistemological question on the conditions of knowledge about our own mental states. Any attempt to define it outside of some folk psychology's language is impossible.

Eliminativism offers an important criticism of ordinary language by pointing out that it is not enough to account for mental phenomena, but the theory goes too far by indicating this inability as the reason of the non existence of the mind. Right ontology demands right semantics; when right semantics, for whatever historical reasons, are lacking, the elimination of ontology and metaphysics of the mind and the reduction of identity to tautology about physical states are unwarranted.

One of the inspirations for semantic proposals among eliminativists (at least in the case of Dennett) is Wittgenstein's criticism of the ordinary language and his dismissal of any private language for any private sensation. In the next chapter I shall examine this thesis in detail. However, while Wittgenstein indeed mistrusted the ordinary language ability to express concisely private sensations he was far from dismissing or eliminating such sensations as being no more than manifestations of the specific physical states. Nor did he leave the semantic problem unresolved. The use of the language–games notion offered a complex solution and an alternative to any ordinary or private language expressions. What is most attractive in this solution is the dismissal of any kind of dualism in his philosophy of mind.

---

[177] Dennett (1991).

[178] Stich (1996).

5. Emergentism and Supervenience

5.1 British Emergentists

The term 'emergence' was originally used in chemistry and biology to denote the result of a combination of two or more substances or properties that give rise to a new substance or property irreducible yet novel to them. Thus, the combination of two molecules of hydrogen and one molecule of oxygen produces water which is no different chemically, that is, in its molecular composition, to the gases but different in properties and, thus, irreducible to its compositional elements. In the psychology and philosophy of science, consciousness is considered to be an emergent property of the brain, in that it is intrinsically dependent on the brain states yet cannot be explained solely in neurobiological terms.

Another example is Richard L. Gregory's analogy between the emergence of ant nests from individual ants' behavior and the emergence of consciousness from individual brain nerve connections. The claim is that a study of individual brain nerves and their connections will not explain the dynamics of human consciousness any more than a study of individual ants' behavior would explain the architecture of the ant nests, even if consciousness is impossible without the brain cells, as nests are impossible without the labor of ants.[179]

Emergentism first appeared in the late–19th, early–20th century Britain and developed largely as a response to the mechanist/vitalist debate. Mechanists claimed that scientific laws in the experimental sciences could be reduced to the lower level scientific laws of physics and chemistry. By the late–19th century amid the general conviction that *all* laws of physics and chemistry were already discovered and known, the major puzzle was the explanation of life. Thus, mechanists believed that all life 'mysteries' could and should be reduced to the physical and chemical structures of living matter. Vitalists contested reductionism and postulated some primary substance as a guiding principle in all living beings. The Emergentists' response was a rejection of both mechanistic reductionism and vital force.

The first systematic account of emergentism was given by John Stuart Mill in his 1843 *A System of Logic*. Mill introduces an important distinction between 'homopathic laws', single laws that explain single events, and 'heteropathic laws', laws not reducible into simple laws. The later are derived from the separate laws and not from the laws of the constituent elements.[180] These 'heteropathic laws' are empirical in that they are formulated solely on the basis of observations of the facts and not *a priori*. In his example, heat, light and magnesium could be considered as a manifestation of a single force, but not as one convertible into another.[181]

---

[179] Gregory (1987).

[180] Mill (1974: III, 10 §4).

[181] The above example was removed by the posthumous editions but reported by Leslie (2004: 116).

These heteropathic laws were considered as emergent. But both heteropathic and homopathic laws were considered to be causal laws. They both explain the causes of the effects in question.

In emergentism nature is presented as layered according to the complexity of its constituents. Physics investigates the most fundamental and simplest constituents, followed by the more specialized sciences which investigate more specific manifestations of nature. Hence, physics is followed by chemistry, then biology and psychology up to sociology.

To this hierarchy in science, which mirrors nature, corresponds each set of laws in their causal explanations. Emergence occurs when the causal interactions in the higher levels are additions to the interactions of the lower, most basic levels.

This account of the world and the emergent mental properties goes in hand with Mill's direct reference theory of meaning, empirical and psychological in nature as was his logic. The meaning of a name is its referent: the name is true if the referent has the attributes which the name expresses. In a way, the name is an emergent property of its referent if it is veridical.

Logic's theoretical foundations, according to Mill, "are wholly borrowed from Psychology, and include as much of that science as is required to justify the rules of the art [of logic]".[182] Thus, the 'laws of thought' are likewise emergent properties of our psychological states as well as of our basic experience.[183]

Both Mill's direct theory of meaning and his logic came under criticism in Frege,[184] Boole,[185] and Peirce,[186] although it did have resonance in Russell's theory of denotation,[187] early–Wittgenstein's 'ostensive definitions',[188] and Kripke's revised notions of reference and necessity.[189]

Following Mill's notion of emergency and his multi–layered vision of the world and sciences, C. D. Broad postulated non–causal interactions between layers. He introduced 'trans–ordinal laws', synonymous to 'emergent laws', which are emergent from the most basic laws but not reducible to them. This method was meant to answer Broad's fundamental questions: "Are the apparently different kinds of material objects irreducibly different",[190] which translates into: "Is chemical behaviour ultimately different from dynamical behaviour?"[191]

The trans–ordinal laws are descriptions of the appearance of a single emergent property of the lower structure in a higher structure. In this way, emergence is contrasted

---

[182] Mill (1974: 359).
[183] Mill (1974: 381).
[184] Frege (1968: §7); Frege (1967: 143–162).
[185] Boole (1952).
[186] Peirce (1931: 18, 92).
[187] Russell (1905: 479–493).
[188] Wittgenstein (2001: 3.263).
[189] Kripke (1980).
[190] Broad (1925: 43).
[191] Broad (1925: 44).

with the reduction of laws as well as with the metaphysical level of discourse on the world.[192]

The major influence on the emergentist philosophy of mind were the works of Samuel Alexander, for whom mental properties were something novel altogether in respect to the neural activity. In this way, the mental is truly emergent instead of resulting from the neural processes.[193]

Emergentism falls under non–reductive physicalism, since mind is interpreted as an emergent property of the neural structure. As O'Connor and Wong emphasize, emergentism is a property–dualism on the basis of its denial of type or token identity between the emergent property and the brain.[194] This is particularly apparent in Alexander's case: emergent qualities are novel while its physical base remains the same. This version of physicalism avoids epiphenomenalism because it accepts a causal relation between the mental and the physical. Mental states cause each other and by that mutual connection they cause the following physical state.[195]

In one aspect all adherents of emergentism agree: even full knowledge of the physical state would not be sufficient for knowing what kind of mental state was its cause. But that makes any talk of causation ('upward' and 'downward' alike) very difficult. If both type and token identity is denied, and if reduction is avoided as well, an explanation of the mental causation does not seem to be plausible.

5.2 Supervenience Thesis

The supervenience thesis comes as a historical and thematic successor of emergentism.[196] Since emergentism, in relation to the mind/body relation, claimed that mental properties emerge from its physical base, supervenience thesis says that if properties of one object supervene upon properties of the second object, any difference in properties of the first object would necessarily imply those differences in the properties of the second object. Let the first object be named '*A*' and the second '*B*'. Then, in a famous dictum, supervenience is defined as 'there cannot be an *A*–difference without a *B*–difference'.

Similarly to emergentism, supervenience rejects reductivist physicalism. Mental properties which supervene upon physical ones are considered to be intrinsically dependent but not reduced to its subvenient base. As Davidson insists, such dependence does not imply reduction to some law or definition.[197]

Under the classification of McLaughlin and Bennett[198] supervenience comes with a certain modal force and extends individually and globally, viz. with different force on particular individuals and to possible worlds. Individual supervenience, then, implies

[192] Broad (1925: 76–78).
[193] Alexander (1920: 14).
[194] O'Connor and Wong (2009).
[195] Alexander (1920: 13).
[196] Kim (1999).
[197] Davidson (2001: 214).
[198] McLaughlin and K. Bennett (2011).

the above made statement that two individuals cannot differ in one property without differing in another as well. In the case of global supervenience it means that two possible worlds cannot differ in one set of properties without differing in another set of properties.

In the works of Kim, individual supervenience is distinguished as weak and strong. Concerning the weak version, the supervenience thesis implies that 'If individuals x and y are indiscernible in P–properties, then necessarily (in all possible worlds), x and y are indiscernible in M–properties'.[199]

In this account, because some properties are present in one world but absent in another, supervenience applies only within one possible world (property supervenes not necessarily). Mental and physical properties are said to depend on each other in one place but not in every. As Kim acknowledges, weak supervenience is insufficient for physicalism which claims that mental properties are determined by the physical properties. If physical properties are determined in a particular way, then the mental properties cannot be distributed in any other way.[200]

To accommodate the physicalist claim, Kim proposed global supervenience, which repeats the physicalist condition that 'When in two worlds physical properties determine two individuals in the same way over their P–properties, they cannot differ how the M–properties are determined'. In this version supervenience is applied to the whole world and not just to individuals. Here, the relations of dependence and determination are settled by the physical conditions of the world rather than relations between individuals as it was in the case of weak supervenience.

In order to accommodate such inter–personal relations in a non–weak supervenience way, Kim introduced a strong (individual) supervenience, according to which 'In any world w_1 and w_2, for every individual x and y, if x in w_1 is indiscernible in P–properties with y in w_2, then x in w_1 is indiscernible in M–properties from y in world w_2'.

Unlike in weak supervenience, strong supervenience applies to the selected individuals from any possible world, and unlike in global supervenience, strong supervenience applies not to the whole worlds but to individuals. Kim points out that strong supervenience entails global and weak supervenience while weak supervenience entails neither.[201] However, Paull and Sider insist that when strong and global superveniences are limited to the intrinsic properties of individuals, they are equivalent.[202]

What the supervenience thesis must explain is how mental properties are dependent without being reduced to their physical properties. The 'type–physicalism' implies reduction (sometimes called 'analytical reduction'). Supervenience, since around the 1970s, seemed to be promising in defining such a non–reductive physicalism. However, supervenience would have to offer a different account of the dependence between mental and physical properties as two identical types (so, type–type reductionism) but without the elimination of the mental properties.

---

[199] Kim (1987).

[200] Kim (1995: 578).

[201] Kim (1987).

[202] Paull and Sider (1992).

Such a solution was proposed by Kim when he speaks of a nomological reduction (as opposed to an 'analytical' one), according to which for each mental property there is a coextensive physical property. He suggested that strong supervenience could carry such a task in providing this important relationship between the mental and physical properties. It says that 'Each mental property *M* has infinitely many physical properties *P* such that if some particular mental state *x* has a particular mental property *M*, that state *x* also, *necessarily*, has physical property *P*'.

In other words, Kim here recalls the multiple realizability thesis, according to which a single mental state can be realized by many physical states. However, neither the above modified strong supervenience thesis nor the multiple realizability thesis per se explain *how* one is realized in the other. A nomological reductivism would have to specify how something law–like in one domain is necessarily law–like in another.

There are two major objections to supervenience: one from internalism and the other from externalism. As was explained before, externalism claims that the relationship between the object in the world and the word for which that object stands is specified by the external causal relations in the world. Thus, someone's belief in a Twin Earth where a substance XYZ is similar in all effects to the H_2O on the earth, and my belief that water, being H_2O, is liquid, does not supervene on our physically identical structures.

Contrary to externalism, internalism claims that 'meanings are in the head', or that the relation between the external object and the name for which it stands is mind–dependent. The internalist interpretation imposes the problem of qualia, qualitative states of mind, such as subjective pains, pleasures, color perceptions, etc. In a case of qualia inversion (as we had seen in the case of Locke's example of the color spectrum inversion), my sensing of the red color of an apple is the same sensation as somebody's with an inverted qualia perception of a yellow plum. Strong and global superveniences cannot respond to the externalist objection, just as weak supervenience cannot answer the internalist one.

The supervenience thesis offers a different dependence relation between mental and physical states, but it offers no casual explanation for the relation. Perhaps the reason why it does not do so comes from the fact that philosophers of supervenience never constructed an ontology, – a general description of the world or of the physical upon which the mind would depend. Although emergentism had an ontology it remained unconnected with semantics and the nature of the mind in general. Thus, emergentism, as was specified already, remained intrinsically dualistic.

The issue of the lack of causal account within the supervenience thesis makes it vulnerable to both internalism and externalism. As a result, the most that the supervenience thesis can offer is to be descriptive of the mind/body relationship, thus remaining on a very general level of discourse.

5.3 Anomalous Monism as Supervenience Thesis

One can be a physicalist monist, presuppose that there is only physical realm, yet at the same time hold that there is no reduction of the mental to the physical. Minds do not exist, but mental properties do. They constantly change and this change constitutes mental events.

Donald Davidson expressed this view of non–reductive physicalism in his argument called 'anomalous monism' and proved it in this three–step argument:

(1) All mental events are causally connected to physical events: agents' beliefs and desires cause them to act and instantiate changes in the world.

(2) Because of the causality between mental and physical events, there must be strict psychophysical laws that explain the interactions.

(3) The nature of mental events is intrinsically different from physical events, which rules out any possibility of stating any laws in physicalistic terms. Therefore, there are no strict psychophysical laws, laws that connect mental events with physical ones under physical descriptions of the causal relations.

Davidson remarks, in connection with the (3), that only physical description can be given to any law but mental events, being different from physical ones, fall only under a mental description, not a physical one. The (3) also excludes two kinds of reductionism: the first is (analytical) of the physical description of the mental events and the second (nomological) of reducing mental and physical to strict bridging laws connecting mental and physical events. However, since the argument as a whole implies that mental events do not constitute any additional reality to the physical one except in our different discourses of 'mental' vs. 'physical', Davidson claims that his 'anomalous monism' represents a sort of an ontological reduction together with a propositional dualism.

The kind of identity claimed by 'anomalous monism' is a token and not a type identity in order to rule out property or substance dualism. The reason for token identity is *a priori*: because there cannot be any strict laws, mental properties must causally interact with physical properties and be token–identical to them without any empirical evidence. A strict law would imply a type identity of some sort but that would contradict the theses of 'anomalous monism'.[203]

Since (3) stands in a seeming contradiction to (2), this version of monism is called 'anomalous'. What makes this argument logically consistent and not contradictory is Davidson's insistence that mental events are non–ontologically different from the physical ones. To insist on this point he presents his theory of meaning that rejects both the Fregean sense/reference theory of meaning and internalism/externalism views. Davidson embraces the Wittgensteinian notion of meaning–as–use as an alternative to the above semantic theories.

In Davidson's account, one's mental attitude toward some physical fact or object does not presuppose the existence of that physical fact or object. (Frege's objectivism requires that one's sense stands in a direct relation to its reference in a real or in some conceptual world of, i.e., arithmetic.) One's calculation of the weight of an object refers as much to the number (i.e. 500 grams) as to the object of that weight, but that object very well can be as abstract as the number indicating its weight. My mere thinking of a half–liter box of milk does not imply it actually exists. What makes my reference to some particular object meaningful is a social agreement settled in the rules of grammar of a particular language and its common use. These rules are common but my use of them is private, just as the existing half–liter box of milk is objective, while my measuring it (although obeying all agreements of measurements) is private.

---

[203] Davidson (2001).

On these grounds Davidson, similarly to Kant, draws a radical difference between the concepts that describe mental events and the concepts that describe physical ones. Linguistic descriptions are private in use but common in following the rules of language, its syntax. He refuses, therefore, the externalist and internalist theory of meaning, since meaning neither resides in the minds nor is it defined by the external causal relations in the world. Instead it depends on the community of the speakers, though at the same time those common concepts are intrinsically the personal, mental property of each of us. Learning presupposes a learner, a teacher and the world from which the learned concepts are drawn. All three are necessary for a personal coherency and shared common understanding: "to communicate is to recognize the existence of other people in a common world".[204]

In sum, 'anomalous monism' postulates a fundamental difference between mental and physical events on the basis that the former is caused by the later and in principle there must be a strict law to account for this causality. Because of that difference and because all discourse on physical events is done in mental description, such laws cannot be formulated.

There are three major objections to 'anomalous monism'. The first is the objection to token identity. The second is in relation to the supervenience thesis and the third is the charge of epiphenomenalism.

In relation to Davidson's *a priori* established token identity between mental and physical properties, it is claimed that unlike mental events, physical ones are defined in space and time with great precision and complexity. If there is any empirical evidence of token identity it must be of the type identity between mental and physical properties, but that contradicts Davidson's theses.[205]

This criticism, in a way, confirms the anomalous characteristic of Davidson's monism. Any token–identity law would have to be defined in a strictly physicalistic language and the only science that so far can be considered as a candidate for expressing such a law is neuroscience, which has not reached the micro–complexity required. In principle, Davidson accepts such a possibility to formulate the token–identity psychophysical law in some unclear future.[206]

In relation to supervenience, Davidson clearly implies that there is a dependency between mental and physical properties. Regardless whether these dependencies are applied cross–worlds or only within one possible world, it must entail the existence of strict laws which would predict mental events, but that is contrary to 'anomalous monism'.

Davidson stated that if two events are alike in all physical respects they must be alike also in some mental respects.[207] Later he modified that claim by saying that if two events do not share a mental property they will likewise not share at least one physical property

---

[204] Davidson (1995a: 234).

[205] Leder (1985).

[206] Davidson (1999a).

[207] Davidson (2001: 214).

either.[208] This modification would employ multiple realizability, since one mental event can be brought by different physical events.[209]

This line of defense is very similar to that of Kim who likewise refers to the multiple realizability thesis and to the complexity of physical properties to account for mental properties. The difference with Kim here is in Davidson's denial of a possibility of formulating any psychophysical law on the grounds of the complexity of the physical properties. An interesting defense of 'anomalous monism' comes from Cynthia and Graham Macdonald who assert the compatibility between supervenience laws and anomalous monism's denial of such, if the possibility of formulating these laws remains only a theoretical possibility, never actually realizable.[210] Yalowitz believes that this would make anomalous monism "a contingent epistemological position" instead of "the necessary metaphysical doctrine" it claims to be.[211]

The third objection to anomalous monism inquires about the status of the explanatory role of mental properties when monism claims that all physical events are explained exclusively in physicalist language. That would imply that all mental events, in fact, are physical events and that the mental descriptions of them cannot be put in the physicalistic terms of a strict psychophysical law. What is then the causal relevance of mental properties? Anomalous monism is charged then with epiphenomenalism.[212]

Davidson thinks that such an objection confuses the metaphysical notion of causation with the epistemological notion of the explanation. Anomalous monism is a metaphysical theory that states the relation between mental and physical properties in terms of causation; it is contrary to its thesis (3) to provide an explanation.

Davidson appeals to the casual efficacy between different mental events and that in turn *must* have casual efficacy on its subvenient physical parts. How this causality is to be *explained* cannot be demanded of a metaphysical theory, especially when such a theory is based on a denial of an explanation.[213]

If the third objection claims that anomalous monism cannot account for *any* causality, it is false according to Davidson's ample theory of semantics. If, on the other hand, the claim is that anomalous monism does not provide a casual explanation of mental events and of its causal role in physical events, the answer would be that it was never supposed to. It is the peculiar nature of anomalous monism to hold an explanation of this kind.

Without Davidson's semantic theory, anomalous monism would have little consistency let alone theoretical relevance for the philosophy of mind. This fact should prove the importance of the semantic question and its intrinsic relevance to epistemological and ontological questions. Anomalous monism does not answer the epistemological question but, in a way, substitutes a semantic one for it. It denies the possibility of an answer on the nature of the mind, given the incomplete status of physics. Its ontological

---

[208] Davidson (1995b: 266).

[209] Davidson (1993: 7).

[210] Macdonald and Macdonald (1986).

[211] Yalowitz (1997).

[212] Honderich (1982); Kim (1993: 20–22).

[213] Davidson (1987a: 444).

reduction must be read together with Davidson's theory of semantics. Meaning takes care of knowledge; semantics substitutes epistemology by encompassing it.

However, even though 'anomalous monism' resists being from the outset any explanatory theory of the mind/body relation, the insistence that the explanation can only come from neuroscience (or biology, physics in general) makes philosophy of science purely a descriptive discipline with little if any innovative theoretical force. The purpose of this present work is to show that philosophy can do better.

Conclusion

To conclude this chapter and this first part, I shall indicate two topics that are essential to the exposition so far provided. First is the dualistic nature of the language of the philosophy of mind, and second is the notion of identity.

As both chapters have demonstrated, the use of the word 'mind', or 'mental' presupposes a dualistic conception of the subject matter of the philosophy of mind. The so–called 'mind–body problem' cannot be solved if such dualistic language is used. This problem was formulated within substance dualism and, as the expositions of this first part have shown, it cannot be solved by property dualism either. The reason is simple: the solution would imply unity between the two, but the formulation of the problem implies separation. The separation is twofold: conceptual and linguistic. The conceptual separation is defined in property dualistic terms, and the linguistic separation is defined in terms of propositional dualism. Within these dualistic limits the problem has no solution.

As to the second problem, my earlier suggestion was to regard physicalistic monism in terms of different identities: from tautology to metaphor, and to translate contemporary theories in the philosophy of mind in terms of different senses of identity. The purpose of this method is to show that when physicalistic monism is defined in dualistic terms the notion of identity – defined in any way – is inconsistent. I also mentioned that *univocal identity* was defined in two ways: as a mathematical identity between *Sinn* and *Bedeutung* by Frege, and as a linguistic identity between types and tokens by Peirce.

Frege's notion of identity works elegantly if it is applied in arithmetic between number and numeral, but when it is applied to logic or semantics we will have the inconsistencies of the Russell Paradoxes, upon which Frege's logicist project stumbled. If this notion of identity did not work in logic, why was it applied in the philosophy of mind, since mind has much less chance to be reduced to a one–to–one relationship with the brain upon this identity?

Peirce's type–token identity in the case of the mind's relation to the brain it either implies the reduction of the elimination that excludes many features of the mind unexplained, or a very vague metaphor that can lead to less radical versions of dualism but still compromises on the explanatory powers of a given theory.

The major problem with the type–token theory is to define what is a type. Linda Wetzel proposed three theories of types[214]: type as a universal (or a kind), as a set, and as a law. However, none of these definitions fit mental states or physical states as types.

First, types of physical or mental states cannot be a universal, if by universal we mean abstract, non–caused entities outside of time and space. A very strong substance dualist theory might accept mind as such type. If by kind we also mean an abstract notion, then it falls under the same description as universal.

Second, types of physical or mental states cannot be sets, since sets are the class of its members, by Quine's definition,[215] just as a species is the class of its members. Sets and species do not exist outside of the class of its members. So, a set of all natural numbers without the number '5' is a different set than the one containing number '5'. If physical states are types–as–sets, then it would imply that they cannot exist without their tokens, the mental states, or at least that the lack of the later would imply a somewhat different type of the former. If, on the other hand, we consider mental states as types–as–sets, then the reverse multiple realizability thesis would not be applicable in their functions.

Third, types of physical or mental states cannot be laws. This last understanding of types as laws comes historically first, due to Peirce's theory of semantics. For Peirce, types are general laws, created by people, whose identity is to be an interpreted sign. Thus, they do not enjoy any existence outside of a purely linguistic use. They are not individual things, nor features, nor properties.[216]

Finally, the relationship between a type and its tokens must be defined by what type is. Generally, there are two ways to define such a relation: idealism or realism. Idealists would say that types do not have any real (spacio–temporal) existence, while its tokens do. A Platonist idealist would insist that types are abstract entities and tokens are their copies. A nominalist would consider a type to exist only in language. Realists accepts type's existence but considers it to be essentially linked to the existence of its tokens.

The recent history of the philosophy of mind attempted implementing different configurations in considering body and mind as types and tokens. Unless there are deterministic psycho-physical laws of identity, neither a type-type,[217] nor a type-token, nor a token-token brain states/mental states description of relations will ever be univocal. And because such laws cannot be defined in principle (due to the lack of empirical evidence of any kind for the mental states), any 'type' and 'token' nomenclature will necessarily fall either into tautology or into ambiguous language of equivocation, analogy, or metaphor. In a way this is evident: 'types' and 'tokens' are mathematical statements and as such indicate mathematical relations. Therefore this relation must be spelled out in a law-like statement, which, in the context of mental states, cannot be formulated.

The situation of identity defined in Fregean terms of sense and reference/meaning is more helpless than with types and tokens. It has already been mentioned that the original distinction was meant to define the identity between numbers (*Anzahl*) and numerals (*Zahl*) against the Kantian intuitionist and Millean psychologist versions of the number

---

[214] Wetzel (2008).

[215] Quine (1987: 216–219).

[216] Peirce (1932: 243–265).

[217] For in–depth discussion see Jackson, Pargetteer and Prior (1981: 209–225).

theory. However, the same notion of identity applied to language and even logic results in semantic as much as in logical paradoxes which made Frege's logicist project inconsistent.

In the case of Frege's distinction, the porridge is too hot and in case of Peirce's distinction, it is too cold. 'Sense' and 'reference/meaning' can only work in number theory, where the relation between numbers and their numerals is immediate. In the case of language it is not, unless we eliminate the meaning. Even Frege's celebrated example does not avoid this problem. Behind the four names for the same planet 'Venus' there is a rich history of cultural and linguistic uses in Ancient Greece, astronomy, etc. Therefore, it cannot be immediate at all.

As for Peirce's distinction, it was only meant for language; it cannot be extended for the brain–mind relation.[218] Whatever (Peircean or not) definition of 'type' we might accept, and whether it is applied to physical or mental states, the above explanation has shown that it cannot be applied in the context of the philosophy of mind.

Despite this grim presentation of identity theories in the philosophy of mind and its intrinsically dualistic language, I believe that the solution is not to abandon the notion of identity altogether, but to redefine it in different terms. Chapter V will offer such a reformulation. The mind–body relation of identity requires a dynamism which neither sense/reference nor type/token can offer.[219]

Finally, the anti–dualistic language commitment must be radicalized. This implies that the 'mind–body' terminology should be substituted by something else. The following three chapters will provide a new language (semantics and epistemology) that could be applied to a new philosophical vision of how the problems are to be met and resolved.

---

[218] C. S. Peirce's own explanation of the distinction between types and tokens is notoriously short and auxiliary to his semeiotics. It must be read only within his explanations of the classification of his theory of signs: "A common mode of estimating the amount of matter in a MS. or printed book is to count the number of words. There will ordinarily be about twenty the's on a page, and of course they count as twenty words. In another sense of the word 'word,' however, there is but one word 'the' in the English language; and it is impossible that this word should lie visibly on a page or be heard in any voice, for the reason that it is not a Single thing or Single event. It does not exist; it only determines things that do exist. Such a definitely significant Form, I propose to term a Type. A Single event which happens once and whose identity is limited to that one happening or a Single object or thing which is in some single place at any one instant of time, such event or thing being significant only as occurring just when and where it does, such as this or that word on a single line of a single page of a single copy of a book, I will venture to call a Token. An indefinite significant character such as a tone of voice can neither be called a Type nor a Token. I propose to call such a Sign a Tone; In order that a Type may be used, it has to be embodied in a Token which shall be a sign of the Type, and thereby of the object the Type signifies. I propose to call such a Token of a Type an Instance of the Type. Thus, there may be twenty Instances of the Type 'the' on a page". "Prolegomena to an Apology or Pragmaticism", 1906 in Peirce (1933: 537). A few pages later Peirce makes clear this relation between types and tokens to his theory of signs: "All general, or definable, Words, whether in the sense of Types or of Tokens, are certainly Symbols. That is to say, they denote the objects that they do by virtue only of there being a habit that associates their signification with them". Peirce (1933: 544).

[219] The type–token identities used in philosophy of mind were already criticized by Saul Kripke in Kripke (1980: 144–155).

The methods of this work are Wittgensteinian and as such are set to be a linguistic analysis. The problems which have been presented so far come as a result of "troubles arising from a particular use of language",[220] whether they are about identity or about the use of its basic terms. Radical problems, viz. problems that define the methods of an inquiry, demand radical solutions. One of Wittgenstein's students, Alice Ambrose, sees the link between philosophical solutions and new language of philosophy this way:

> To find our way among these questions, confusions engendered by our use of language have to be got round, the aim being "*complete* clarity" which "simply means that the philosophical problem should *completely* disappear" *[Philosophical Investigations §133]*. When one considers the many centuries of controversy over every philosophical thesis, without resolution, and seemingly without hope of resolution, it would seem that the failure to come to truth indicates that a different approach to the disputes is necessary.[221]

Ambrose's interpretative statement of Wittgenstein's philosophy will be my guide for the rest of the work to follow.

[220] Wittgenstein (1979a: 13).

[221] Ambrose (1993: 136). See with another statement of Wittgenstein: "For philosophical problems arise when language goes on holiday". („Denn die philosophischen Probleme entstehen, wenn die Sprache *feiert*".) Wittgenstein (2001: §38).

PART II

LANGUAGE–GAME FOUNDATIONS FOR THE PHILOSOPHY OF COGNITION

CHAPTER III

Language–Games and the Semantics of Cognition

Introduction

In presenting the next two philosophers' thoughts on mind and body, viz. Wittgenstein and Peirce, their contributions to the philosophy of mind, the significant difference between them and all other philosophers presented so far must be stressed. Until now the division was between monism and dualism with further distinctions between physicalism and idealism within monism, and the three types (propositional, property and substance) within dualism. These distinctions are wholly uncontroversial within the historiography of the philosophy of mind.

Wittgenstein and Peirce belong to neither of the above–sketched categories. They both boldly rejected monism and dualism. They did not explain the higher cognitive functions by references exclusively to either the brain or mind. Both of them, although independently from each other (we have no indication that Wittgenstein had ever read anything of Peirce with the exception of one marginal remark on Peirce's treatment of induction in his conversation with Rush Rhees in 1943[1]) have ridiculed the notion of mind being located either in some part of the brain or in the brain at all. Likewise, the Cartesian notion of *res cogitans* was totally extraneous to both.

For Peirce cognition is the progressive growth of signs. For Wittgenstein cognition is the interplay of different but related language–games. These two limited definitions of cognition are much closer to each other than they may seem at first.

As J. van Heijenoort and Jaakko Hintikka remarked, Peirce and Wittgenstein belonged

[1] Rhees (2002: 13).

to two opposite traditions of logic and philosophy.[2] Peirce was a firm believer in language as a modal calculus, open to interpretations by that same language. For him logic was semeiotic in nature, its purpose was to give many potential interpretations of the meaning. Wittgenstein, on the contrary, belonged to the tradition of language as universal medium and the ineffability of semantics. Logic was seen as a special language, radically separated from the ordinary language, whose functions were to aid science and whose propositions did not reflect reality any more than tautology reflects truth. For the author of the TLP (just as for the logicists such as Frege, Russell and the members of the Vienna Circle), truth depends on how facts are in the world; logical propositions, based on mathematical ones, do not tell anything about these facts but only restate in their phenomenological languages what the ordinary language already describes. This is why all logical and mathematical propositions for Wittgenstein are tautological.

Yet, in their studies of human cognition these two philosophers meet at the common point of treating it essentially as multiple sets of signs operating by certain rules of grammar. The next two chapters will expand significantly on these claims.

In Chapter III I will present in brief Wittgenstein's early philosophical and logical developments and the reasons why he changed his mind about the most fundamental philosophical topics. This will lead to the explanation of his central idea in philosophy, viz. meaning as use and use as language–games. This idea will serve as a tool for solving the ontological, metaphysical and epistemological questions asked in the beginning of this work. The central part of this chapter will focus on the so–called 'Private Language Argument' that is often considered as his major contribution to the philosophy of mind. I will then present Wittgenstein's critique of identity in Frege, Russell and Ramsey and will indicate the consequences of this critique for the problem of the mind/body identity.

Chapter IV will begin by examining Peirce's semeiotics and logic as foundations for his remarks on the mind/body identity. The remaining parts of the chapter will concentrate on Peirce's idea of cognition as a continuum in relation to other continuum theories in philosophy and will end by providing a study of language–games and the semiotic notion of cognition.

---

[2]First to make a distinction between a *lingua characteristica universalis* as a universal language of thought, and a *calculus ratiocinator*, as a calculus of reasoning, was Leibniz in his "Zur allgemeinen Charakteristik" (1904: 30–8). Philip E. Jourdain was the first to associate the *calculus ratiocinator* with the traditional works of Boole, de Morgan, Jevons, Venn, C. S. Peirce, Schröder and Ladd–Franklin, and the *lingua characteristica universalis* with Frege, Peano, and Russell. See his "Preface to the English translation" of Louis Couturat's *L'Albegrè de la logique* (1914: viii). Jean van Heijenoort postulated to extend Leibniz's distinction into two different views on how language relates to the world. The first considers language as a universal medium in which interpretation is already fixed, while the second is a model–theoretic view that insists on a variety of many possible interpretations of our language depending on the users' agreements and diversity of possible worlds (1967). However, Frege himself associated his own *Begriffsschrift* with both traditions (1967: 220–233). Finally, J. Hintikka extends the list to include within the 'universalist' tradition Wittgenstein, the Vienna Circle, Quine and Heidegger, and within the tradition of 'language as calculus', Löwenheim, Tarski, Gödel and later Carnap in Hintikka (1988), Hintikka (1997: 142), and Hintikka (1980). For an in–depth analysis of Peirce's role within the language as calculus tradition see Hilpinen (2004).

1. From Meaning as Picture to Meaning as Use

TLP (1921) declares that the world, as a totality of facts, imposes its ontological structure on the nature of our language. The facts of the world are described by a proposition, and the function of language is to describe reality as this totality of facts. Words of our language refer ostensibly to these facts in the world and propositions are their descriptions.

Objects of the world are elements of facts. These objects, thus, are phenomenological in nature.[3] The meaning of a word is the object for which it stands. This relationship is isomorphic; it is defined by the picturing of the facts in propositions. This picturing of the facts is thought. This relationship between words and objects is ostensive, immediate. Since objects, the constituent elements of the facts of the world, are phenomenological, the language that describes them is also phenomenological.

Language is a picture of the world if the elements of language stand in veridical relation to the elements of the world. There must be an internal similarity between the two. It can be so if the picture is logical, not physical (it is a picture, not a photograph). Its characteristic is descriptive and its veridicality is approximate, not exact (the relationship is not of an univocal identity but of analogy[4]).

A name is a sign becoming symbol in a proposition representing facts. Its function is to describe reality and thus establish its meaning in language. Combinations of names are propositions in language. From elementary to complex names as pictures, from elementary objects to elementary and complex facts, – this is how language and the world are related. The function of language is to represent reality.

The objects in the *Tractatus* are phenomenological objects: they are thoughts isomorphicaly resembling the facts in the world. Therefore the language of the *Tractatus* about these objects is also phenomenological. Wittgenstein knew about the mismatch between phenomenological language as the second order language of description of what thought is about and the facts of the world composed of elementary objects. This is why at the end of the *Tractatus* we are told that the book itself is nonsense, that it serves only to elucidate its readers in seeing reality rightly. Physical language of ordinary discourse, on the other hand, is that of ostensive definitions or simply the first order language about facts and their objects in the world.

What can be said, must be said clearly, in an ostensive way. What does not come as

---

[3] Some realist interpretations, however, will insist on the simple nature of the *Tractarian* objects without affirming their phenomenological nature. For Harré these objects constitute the state of affairs (the relationship of things) by simple arrangements between them. Picture theory, thus, is the representation of these states of affair. It is not clear, however, how such picture theory can account for the meaning of propositions if the objects are simply physical ones. The propositions 2. of the TLP where Wittgenstein states the difference between objects (*Gegenstände*) and things (*Dingen*) remain obscure in this realist interpretation. Harré seems to build his interpretation on the mistaken views that Tractarian objects only include particular things and do not extend to relations and properties in Harré (2001). Such views are common in Copi (1958) and Anscombe (1959). They were criticized in the works of Hintikka and Hintikka (1986: ch. 2 and ch. 3) and Stenius (1964: ch. 3 and ch. 5).

[4] See Wittgenstein (1974a: 4.063); Wittgenstein (1979a: 23).

elements of facts or facts themselves, cannot be described by isomorphic (phenomenological) language. An attempt to say it would result in nonsense. But whatever (outside of facts) cannot be said must be shown.

Nevertheless, if our language of description is phenomenological, it cannot be direct. Even if we agree that what language describes (viz. pictures) are phenomenological objects and facts, these facts and objects refer to the physical world. However, since this physical world is described in a non–physical language, the entire discourse is then a mismatch of language and reality: "The assumption that a phenomenological language is possible and that only it would express what we really must/want to say in philosophy, is – I think – absurd. We have to manage it with our ordinary language and only understand it correctly. I.e. we must not let ourselves be tempted by it to talk nonsense".[5]

This nonsense is that mismatch between language and reality. The mismatch occurs when phenomenological language does the job of describing physical reality.

In the late 1920's and early 1930's Wittgenstein moves from the idea that phenomenological language is false in expressing meaning or unnecessary to the idea that it is simply impossible. What Wittgenstein means by "no longer possible" is not that the language of logical analysis is not possible but that such language cannot express how things are in the world. By '*our*' language is meant physical language that alone must have the job of describing physical reality.

The major achievement of the *Tractatus*, as he believed, was that such language was given as an alternative to logicisms of Frege and Russell. But isomorphism alone cannot convey the meaning of its propositions; hence it is impossible as language in the ordinary sense, even if it is still language expressible by logical syntax.

The futility in inventing such language was expressed by Wittgenstein in his conversations with the members of the Vienna Circle in 1930: "I used to believe that there is the colloquial language in which we all usually speak and a primary language that expresses what we actually know, viz. the phenomena. I have also spoken of a first system and a second system. I would now like to explain, why I no longer hold fast to this view".[6]

Here Wittgenstein refers to the fact that the phenomenological language of logic of the *Tractatus* was only to refer to the phenomena and not to the ordinary experiences we have. For the later there is the 'colloquial' language. Notice the separation he makes between the two languages in his description of his earlier views. In the background we can hear Frege's mistrust in the ordinary language and his predilection for the ideal language. Only that ideal language (freed from uncertainties) was to be set in the logic of the *Begriffsschrift* or, in the case of Wittgenstein, in the logic of the *Tractatus*: "I believe that we essentially have only one language and that is the ordinary language. Not only

[5] „Die Annahme dass eine phänomenologische Sprache möglich wäre und die eigentlich erst das sagen würde was wir in der Philosophie ausdrücken müssen/wollen ist – glaube ich – absurd. Wir müssen mit unserer gewöhnlichen Sprache auskommen und sie nur richtig verstehen. D.h. wir dürfen uns nicht von ihr verleiten lassen Unsinn zu reden". Wittgnstein (2000: MS 107: 176, 22.10.1929).

[6] „Ich habe früher geglaubt, daß es die Umgangssprache gibt, in der wir alle für gewöhnlich sprechen und eine primäre Sprache, die das ausdrückt, was wir wirklich wissen, also die Phänomene. Ich habe auch von einem ersten System und einem zweiten System gesprochen. Ich möchte jetzt ausführen, warum ich an dieser Auffassung nicht mehr festhalte". Wittgenstein (1979b: 45).

do we not need to invent a new language or to construct a symbolism, but the spoken language *is* already *the* language, provided that we free it from the lack of clarity that lies in it".[7]

Phenomenological language fails to represent reality in a meaningful way. Later Wittgenstein in the *Philosophical Investigations* (Part I written mostly between 1936 and 1946, published in 1953) will express his skepticism that the ordinary language alone is successful in expressing the meaning (his anti–referential meaning–as–use semantics). But at this time, in the early 1930's, his main point is to insist that phenomenological language is not only unnecessary but also *impossible* in its conveyance of meaning.

If we allow for phenomenological language to describe physical reality then this language would be private, whose meaning would be inaccessible in public use. But that is a contradiction: how can language with the rules of grammar being public have a private semantics? Not only would private psychological states be inaccessible to anyone except its user, but even the description of objective facts would be private, resulting in some kind of solipsism and idealism.

To avoid the obvious contradiction, phenomenological language, including its description of private psychological states, must be given up for physical language. Another reason for this substitution is the fact that phenomenological language whose main function is to describe reality by ostensive definitions of objects by names, is very limited even in its expression. Such language would not be able to include verbs, adverbs, adjectives, connectives, and indexicals:

> One of the implements of our language is ostensive definition. But with such ostensive signs we have only a mere calculus.
>
> What we call a connection between language and reality is the connection between spoken language and, for example, the language of gestures. If we had no written or spoken language, where then would be the connection? How can you explain one gesture by another?[8]

The importance of this remark from Wittgenstein's lecture from 1931 is that he makes clear what the connection between language and reality is and what is not. What it is he will call 'language–games' and what is not are the ostensive definitions. For instance, even when restricted to nouns, such language with the meaning of a name standing in ostensive definition to an object (or fact) would be lost whenever that object ceases to exist. The meaning of the name 'Socrates' exists even if its bearer does not (PI §40).

In order to account for language's complexity a simple exchange from phenomenological to physical language is not enough. Logical positivism made such a turn in the works of Neurath and Carnap by suggesting that physical language should procure with protocol sentences as the foundation of all scientific knowledge. These sentences were supposed to capture the observation of reality as it is. The formulation of protocol sentences does not insure, however, that their use will not be somehow dependent upon the

---

[7] „Ich glaube, daß wir im Wesen nur eine Sprache haben und das ist die gewöhnliche Sprache. Wir brauchen nicht erst eine neue Sprache zu erfinden oder eine Symbolik zu konstruieren, sondern die Umgangssprache *ist* bereits *die* Sprache, vorausgesetzt, daß wir sie von den Unklarheiten, die in ihr stecken, befreien". Wittgenstein (1979a: 45).

[8] Wittgenstein (1980: §102).

subjective experience of the scientist or that they will not be theory–laden.[9]

Instead of excluding subjectivity from physical language, or focusing on the subjective part of language, one can embrace observational relativism and concentrate on the meaning of language as it is used. Wittgenstein's proposal of meaning–as–use substitutes his earlier theory of meaning of an expression as the object to which it refers. This earlier theory, called 'the picture theory of meaning', had been criticized by Wittgenstein in PI §1 and this criticism is referred to in the secondary literature by the name of the 'Augustinian picture of language'. The idea is that sign (and its use) pictures different possibilities of object–world relations. In his later philosophy, instead, a sign acquires its meaningfulness not in the way it stands to the object it signifies, but by having a specific (but by no means unique) rule–governed use.

"Every sign *by itself* is dead, but *what* gives it life? – It *lives* by use. Does it have in itself the living breath? Or is the *use* its breath?"[10] The sign is inspired (from *in–spirare*, breathed upon; Anscombe's translation is "breathed upon it") by the way we use it. The meaning of the sign is not inspired by some abstract meaning–ascription to it but by a concrete use of it.

The meaning of a sign, thus, depends on the way it is used. But also the essence of a sign depends on its use, not on the internal structure of the sign, the way it appears, etc. In this regard, Wittgenstein departs significantly from the ostensive nature of the picture theory, according to which the atomism of a sign isomorphically corresponds to the structure of the world. In this 'meaning as use' notion, sign's meaning depends on how it is used: "The meaning of a word is its use in the language".[11]

Edward Minar speaks of the meaning as use in terms of what practice the sign has in our lives, what role it plays: "The life of the sign lies in our lives in language".[12] But this practice, life and use, is none other than language–game being played all the time. There is no need of separating rules and grammar from use and meaning, and then try to reassemble it somehow in our practices and customs.

The above passage quoted from *Philosophical Investigations* has two received and at the same time self–excluding interpretations. First, with these words Wittgenstein is often credited with providing the theory of meaning as use.[13] Second, it is claimed that Wittgenstein did not even provide any notion of meaning, instead offered advice: "Don't ask for the meaning; ask for the use".[14]

---

[9] Carnap (1931a: 432–465); Neurath (1932). Carnap's and Neurath's "*physikalische Sprache*" sometimes goes by the name of "*Physikalismus*", physicalism, which, however, must not be confused with physicalism as the position in the philosophy of mind, treated in the previous chapter. The uses of "physical language" in Carnap and Wittgenstein are also different. David Stern makes the difference very clear: "Wittgenstein's physicalism in the early 1930s amounts to a commit-ment to the primacy of the objects we discuss in our ordinary language, while Carnap's physical-ism turns on the primacy of the objects posited by the physical scientist". Stern (2008: 329).

[10] „Jedes Zeichen scheint *allein* tot. *Was* gibt ihm Leben? – Im Gebrauch *lebt* es. Hat es da den lebenden Atem in sich? – Oder ist der *Gebrauch* sein Atem?" Wittgenstein (2001: §432).

[11] „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache". Wittgenstein (2001: §43a).

[12] Minar (2011: 289).

[13] Akmajian and Demers (1990: 201).

[14] Brown (1962).

If Wittgenstein suggested a theory of meaning he never specified what this use is, nor how exactly the meaning is connected to the use. If, on the other hand, the meaning should be eliminated somehow from our language, or reduced to the mere use by the linguistic community (as Quine suggested), then why is this attention to the meaning in *Philosophical Investigations* at all? This later suggestion is reinforced when one considers the existence of some object beyond the sign for which the meaning is supposed to stand. Wittgenstein considers it from the perspective of the German word *Bedeutung*, which derives from *deuten*, 'point', 'indicate'.[15]

If we look at the larger context of §43, we notice that it stands after Wittgenstein's criticism of the Augustinian picture of language theory (or in effect his own picture of language theory of the *Tractatus*), and as a series of short responses and suggestions of new ways of looking at language acquisition and use. If we look at the immediate context of §43, we see that it is sandwiched between or rather *within* his explanation of how names are used in primitive language–games. Looking at §42 and §44 we would notice that these texts are quite concrete explanations of some primitive language–games and the examples of how names are ordinary used in them. Thus, §43 where the famous definition of meaning occurs, can indeed be read as an addendum to the text rather than examples, amplifications of §43. It can be easily seen if one would read §42 and then skip §43 to read §44: no meaning is compromised, the text seems to flow perfectly and refer to the analysis of the primitive language–games which Wittgenstein started already at §1.

What this means is not that the landmark *Investigations'* definition in §43 should be disregarded as marginal, but rather that this definition should be read in the context of the general discussion on language–games and that 'meaning as use' instead of being a novel theory of meaning, is indeed a methodological explanation of the *modus operandi* of the language–games.

It should come as no surprise that this somewhat chaotic way of assigning important definitions, explanations, notions and theories had been done by Wittgenstein at least since the *Tractatus*. The later is famously divided into sets of numbered paragraphs under seven major propositions of the book. Erik Stenius remarked that quite often the more important passages are assigned to the paragraphs with more numerals, but sometimes this ordering rule is reversed.[16] In the *Investigations* just as in the *Tractatus*, Wittgenstein 'feels free' to follow this rule or the opposite of it. I believe that §§42–44 are examples of the same technique as in the *Tractatus*, where the more important point can serve as an example or as an explanation of the less weighty or more technical one.

The paragraphs of the *Investigations* are not divided hierarchically as in the *Tractatus*, but one can get the rhythm from the style of the narrative and sense the emphasis of the text. The important point here is to regard the notion of use as a use in and by the language–game. Even if Wittgenstein does not provide here (nor anywhere else in the PI) a theory of meaning, this use in the language–game can be made precise and clear by making concise analysis of each particular language–game.

Picture theory in the *Tractatus* was meant to explain how words (names) are to relate

[15] Wittgenstein (1974b: I, §19).

[16] Stenius (1964: 4–5).

to reality. Although the term 'ostensive definition' never occured in the *Tractatus* (it will, however, occur in the PI), language refers to reality *ostensively*, or, in Wittgenstein's words, language projects itself on reality by picturing it.[17] This isomorphism is closely related to Frege's and Russell's formalism of logic and language at least in the aspect of its immediacy of relation between word and (its) object. Logical formalism in general does not explain what makes up these projections of language on reality.

The same omission was committed by Wittgenstein in the *Tractatus*, and in the transitional years (1924–29) Wittgenstein was beginning to realize that the immediacy of language and reality is unwarranted especially if one prefers physical language of description over phenomenological language of the likewise physical reality. Picture theory was true, for it indeed does the job of picturing the world of facts and objects, but it was not enough when, i.e. mental events entered into that same realty of the world of facts. These 'mental facts' under phenomenological language would imply a strict solipsism and the language that would describe them can only be private in nature. Yet, 'private language' is a contradiction of terms for all language is public, based on the public rules of grammar and semantics.

Furthermore, 'picturing' involves not just *naming* but most importantly *showing* how things are in the world: "The limits of my language mean the limits of my world".[18] Whatever is expressed in language itself requires linguistic representation, but that would not be possible: we cannot step outside of our language by means of language: "What expresses *itself* in language, *we* cannot express through language. The proposition *shows* the logical form of reality. It makes it apparent".[19] What this showing (*zeigen*) is, cannot be specified by means of language, but it will be developed into his later notion of the language–games as the promised showing. Thus, something more to isomorphism was needed in the use of physical language of description. Logic cannot take care of itself (see NB 4; TLP 6.123) if logic is not based anymore on isomorphic relations. These relations are not sufficient for the explaining the relationship between language and the world.

From the early 1930's Wittgenstein begun to speak of linguistic activities in terms of games, from the most primitive ones as calculation to more complex ones, i.e., children's games, chess, football, etc. These games are rule–governed human activities that indicate the complex relationships between language and reality. These activities are linguistic uses and they are the ones that represent the relationship between language and reality.

Compare these two sets of passages from TLP and PR:

> The configuration of the objects forms the relationship of things.
> In the relationship of things objects hang on each other, like links in a chain.
> In the relationship of things the objects are combined in a certain way.
> The way in which objects are related in the relationship of things is the structure of the relationship of things.

---

[17] Wittgenstein (1974a: 2.172ff).

[18] „*Die Grenzen meiner Sprache* bedeuten die Grenzen meiner Welt". Wittgenstein (1974a: 5.6).

[19] „Was *sich* in der Sprache ausdrückt, können *wir* nicht durch sie ausdrücken.
Der Satz *zeigt* die logische Form der Wirklichkeit.
Er weist sie auf". Wittgenstein (1974a: 4.121).

> Objects I can only *name*. Signs represent them. I can only speak *of* them; *I cannot express them*. A proposition can only say *how* a thing is, not *what* it is.[20]

> By application I take it to mean what makes the connections of sounds or lines in the first place a language. I mean in the sense that it is the application that puts a bar with lines to a scale: *matching* language with reality.[21]

In the passage from PR we can now begin to see how "the objects are combined" and what is this "certain way": it is, "by application", by the relation one to another that this combination is made possible (in the TLP quoted above, this relationship is immediate, isomorphic). Furthermore, the expression of the signs is made possible through the same application. This application of signs and sounds is the use of language, and the use of language is a language–activity, like a theater (*Theaterspiel*), or an actor's acting (n. *Schauspieler*, adj. *schauspielerisch*, vi. *schauspielern*). Wittgenstein's term for that new relationship between word and object is *Sprachspiel*, language–game.

There is an important text by Frank Ramsey from 1923 which is naturally read as a critical commentary to the TLP.[22] However, closer to the end of the article Ramsey suggests a solution to the Tractarian doctrine of the ineffability of semantics in ordinary language and the inability of stepping out of the language by means of language. The solution for Ramsey is to look into the use of language and not the language itself. This, by my understanding, is the prolegomenon to any future notion of meaning–as–use and, perhaps, even a foresight of the upcoming language–games:

> We must agree with Mr. Wittgenstein that "*p* is of such and such a form" is nonsense, but "'*p*' has a sense of such and such a form" may nevertheless not be nonsense. Whether it is or not depends on the analysis of "'*p*' is significant", which seems to me probably a disjunctive proposition, whose alternatives arise partly from the different possible forms of the sense of "*p*". If this is so, we can by excluding some of these alternatives make a proposition as to the form of the sense of "*p*". And this in certain cases, such as when "*p*" is "He thinks *q*" or "He sees *a*", could be appropriately called a philosophical proposition.[23]

These philosophical propositions of "thinking" something and "seeing" something are an activity in use of language. This activity alone, according to Ramsey, can make propositions meaningful, acquiring sense through different uses by its subjects.

---

[20] „Die Konfiguration der Gegenstände bildet den Sachverhalt.

Im Sachverhalt hängen die Gegenstände ineinander, wie die Glieder einer Kette.

Im Sachverhalt verhalten sich die Gegenstände in bestimmter Art und Weise zueinander.

Die Art und Weise, wie die Gegenstände im Sachverhalt zusammenhängen, ist die Structur des Sachverhaltes". Wittgenstein (1974a: 2.0272–2.032).

„Die Gegenstände kann ich nur *nennen*. Zeichen vertreten sie. Ich kann nur *von* ihnen sprechen, *sie aussprechen kann ich nicht.* Ein Satz kann nur sagen, *wie* ein Ding ist, nicht *was* es ist". Wittgenstein (1974a: 3.221).

[21] „Unter Anwendung verstehe ich das, was die Lautverbindungen oder Striche überhaupt zu einer Sprache macht. In dem Sinn, in dem es die Anwendung ist, die den Stab mit Strichen zu einem Maßstab macht. Das *Anlegen* der Sprache an die Wirklichkeit". Wittgenstein (1975: 85).

[22] Ramsey (1923).

[23] Ramsey (1923: 477).

2. Analysis of the Language–Games

For the first time Wittgenstein mentions language–games in BB 17 as an example of the notion of use. But at this time language–games were not considered by him as representations of reality by language. This idea matured only from PI I on. Wittgenstein's idea that language–games constitute the representative relation between language and reality was brought forward by the interpretations of Jaakko Hintikka and Merrill Hintikka. Likewise, they indicated an important distinction between 'primitive' and 'complete' language–games, used quite often by Wittgenstein but largely ignored by the secondary literature.

Their vocabulary is of 'primary' and 'secondary' language–games. In Wittgenstein's writings, however, the terms are 'primitive' (*das primitiv Sprachspiel*) and 'complete' (*das Ganze* (or *vollständig*) *Sprachspiel*). This vocabulary will be used in this work as well:

> We can also think that the whole process of using words in (2) [PI §2] is one of those games by means of which children learn their mother tongue. I want to call these games '*language–games*', and sometimes speak of a primitive language as a language–game.[24]
>
> I will also call the whole, how the language and the actions are interwoven, the 'language game'.[25]

From the passage is also clear that Wittgenstein often uses interchangeably 'language' and 'language–games'. This nuance is important, however, keeping in mind that his use of 'language' is not the same as the regular usage in the ordinary language as "the method of human communication, either spoken or written, consisting of the use of words in an agreed way" (*The Concise Oxford Dictionary*). Thus, Wittgenstein's use of 'language' comprises that of the ordinary language but is much larger than the later (see PI §491).

From *Blue Book* (1930s) to *On Certainty* (1951) Wittgenstein explains many uses of primitive and complete language–games. A comprehensive list would be very long, but these few examples might suffice for having an idea of the universal applicability of language–games in Wittgenstein's later philosophy. Primitive language–games include games of lying (PI §§249–50), expectation (PI §442, §§576–77, §§581–83), certainty of knowledge (OC §§56–57), imagination (OC §18), giving and receiving orders (PI §2, §8, §630). Complete language–games include sensations (PI §290), use of signs (BB 17), naming (PI §§2–7, PG 62–63), counting (PI §146), children learning (PI II: 200, RPP II: §453), color vision (RPP II: §330, Z §420). In PI §23 Wittgenstein lists some of the most common complete language–games.

Perhaps the most vivid passage that illustrates primitive vs. complete language–

---

[24] „Wir können uns auch denken, daß der ganze Vorgang des Gebrauchs der Worte in (2) eines jener Spiele ist, mittels welcher Kinder ihre Muttersprache erlernen. Ich will diese Spiele ‚*Sprachspiele*' nennen, und von einer primitiven Sprache manchmal als einem Sprachspiel reden". Wittgenstein (2001: §7b).

[25] „Ich werde auch das Ganze: der Sprache und der Tätigkeiten, mit denen sie verwoben ist, das ‚Sprachspiel' nennen". Wittgenstein (2001: §7d). For the use of '*vollständig*' see also, i.e., Wittgenstein (2001: §18).

games is from PI §18 where Wittgenstein compares our language to an ancient city:

> Let it not disturb you that languages (2) and (8) [PI §2, §8] consist only of orders. If you want to say that they are not complete (*nicht vollständig*), then ask yourself whether our language is complete (*vollständig ist*), – whether it was complete before the symbolism of chemistry and of infinitesimal notation became part of it, because these are, so to speak, suburbs of our language. (And with how many houses or streets, does a city begin to be a city?) Our language can be seen as an ancient city: a maze of narrow streets and squares, of old and new houses, and houses with additions from various periods, and this is surrounded by a multitude of new boroughs with straight regular streets and uniform houses.[26]

In this passage language is seen as a historically and culturally developing activity, incorporating within itself new forms and new sub–languages as it grows larger and richer. Being complete would mean in the sense of its complete particular 'boroughs', such as symbolism of chemistry or of the infinitesimal notation, but also in terms of the language as whole, just as a whole city. In the *Lectures on Aesthetics*, Wittgenstein once referred to the whole of culture as one language–game: "What belongs to a language–game is a whole culture".[27]

Notice, however, that the language of "the symbolism of chemistry and of the infinitesimal notation" is that phenomenological language which now is given the status of a suburb in the city of physical language. Suburbs depend on the city as a whole and the semantics of a phenomenological language depends on the semantics of physical languages: i.e. Aventino belongs to and depends on Rome as a *rione* of the city.

Every language–game, just as any game, is constituted by rules that define the game and strategies that employ these rules in order to win the game (i.e. acquire the meaning, coming to understanding). To the notion of rule–following Wittgenstein dedicates an important part of the *Investigations*, §§185–242. These precede immediately his so–called 'Private Language Argument' (§§243–315).

What does it mean to follow a rule of a particular language–game? In the second chapter it was already mentioned that mere mechanical rule–following is not enough to account for thinking. The problem with unreflective rule–following is that one does not know whether s/he is applying the rule correctly. If language–games are to be the representative links between language and reality then rule–following in a particular language–game must be mirroring the rules of a particular linguistic community. In PI §206 Wittgenstein asks us to imagine an explorer who comes across a language unknown to him/her. How one can say that people in that community behave according to certain rules? He explains:

---

[26] „Daß die Sprachen (2) und (8) nur aus Befehlen bestehen, laß dich nicht stören. Willst du sagen, sie seien darum nicht vollständig, so frage dich, ob unsere Sprache vollständig ist; – ob sie es war, ehe ihr der chemische Symbolismus und die Infinitesimalnotation einverleibt wurden; denn dies sind, sozusagen, Vorstädte unserer Sprache. (Und mit wieviel Häusern, oder Straßen, fängt eine Stadt an, Stadt zu sein?) Unsere Sprache kann man ansehen als eine alte Stadt: Ein Gewinkel von Gäßchen und Plätzen, alten und neuen Häusern, und Häusern mit Zubauten aus verschiedenen Zeiten; und dies umgeben von einer Menge neuer Vororte mit geraden und regelmäßigen Straßen und mit einförmigen Häusern". Wittgenstein (2001: §18).

[27] Wittgenstein (1978: §26).

> Following a rule is analogous to following a command. One is trained to do so and one reacts to it in certain ways. But what if someone reacts to instruction and training *one way*, and another in *a different way*? Who is right?
>
> Imagine you came as a researcher to an unknown country with a language totally foreign to you. Under what circumstances would you say that the people [there] give commands, understand and follow commands, to rebel against orders, etc.?
>
> The shared way of human acting is the reference system by which we interpret foreign language [unknown] to us.[28]

To answer the question about how one knows how to follow the rule, one must investigate first the way in which a particular language–game is played. Learned rule–following is playing the game; this is the very same activity. If one knows nothing about the game, one cannot follow the rule, unless one follows the rule automatically, i.e. by someone else's command.

Although rules define each game, rules are made in regard to the entire game, in regard to the many strategies the game has.[29] Quite often the strategies modify their own rules. For example, in the game of chess two rules introduced to avoid the stalemate situations in the game. These rules are, i.e., *en passant* and promotion of the pawn into the queen. These two rules facilitate a more swift development of the game and can be employed freely in numerous chess strategies. In the game of cricket a Law 21 was introduced to decide on the tie in case both teams have the same score for the period of more than 7 days of playing the same match. Although figure skating is an old sport, after the infamous scandal in the 2002 Winter Olympics an entirely new judging system was introduced to determine the winner(s) in competition, viz. the old strategies of the sport required a new system of evaluation.

---

[28] „Einer Regel folgen, das ist analog dem: einen Befehl befolgen. Man wird dazu abgerichtet und man reagiert auf ihn in bestimmter Weise. Aber wie, wenn nun der Eine *so*, der Andere *anders* auf Befehl und Abrichtung reagiert? Wer hat dann Recht?

Denke, du kämst als Forscher in ein unbekanntes Land mit einer dir gänzlich fremden Sprache. Unter welchen Umständen würdest du sagen, daß die Leute dort Befehle geben, Befehle verstehen, befolgen, sich gegen Befehle auflehnen, usw.?

Die gemeinsame menschliche Handlungsweise ist das Bezugssystem, mittels dessen wir uns eine fremde Sprache deuten". Wittgenstein (2001: §206).

Jaakko and Merrill Hintikka translate the last sentence "die gemeinsame menschliche Handlungsweise" as "people's shared way of acting" or "the shared human way of acting" against Anscombe's translation "the common behavior of humankind" Wittgenstein (2001: §70). The last translation would go against Wittgenstein's linguistic and cultural relativism that would *not* presume that there is a common way of behaving for the entire humankind. In fact, in this context, 'common' or 'shared' stands for that imaginary linguistic community met by the imaginary researcher in Hintikka and Hintikka (1986: 209, n.17). Hacker's and Schulte's translation is slightly amended but the words 'shared' and 'behavior' are not: "Shared human behavior is the system of reference by means of which we interpret an unknown language". Wittgenstein (2009: §88^{e}).

[29] Wittgenstein himself has used the word 'strategy' in the manuscripts in relation to the war trategies Wittgenstein (2000: MS 210:20, 1 Jan, 1930) and in relation to the game strategies. Wittgenstein (2000: MS 136: 141b, 23 Jan., 1948).

In short, rules of the game determine and define each and every strategy of the game and the same rules are made in view of the strategies. Being proficient in knowledge about the rules of the game is not enough for playing the game. This is the point of Wittgenstein in PI §206 as it was in the quoted paragraphs of the RFM in chapter II. We understand (and not just know in the sense of being informed) the rules of a particular language–game only if we studied its implementations in the strategies of that game. Being able to discern whether in a particular community the people's behavior corresponds to obeying orders means knowing their way of acting first.

Let's take as examples two very different games: chess and soccer. It does not take much time to learn how to play chess and, perhaps, even less how to play soccer. The sets of rules for both games are very limited and each game's rules can be learned from a book or some other list of rules. Yet knowing every rule of chess still does not make one a chess player. One must master strategies of the game of chess in order to play the game. The same goes for soccer or any other game. Whether the game consists in moving wooden pieces on the board or intense physical activities, each game demands the mastery of strategies and a very long time of learning them in actual playing. In this sense one may use the term 'strategy' and 'game' interchangeably.

There is a difference between the middle and later periods of Wittgenstein's philosophy in terms of considering rules to be primary to their games. In PG he speaks of rules as the grammatical description of language. If rules define the game, they also necessarily define its strategies: "We regard language *from* the game's *point of view*, as it follows fixed rules. We compare it with such a game like this, by measuring against it".[30]

But in the period of the *Investigations* Wittgenstein gives priority to the games (or their strategies) over their defining rules. The idea of the game's priority over its rule, it seems, emerges when Wittgenstein realizes that one can learn a game without learning its rules. This passage from the very same PG goes in harmony with PI §31: "Indeed, the grammar of a language is not recorded nor comes into existence until the language has been spoken for a *long* time by the people. Likewise, primitive games are also played without having put up a list of rules; perhaps even without a single rule having been formulated for it".[31]

Having indicated that primitive games *can* be played without rules being formulated, Wittgenstein immediately reaffirms the primacy of the rules, not of the games themselves: "But we consider the games and the language from the point of a game that performs according to rules. That is, we always compare the language with such a process".[32] Compare the first paragraph with this passage from PI §31: "One can also imagine: one learned the game without ever learning or formulating rules. He has learned by

---

[30] „Wir betrachten die Sprache *unter dem Gesichtspunkt* des Spieles nach festen Regeln. Wir vergleichen sie mit so einem Spiel, messen sie an ihm". Wittgenstein (1974b: I, iii, §36).

[31] „Ja so, wie die Grammatik einer Sprache erst aufgezeichnet wird und erst in die Existenz tritt, wenn die Sprache schon *lange* von den Menschen gesprochen worden ist, werden primitive Spiele auch gespielt, ohne daß ihr Regelverzeichnis angelegt wäre, ja wohl auch, ohne daß eine einzige Regel dafür formuliert worden wäre". Wittgenstein (1974b: II, ii, §26).

[32] „Wir aber betrachten die Spiele und die Sprache unter dem Gesichtspunkt eines Spiels, das nach Regeln vor sich geht. D. h. wir *vergleichen* die Sprache immer mit so einem Vorgang". Wittgenstein (1974b: II, ii, §26).

watching first some simple board games and has progressed to more and more complicated ones".[33]

In PI §31b Wittgenstein, indeed, reports a possibility of knowing how to play the game without ever learning any rules, i.e. by observing it being played by others. In that case we can say that the person had learned the rules of the game (even without formulating them) from the observation of the strategies of the game. But the converse is not possible: by mere knowledge of all rules (limited in number) one cannot know the strategies of that game, potentially infinite in many games (see also PI §556, n. b).

How can we define rules and strategies not from particular language–games but in general? In what way can we speak of rules and strategies that would apply to all language–games regardless if they are primitive or complete, whether they refer to language or just to any human activity? According to Wittgenstein we cannot, or at least, we should not. The passage from *The Blue Book* that introduced the notion of the language–game for the first time in his writings urges us to give up our craving for generality:

> If we say thinking is essentially operating with signs, the first question you might ask is: "What are signs?" –Instead of giving any kind of general answer to this question, I shall propose to you to look closely at particular cases which we should call "operating with signs". Let us look at a simple example of operating with words. I give someone the order: "fetch me six apples from the grocer", and I will describe a way of making use of such an order: The words "six apples" are written on a bit of paper, the paper is handed to the grocer, the grocer compares the word "apple" with labels on different shelves. He finds it to agree with one of the labels, counts from 1 to the number written on the slip of paper, and for every number counted takes a fruit off the shelf and puts it in a bag. –And here you have a case of the use of words. I shall in the future again and again draw your attention to what I shall call language games. These are ways of using signs simpler than those in which we use the signs of our highly complicated everyday language. Language games are the forms of language with which a child begins to make use of words. The study of language games is the study of primitive forms of language or primitive languages. If we want to study the problems of truth and falsehood, of the agreement and disagreement of propositions with reality, of the nature of assertion, assumption, and question, we shall with great advantage look at primitive forms of language in which these forms of thinking appear without the confusing background of highly complicated processes of thought. When we look at such simple forms of language the mental mist which seems to enshroud our ordinary use of language disappears. We see activities, reactions, which are clear–cut and transparent. On the other hand we recognize in these simple processes forms of language not separated by a break from our more complicated ones. We see that we can build up the complicated forms from the primitive ones by gradually adding new forms.
>
> Now what makes it difficult for us to take this line of investigation is our craving for generality.

[33] „Man kann sich aber auch denken. Einer habe das Spiel gelernt, ohne je Regeln zu lernen, oder zu formulieren. Er hat etwa zuerst durch Zusehen ganz einfache Brettspiele gelernt und ist zu immer komplizierteren fortgeschritten". Wittgenstein (2001: §31b).

> This craving for generality is the resultant of a number of tendencies connected with particular philosophical confusions.[34]

In this passage Wittgenstein brings up for the first time his new notion of language–games. He defines them as the representative links between language and reality and as an activity of use. But he refuses to make out of this new notion a new philosophical theory. The essence of a language–game is to be played in many activities. It is in these activities that the meaning reveals itself. Defining rules and strategies would be counter-productive. First, because one game is always composed of many other games and in our activities we are always free to pass from one game to another. This implies that new games are introduced by changing rules. Second, because strategies of games change even more often than the rules. Old strategies give rise to new rules and with that to the new games entirely. As a consequence of this unending dynamism, instead of fixing our attention to the ever–changing rules and strategies of games we should, instead, pay attention to the games themselves. Meaning is not expressed in the rules nor even in the strategies of games but in games themselves.

A potential definition of rules and strategies in the game is redundant, at least for philosophical discourse, because once we understand the game, we do not need any definition of the rules and strategies of that game. All we need to do is to play it or learn from it. On the other hand, a potential definition of *all* rules and strategies of *any* language–game in general will inevitably exclude some future eventual instances of making up new rules and strategies and thus, exclude in principle some potential games from our philosophizing. In both cases, the craving for generality must be satisfied by our active *playing* of these games each moment of our lives, not by our passive *studying* them.

Later Wittgenstein's anti–theoretical approach to philosophy has its origin from the earlier belief in the ineffability of semantics but it is also based on his notion of the language–games. Theory is "a supposition or system of ideas explaining something, especially one based on general principles independent of the particular thing to be explained" (*Oxford English Dictionary*). In other words, theory is an exercise in stepping outside of language by means of language. The Greek *theōria* comes from *theōros* and means 'spectator' from *theōreō* 'look at'. On the contrary, language–games are for playing and even a language–game spectator as a spectator has a role to play in that game: "The origin and the primitive form of the language game is a reaction; only from this can grow the more complex forms. The language – I want to say – is a refinement, 'In the beginning was the deed'. (Goethe: *Faust* I)".[35] Strictly speaking, there are no independent spectators in language–games.

But some generality is allowed by Wittgenstein. First, it is obvious that all games have rules and strategies. Second, language–games are the link between language and reality. Third, there are primitive and complete language–games. The last point is stressed by Wittgenstein in the above passage: "we can build up the complicated forms from the

---

[34] Wittgenstein (1958: 16–17); Wittgenstein (2001: §1).

[35] „Der Ursprung und die primitive Form des Sprachspiels ist eine Reaktion; erst auf dieser können die komplizierteren Formen wachsen.

Die Sprache – will ich sagen – ist eine Verfeinerung, ‚im Anfang war die Tat.' (Goethe: *Faust* I)". Wittgenstein (1969: §31).

primitive ones by gradually adding new forms". This goes in hand with the examined passage PI §18, where complexity of language is demonstrated by applying the notion of language–games. Just as the city is growing by the addition of new houses and parts, so is our language developing by adding more complexity to it.

Perhaps that much generality in the study of language–games can be allowed. We can look at the language–games as existing in three–dimensional space. The first dimension is that of rules and strategies. The second dimension is that of primitive and complete. The third dimension is that of different language–games in sets of their family resemblances (*Familienähnlichkeiten*) and different forms of life (*Lebensformen*).[36]

The distinction between primitive and complete language–games will be explained further in the next section in the context of private sensations. The second point about language–games as the link between reality and language needs to be explored further in connection with the first point, that all games are composed of rules and strategies. One of the most telling passages about this relation comes from the BB:

> What now is the relation between a name and the object named, say, the house and its name? I suppose we could give either of two answers. The one is that the relation consists in certain strokes having been painted on the door of the house. The second answer I meant is that the relation we are concerned with is established, not just by painting these strokes on the door, but by the particular role which they play in the practice of our language as we have been sketching it. –Again, the relation of the name of a person to the person here consists in the person having been trained to run up to someone who calls out the name; or again, we might say that it consists in this and the whole of the usage of the name in the language–game.[37]

Jaakko and Merrill Hintikka compare this passage against Wittgenstein's earlier ostensive use of names in the TLP.[38] While in the TLP name and object stand in an immediate, isomorphic relationship, here the relationship is explained in terms of the language–games as some activity (i.e. having a name and the response to someone's calling that name). In this later case, the relationship is not immediate but dynamic and complex. It can be described by the rules and strategies of the particular game. As was explained earlier, the meaning and the game coincides if the meaning is understood as its use and the use is synonymous with playing the game.

One of the most telling texts on the language–game being the main link between language and reality is PI §55: "–What corresponds with a name, and without which a name would have no meaning is, for example, a paradigm that is used in conjunction with the name in the language–game".[39] The name is a paradigm used in the language–game, and that language–game employs and elaborates the paradigm according to the use

---

[36] For forms of life being a dimension of language–games see Putnam (1992: ch. 7 and ch. 8); see also Hark (1990: ch. 2). Hark considers two dimensions in language–games: primary vs. secondary (primitive vs. complete) and forms of life (between language–games).

[37] Wittgenstein (1958: 172).

[38] Hintikka and Hintikka (1986: 193).

[39] „–Das, was dem Namen entspricht, und ohne den er keine Bedeutung hätte, ist, z. B., ein Paradigma, das im Sprachspiel in Verbindung mit dem Namen gebraucht wird". Wittgenstein (2001: §55).

imposed by the common or personal criteria and the rules of that game.

But notice that it is not the language–game being a metaphor, paradigm or some kind of analogy; it is the paradigm of the given reality, fact or event that is being employed within each language–game and elaborated by its rules and strategies. This way language–games are instruments in the game of identity between the world and language.

3. The Public Language(–Game) Activity

In the secondary literature, the passage of PI §§243–315 goes by the name of the Private Language Argument. This name is as accurate as that of the Holy Roman Empire (962–1806): just as the collection of lands was neither holy, nor Roman, nor an empire, the passage is not about private language, nor an argument.[40] In the following sections I will demonstrate that the relevant passage refers to the public character of the language–games rather than to the inexistent and impossible private language. In fact, I would insist that the subject of this passage is not language in general but on the private sensations and their expression in the language–games rather than in ordinary language. Finally, Wittgenstein never proposed any argument in which premises are followed by the conclusion.

3.1. Against Metaphor

In the previous section, based on the reading of PI §7, I briefly mentioned an important distinction Wittgenstein makes between primitive and complete language–games. From the middle period on, Wittgenstein did not regard ostensive definitions as the representative link between language and reality but ascribes such a link to the language–games. As was explained before, the reason for such preference and rather radical change in his philosophy was that ostensive definition as such could not give an analysis of the relationship between word and object; language–games can. Language–games refer to the paradigmatic use of the word in them being played in everyday activities of people. But Wittgenstein suggests (from PI §1 onward) that language–games themselves come in many varieties. Thus, not only playing but also studying them is the semantic and epistemic activities anyone, especially philosophers, can undertake.

The best place to study primitive and complete language–games is in the context of private experiences introduced by Wittgenstein in the PI §§243–315, often called as the 'Private Language Argument'. Perhaps no other passage from the PI attracted so much scholarship and invoked so many different interpretations. As Meredith Williams rightly indicates, in the argument Wittgenstein refutes three major theses: that ostensive definitions are the representative link between language and reality (against TLP, Frege and Russell); that linguistic representations require some privileged objects situated in the mind (against Descartes and phenomenology); and that 'reference' and 'meaning' are

[40] See Hintikka's and Hintikka's similar remark on the 'picture theory of language' (1986: 97).

synonymous (again, against Frege).[41] In relation to the private sensations, these three arguments can be summed up in one: that the expression of a sensation does not occur in ordinary nor in an imaginative private language. In a more positive way this last point can be expressed by saying that meaning is expressed in the active play of the language–games.

When we look at the place of the §§243–315, we notice that they come after Wittgenstein's lengthy discussion of the rule–following (§§185–242). These later paragraphs can also be seen as a conclusion of Wittgenstein's critique of the Augustinian Picture of Language (§§1–64) and general discussion on the nature of philosophy (§§65–184). We should also recall from the discussion above that these topics were explored by Wittgenstein in the light of the introduced notion of the language–games (§7, but in effect from §1, where the first primitive language–game was explained). It is also important to see that the passages following the Private Language Argument deal primarily with mental events and language, viz. on thought and imagination (§§316–397), the self and consciousness (§§398–427), intentionality (§§428–465, §§629–660), meaning (§§547–610, §§661–693) and the will (§§611–628). These large texts constitute most of the remaining space in the first part of the PI.

If we ignore the centrality of the notion of the language–games from the Private Language Argument, the interpretation of the argument would rely on false assumptions and would come to false conclusions. In fact, the received view of the argument speaks of the consistency of expressing one's sensations.[42] The skeptical interpretation of the new–Wittgensteinians insists on the resolute reading and underlines the ineffability of expression of one's sensations in principle.[43] The anti–skeptical interpretation insists essentially on the impossibility of having a private sensation.[44] This later interpretation is extended by saying that a unique meaning of sensation is impossible and that all meaning must be expressed by the community's rules of use and grammar.[45]

There is a more simplified view in the secondary literature on the Private Language Argument recently presented by David Stern. He suggests viewing the immense literature in two opposite interpretations. The first is the 'orthodox' interpretation that considers the relevant text in terms of "a deductive proof that the idea of a private language leads to contradiction".[46] The second is the 'unorthodox' reading which denies there is any argument at all.

As different as these interpretations are, they all have in common a peripheral (if any) view of the language–games in Wittgenstein's treatment of the private sensations. As some recent investigations of the manuscript material has shown, Wittgenstein himself treated the §§243–315 rather as an implementation of his earlier discussions in the PI I and not as the topic of its own.[47] No resolute reading can be limited to that passage alone without remaining obscure and ambiguous in meaning.

---

41 Williams (1999: 16).

42 Malcolm (1966: 68).

43 Crary and Read (2000); Conant (2004). Mulhall (2007).

44 Hacker (1990); Hacker (2010).

45 Kripke (1982).

46 Stern (2011: 335).

47 Stern (2010); Nielsen (2008).

If, on the other hand, we apply the notion of the language–games in its full force to the problems surrounding private sensations, we will see the text in a new light. It is in this regard that the previous analysis of the language–games can be applied to the text and the intentions of its author of the text be explored without running into contradictions.

Interpretation of the Private Language Argument is closely related to interpreting the text in which the text is located. If one thinks of the text as a series of premises followed by conclusion, then §§243–315 is hardly any argument at all. Premises are mixed, often mingled in different sections. Wittgenstein often returns to the same points discussed in previous sections. Some authors tried to compose the text into a coherent argument, but each composition necessarily brings too much or too little reading of the actual text of the §§243–315.[48]

Another difficulty of the text are, the so called 'Wittgenstein's voices'. These are two or three general lines of thought always in contradiction, always bringing different points of view to the ongoing discussion of the text.

Generally, there are two ways of interpreting Wittgenstein's 'voices' within PI: the narrator, the interlocutor and the commentator. First, most prominently presented by Backer and Hacker, is to regard the interplay of all three voices (strictly speaking, Backer and Hacker do not consider the commentator's voice) as purely rhetorical or metaphorical uses of Wittgenstein's style. The text of the PI is full of metaphors working on to elucidate one's comprehension of the author's ideas. The author here is identified with the narrator.[49]

Second interpretation given by Stern is to *not* identify Wittgenstein with any of the three voices but that all three are different characters in the text. We are invited to participate in the many examinations of the philosophical problems without necessarily identifying any particular voice with that of the author.[50] This interpretation is partially based on that of Cavell, for whom the reader is invited to internalize her/himself with the voices within the text.[51]

The problem with the first interpretation is that it often ends up in contradictions. If the narrator is Wittgenstein then the narrator's voice is changing on more than one occasion on the same topic. The problem with the second interpretation is that no final point of the text can be clear at all. If we are external spectators of the play between the

---

[48] Among the best known such compositions of the argument were done by Malcolm (1966), Rhees (1966), Strawson (1966), Ayer (1966), Kenny (1966) and Kenny (1971).

[49] Backer (2004); Hacker (2001); Backer and Hacker (1980).

[50] Stern (2004: 10–28).

[51] Cavell (1979). Cavell's interpretation looks less mysterious if one (rightly) identifies him with the New–Wittgensteinian 'resolute' interpretation of the TLP and PI, in line with Diamond, Conant, Goldfarb and Floyd. According to these interpretations, Wittgenstein is not making any points at all and all three voices serve as the ladder to reach to the elimination of metaphysical questions as senseless (TLP 6.54). This interpretation would work if one ignores completely Wittgenstein's language–games as the representative link between language and reality or, at most, considers them to be no more than rhetorical metaphors. Surprisingly, this dismissal of language–games could be the common point between the two rival interpretations of Peter Hacker and the New–Wittgensteinians. Crary and Read (2000).

three voices without identifying any of them with the author's, then what should we take as the conclusive response of the author? Cavell's interpretation is also obscure since it gives no indication as to how this internalization of the reader with the voices would proceed. Besides, if the voices are constantly changing, keeping to just one 'narrator' or 'interlocutor' or 'commentator' voice will soon result in inconsistency of my own voice with myself!

I suggest a different interpretation but somewhat close to that of Stern. If we take Wittgenstein's definitions of the language–games in PI §1 and PI §7 as the interpretative models for the entire text of the book, then the three voices in the text should be seen as the three players voicing their parts in the overall play we are witnessing. This is why Wittgenstein is usually in none of them and his final point would be the silent obvious conclusion of the argument in discussion. The point is to convince the reader without putting in words that of which we are to be convinced, but to communicate the certainty and obviousness of conclusion: "In philosophy, it is always good to *question* rather than answer a question. For an answer to the philosophical question can easily be unfair, while its settlement by means of another question, is not".[52]

The author's conclusion, which is the point Wittgenstein makes to elucidate philosophical issues in question, is also the pay–off of the language–game we are invited to participate in. These examples are, indeed, language–games that resemble more language–plays. It is possible that all three voices would *not* convey the author's point and get it all wrong. But if we were reading off the language–plays carefully, following its arguments step by step, we might be able to see the point from all three voices getting it wrong. Wittgenstein shows through these three voices *our own* philosophical and linguistic uses and misuses. If we can see ourselves in one, two or even all three of the voices, then we should be able to understand our own symptoms of our philosophical condition. Wittgenstein's seemingly absent voice is our understanding and seeing the way out of the bottle (PI §309), which is also his therapy for us.

Another point regarding the contradictions was made by L. Caruana,[53] suggesting that contradiction (such as between two opposite propositions defying the Law of Non–Contradiction, 'P and ∼P'), does not have to be regarded as an insurmountable obstacle in constructing meaningful discourse. While some contradictions are postulating two opposite and conflicting statements, such as 'It is raining and it is not raining at the same time', indeed resulting in nonsense, there are other situations in which final judgment cannot be expressed due to the incomplete knowledge in an open–ended inquiry. Building on N. Rescher's and R. Brandom's ontological and semantic distinctions,[54] Caruana calls the first case of contradiction "superimposed" and the second "perforated". To the second case we can attribute many scientific inquiries that do not provide us with complete knowledge of the matter at hand and, indeed, often result in contradictions. The

[52] „In der Philosophie ist es immer gut, statt einer Beantwortung einer Frage eine *Frage* zu setzen.

Denn eine Beantwortung der philosophischen Frage kann leicht ungerecht sein; ihre Erledigung mittels einer andern Frage ist es nicht". Wittgenstein (1983: III, §5).

[53] Caruana (2004).

[54] Rescher and Brandom (1980).

author's examples are from chemistry, quantum mechanics and logic. I suggest regarding Wittgenstein's own style of apparent contradictions as indeed perforated ones in which contradictions have pedagogical use, instrumental in making us understand where our own misuses of language lead to our contradictions on the level of superimposition.

Contrary to the common craving for generality, Wittgenstein has no common medicine for all and for every condition. He is not that kind of a philosopher. Language–games are not designed to satisfy one's craving for generality but rather to get rid of it. His point is to play within at least one of these voices our own voice and show where we went wrong. The answers to our personal voices might be personal as well. If we must find one general answer for all of us, it is in Wittgenstein's insistence that meaning is in the use and the use is the games we play (yet, this is hardly a generalization). Then, each one's game can and probably will be quite different and unique. What is common between these games are rules settled by the author; what is unique are the strategies which we develop as we go on playing.

Before applying game–linguistic solutions to the argument I shall first briefly introduce the argument as it stands in manuscripts and the PI I, say what it is about and what it is not.

The argument is one of the last pieces of the PI I, written between 1937 and 1945. Most of it was written in the second half 1944. The earliest discussion takes place in the "Notes for Lectures on Sense Data and Private Experience" (PO 202–288), composed as notes for his lectures in Cambridge in 1935–36, and were written mostly in English. Rush Rhees' lecture notes cover these lectures in PO 290–367. Some manuscripts also contain early drafts of the PI §§243–315.[55]

The text refers to the idea of an unsharable and unteacheable language in principle to anyone. The 'language' is supposed to refer to one's immediate personal experiences, known only to its user. The passage is not about a language as a secret code, neither spoken by only one person or in a soliloquy (§243a).

Wittgenstein is quite clear, especially after §§33–35, that ostensive definition, that was demonstrated to be insufficient for word–object relation in general, is even more so insufficient in the case of the description of private experiences. Writing 'S' for a particular sensation would not do the job of description not only for the public but even for the person who has this sensation. The word, or symbol, 'S' would change the meaning over time. The problem is not about remembering what the person meant the other day by 'S', but it is rather in ascribing the 'S' to the same sensation at all.

If the ordinary language fails to express the private sensation and neither there is some private language that can do the job, all that there remains is the use of the language–games in which meaning (private) is expressed by the rules and strategies of the games (public). As long as I can employ the expressions my private sensations within public rules of the likewise public language–games, they are indeed meaningful and not only to myself. This does not invalidate in any way Wittgenstein's consistent belief in the ineffability of semantics. The public language continues to fail to express the meaning, just as the symbol (or word) 'S' fails to express the meaning of a sensation even to its

[55] The full list of every section from PI §243 to §421 in Hark (1990: 19–24).

proprietary. It is not in language that the meaning is expressed but in the language–games.

How do language–games express the meaning of a private sensation? In order to answer this question it is necessary to come back to the introduced notions of the primitive and complete language–games. In fact, in the §244 he refers to the primitive game of pain and its expression. But this is by far not the only one primitive language–game in the argument. It would be appropriate to enumerate the games mentioned by Wittgenstein in the §§243–315. Five passages will be then examined in details.

Before that it would be perhaps helpful to recall what Wittgenstein means by primitive and complete language–games. Primitive games are responsible for establishing the basic link between our experiences and the language, while complete language–games are composed out of the primitive games and modify them. Jaakko and Merrill Hintikka point out that only through the complete language–games "that we bring such notions as knowledge, certainty, evidence, and justification to bear on our talk about mental experiences, for instance sensations".[56]

Sections of the Private Language Argument include many examples of the both kinds of the language–games.

1. Among the primitive language–games there are:
§§244[a1–7], 257a, 288, 311–13 (pain),
§258 (ostensive naming sensation),
§§270a, 278 (physiological games),
§273 (seeing colors),
§275 (seeing objects, pointing),
§277 (visual impressions),
§§285–6, 311–13 (physiognomic games),
§295 (games of non–factual statements),
§299 (saying without knowing or mere perceiving something),

2. Among the complete language–games there are:
§247 (intention),
§249 (lying),
§260 (belief that one behaves in some way),
§261 (verbalizing sensations),
§264 (knowing the meaning/use of the word),
§265 (recollection of some information),
§280 (painting an imaginary picture),
§282a, b (language–games of fairy tales),
§282c (attribution of pain to inanimate objects)

3. There are also combinations of both, primitive and complete (out of the primitive), such as:
§244 [a7] (child sensation/behavior and adults teaching pain–behavior),
§257 ("the stage–setting in the language"),

[56] Hintikka and Hintikka (1986: 345).

§290 (descriptions of mental and physical),
§300 (identity between pain experience and picture of pain experience),

4. Finally, there are also pseudo language–games, or false language–games that Wittgenstein uses quite often to demonstrate some point as his *argumento ad absurdum*. These pseudo language–games are without any meaning or conveying any information. Their rules are meaningless and self–serving, without any strategy in view. In the Private Language Argument there are five such games and their absurdity is to show the irrationality of the idea of a private language:

§257b (naming private sensation by private name),
§267 (justifying the choice of dimensions for an imaginary bridge),
§268 (right hand gives money to the left hand),
§§269–70 (inability to use the meaning one understands),
§271 (unable to make the use of the word 'pain'),

Wittgenstein's complete language–game in BB 17, mirrored in PI §1, and the primitive language–game in PI §2 are, in fact, pseudo–language–games. They are impossible to play: if the point of a language–game is to convey a meaning, then both examples fail to do so. The failure can be stated in two ways. First, and most complicated, is that the rules of the games are incapable of producing any strategies. These rules are strategies themselves, but meaning is transmitted not by rule–following but by strategy–building. Second, more straightforwardly, both pseudo language–games are mere examples of the Augustianian picture of language and of Frege's referential theory of meaning. In them, grammar takes care of itself and in pseudo language–games *rules* take care of themselves and of the game as whole. Both games in PI §1 (BB17) and in PI §2 should be read in the context of the quotation from St. Augustine's *Confessions* in which a theory similar to the TLP and Frege is given.

What distinguishes pseudo from regular language–games is that the former are impossible to play. To know the rules without knowing any strategies would mean not knowing how to play the game. To have no possibility of constructing any strategy from the rules given makes it no game at all.

What are these impossible elements in PI §1 and PI §2? In PI §1d the shopper with the piece of paper and the shop keeper with the color charts and drawers with apples must constantly compare words with objects, objects with numbers and colors in order to make anything work in that game. This is the description of how the referential theory of meaning works. These are the rules without strategies and, indeed, a ridiculous situation of a continuous (if not infinite, like the halting problem in the Turing Machine) comparison of words for that which they stand for. Even if one's walk to the strange grocery store works and both shopper and the store keeper could function for that short scene together by this referential theory of meaning, life in general would be impossible because questions such as "What does it mean...?", or "To what this refers to?", would never stop.

In the PI §2b the situation is even more absurd and the game is slightly shorter in its list of rules that in PI §1d:

> The relation between name and object. Language game of builders. What is the

> relation between names and actions names and shapes? The relation of ostensibly defining. That's to say, in order to establish a name relation we have to establish a technique of use. And we are misled if we think that it is a peculiar process of christening an object that makes a word the word for an object.[57]

This game would be possible if the building of anything would consist in asking and bringing necessary parts. The analogy with language is clear: it cannot consist in a limited number of words in a sequence. At some very near point the builders will encounter the problem not expressible by any of the limited words nor by any ordered sequence (i.e. as in a computer program). At some point we will have rules that are postulated by the meaning and not by the rules themselves, nor are they expressible by the rules of language. Ostensive definition is not enough: we need to establish "a technique of use" and that is the job of a genuine language–game.

In the sections §§243–315, to each game and to each respective paragraph an impressive history of philosophy precedes but also follows the argument to which Wittgenstein offered his own views. Studies have been done to cover both textual and historical analyses of the argument. Space and purpose does not allow reviewing these studies nor permit a detailed analysis of the respective manuscripts. I shall only focus on the game–linguistic part of the argument that, as I believe, is the key for its interpretation. But a comprehensive study from the game–linguistic perspective would have to analyze the respective paragraphs of the games just listed above and for this too, there is not enough space in this work. I shall limit myself to five passages. The first passage (§§244–5) asks the question 'How can word relate to sensation?' The second passage (§257–8) answers that it cannot be by an ostensive definition. The third passage (§265 and §270) also denies that such connection can be done by naming in language. The fourth passage (§282 and §289) looks at the possibility of attributing sensations to unanimated objects. The fifth and final passage (§293) affirms that private sensations are real, and that the connection between word and sensations can only be in a particular language–game which is always public: a public language–game for a private sensation.

3.2 Against Argument

3.2.1. *§§244–45: The Primitive Game of Pain*

After introducing the problem and the question of a private language for a private experience in §243, Wittgenstein continues in the following two paragraphs by asking twice about how words, names and language in general refer to sensations. He provides the most obvious explanation: words refer to sensations in our naming them in everyday use. That which mediates the connection is not, however, naming but a language–game: "This is one way: words are connected with the original, natural expression of sensation and used in its place". After that Wittgenstein immediately provides us with an example of a language–game: "A child has hurt himself, he cries, and now the adults talk to him and teach him exclamations and then sentences. They teach the child new pain–behavior.

[57] Wittgenstein (1993: 448).

'So you say, then, that the word 'pain' really means crying?' – On the contrary; the verbal expression of pain replaces crying and does not describe it".[58]

Notice the question asked by the author's imaginary interlocutor. It is based on the common assumption that we have here two separate events: pain and crying. Wittgenstein rejects this assumption right away. There is no description here (although one can, of course, apply some further description of the event), nor are there two separate events. Being in pain and crying *are* the verbal expressions. But what is their unity based on? It cannot be conceptual because in that case we should be able to return to the synonymous description of one event by two concepts. The unity is in the syntax of the event. Whenever the child is in pain he cries: pain and crying are the constitutive parts of the game, they are the game's defining rules. Any further elaborations of the spontaneous pain–behavior, such as words, gestures, movements, are strategies of the primitive language–game of pain, the elaborations of the rules of the game.

However, 'pain' and 'crying' are not the same fact, for I can have pain without crying and *vice-versa*. Saying that the event is one is not the same as saying that we have the same fact here. The fact of pain and the fact of crying are different, just as the rules of the game and its strategies are different, while the game is one and the same. This distinction between events and facts is fundamental in rejecting any behaviorist interpretations of Wittgenstein. Insisting on the unity of an event, viz. on the unity of the language–game, is important in rejecting any dualistic interpretations of the PI.

Similar passages from Z §527, §568, RFM IV, 50 and VII, 60, and LCA 31 speak of the physiognomic manifestations to different events, which Jaakko and Merrill Hintikka call 'physiognomic language–games'. These are primitive language–games that set up the link between reality and language directly. Pain behavior establishes the public framework that can be shared among the individuals to understand private sensations. It is on this level that such notions as knowledge, information and meaning enter in. But these primitive language–games themselves do not have any epistemic value. In RFM VII, 60 Wittgenstein speaks of the mechanical rule–following in the physiognomic games. Meaning does not occur in them unless we employ these rules in a complete language–game that alone can tell us what sort of meaning the physiognomic expression is supposed to convey. In the end, pain and pleasure can have the exact same physiognomic manifestations and these manifestations by themselves cannot be veridical evidences of either pain or pleasure. One must pass into the complete language–game in order to *know* whether these manifestations are of pain or pleasure.

In this context §245 is of essential importance. It speaks of the separation between pain and the expression of pain and Wittgenstein sees that to be quite problematic: "But how can I then want to use language to get between pain and the expression of pain?" „Wie kann ich denn mit der Sprache noch zwischen die Schmerzäußerung und den Schmerz

[58] „Dies ist eine Möglichkeit: Es werden Worte mit dem ursprünglichen, natürlichen, Ausdruck der Empfindung verbunden und an dessen Stelle gesetzt. Ein Kind hat sich verletzt, es schreit; und nun sprechen ihm die Erwachsenen zu und bringen ihm Ausrufe und später Sätze bei. Sie lehren das Kind ein neues Schmerzbenehmen.

‚So sagst du also, daß das Wort ›Schmerz‹ eigentlich das Schreien bedeute?' – Im Gegenteil; der Wortausdruck des Schmerzes ersetzt das Schreien und beschreibt es nicht". Wittgenstein (2001: §244 a²–c).

treten wollen?“ Anscombe's translation is: “For how can I go so far as to try to use language to get between pain and its expression?” J. Hintikka critisized this translation on the bases of inacurate use of the word *denn* as ‘for’ instead of ‘then’ in the presence of the *noch*, which is often separated from its full form *dennoch*, ‘nevertheless’. This is furthermore supported by *kann* instead of *könnte*, which would render the translation “How can I then”, instead of “For how can I”, in which “then” is omitted.

The difference between these two translations is evident. In Anscombe's interpretative translation Wittgenstein would deny the possibility of separating pain and its expression, making the question–form of §245 to be rather rhetorical. In Hintikka's version, this separation is taken for granted, especially in view of the last sentence from §244c: “the verbal expression of pain replaces crying and does not describe it”. For these reasons, J. Hintikka's translation is used here for the entire text of §245.[59]

Hacker's and Schulte's amended Anscombe translation is even more radical: “How can I even attempt to interpose language between the expression of pain and the pain?”[60] According to this translation any separation sounds somewhat ridiculous. Hacker explains elsewhere that such a separation is impossible indeed.[61] Since Hacker does not have a distinction between primitive and complete language–games, his inseparability thesis of sensation and its behavior would apply *across all* language–games. Anscombe's, Hacker's and Schulte's translations inevitably give a bad name to Wittgenstein as a logical behaviorist for whom the sensation and its behavior are inseparable in principle and not just in primitive games.

In §244 $a^{1–7i}$ Wittgenstein considers *only* primitive games in which these distinctions do not apply. But does it apply in other passages where complete language–games are reported? How else could we separate the real pain from the simulated pain behavior of, i.e. an actor in performance?

These two facts of separation between language and sensation in complete games and lack of such separation in primitive games are very important. The first fact insists on the complexity of our language as such and on the semantic relativism in general. The second is even more important because it sets up the relationship between language and reality but then it indicates that meaning and epistemic features of language should not be confused with that which sets up this relationship. We must look for the employment of such notions in complete language–games instead.

To make matters clear, let us put the two texts together in the way of question and answer. Questions in:

– §244a^1 (‘how do words refer to sensations?’),
– §244a^3 (‘how is the connection between the name and the named set up?’),
– §244a^4 (‘how do we learn the meaning of names for sensations?’),
– §245 (‘how do we use language to distinguish pain and pain–behavior?’)

are answered twice in:

---

[59] Hintikka (1969b).
[60] Wittgenstein (2009: §95^e).
[61] Hacker (1990: 43–45).

– §244 a^{6} ("words are connected with the original, natural expression of sensation and used in its place"),
– §244 b^{2} ("the verbal expression of pain replaces crying and does not describe it").

It does not describe it because there can be no description in a primitive language–game. If sensation is described it is already a complete language–game. The example of a child that is hurt and cries is the example of a primitive language–game. If the child adds a description of his pain (because the adult teaches him to do that) he passes, so to speak, to a more complex game in which the meaning is not only given but explicitly communicated: we pass from mere perception to evidence.

This is why, I believe, Anscombe and Hacker misinterpreted and mistranslated the two sections: they do not make any distinction between sensation/behavior and the naming of sensation and behavior. In the first case sensation is inseparable from behavior (and in this case we can mistake what kind of sensation it is), while in the second case separation is clear by means of explicit language (exclamations and sentences).

Once again, in both cases we do learn something about the sensation of the child. In the first case (§244 a^{7i}) we perceive his sensation in which sensation and behavior are inseparable. In this case language is not used to get between the pain and its expression, because they are one and the same. Indeed, as with any primitive language–game here too we can easily misinterpret the sensation. In the second case (§244 a^{7ii}), when adults teach child exclamations and sentences of the pain–behavior, we come to know about the child's sensation, not just perceive it.

Coming back to the RFM, it must be remembered that the notion of mechanical rule–following can only be implied within the primitive and not complete language–games. Conscious rule–following as conscious actions must also be applied only within the complete language–games.

With this method in mind, we can also return to Wittgenstein's remarks on Turing machines which "are humans who calculate". The main difference between a machine that follows rules blindly and the human intellect is the human ability to create strategy out of the rules of the game. If thinking consists in the ability to follow rules in a certain sequence then we are no different from the machines. If thinking means the ability to create strategies and create new connections out of the old rules, then machines are falling behind. However, as mentioned already, thinking compromises in the performance and here we, humans fall behind even the simplest calculator. Since the Turing machine is a description of one such pseudo language–game where the only activity is that of rule following and is without a single strategy, this game cannot serve even as a metaphor for human intelligence.

3.2.2 *§§257–58: Impossibility of a Private Language(–Game)*

Wittgenstein's earlier rejection of phenomenological language for physical removes the ground for any possibility of a private language. Any solipsistic language would have to be phenomenological for it would need to apply private grammar and context in order to convey the private meaning of one's experiences. On the other hand, if the

only language there is, is public, then developing a private grammar by means of public grammar is a contradiction. In this regard, the Private Language Argument is not just about interpersonal communication necessarily being public but it is rather about the impossibility of any private language in general.

But how certain is this impossibility of developing a private language by means of public grammar? Is it that contradictory? Carnap had spoken about the possibility of creating one's own logic. All that is needed for that private logician is to state clearly the rules of his/her new logical system. This Carnap expresses by his pluralistic Principle of Tolerance: "*In logic there is no morality*. Everyone can make his own logic, that is, his own form of language, as he wants. He only must, if he wants to discuss with us, clearly specify how he wants to do it, give syntactical rules instead of philosophical discussions".[62]

In §257 Wittgenstein shows this by the example of a language–game in which the child invents a name for a sensation. The trouble comes when the child must explain to himself (at least) the meaning of the sensation. Creating a new term would not substitute semantic designation of the sensation. Wittgenstein's objection to this pseudo language–game is the following: "–When you say 'He gave a name to the feeling,' one forgets that a great deal must be prepared in the language so that the mere naming is to make sense. And when we talk about someone giving a name to pain, the grammar of the word 'pain' must be already presupposed; it indicates the pillar towards which the new word is attached".[63]

As we can see, §257 has two language–games. The first is the simple game of inventing a new name for a sensation by the child. The second is the complex game of the great deal of language preparation just for the name to appear. The first game is a pseudo–game; the second is the complete language–game behind the naming of the sensation. The first game is impossible to perform by the criteria set up by the child (to have a private name for a private sensation) because one still applies the public grammar in the allegedly private game.

Making one's own logic or language is only possible if one explains her/his methods of proceeding: first to himself, then to everybody else to whom the new logic or language is presented. That too requires great deal of language preparation. Thus, the new logic or language would not be private, just new in terms of its methods, rules and proceedings.

The "grammar of the word 'pain'" is further explained in §258. In that section Wittgenstein begins by dismissing the possibility of assigning any ostensive definition to the sensation:

> Let us imagine this case. I want to write a diary about recurrence of a certain

---

[62] „*In der Logik gibt es keine Moral.* Jeder mag seine Logik, d. h. seine Sprachform, aufbauen wie er will. Nur muß er, wenn er mit uns diskutieren will, deutlich angeben, wie er es machen will, syntaktische Bestimmungen geben anstatt philosophischer Erörterungen". Carnap (1968: §17).

[63] „– Wenn man sagt ‚Er hat der Empfindung einen Namen gegeben,' so vergißt man, daß schon viel in der Sprache vorbereitet sein muß, damit das bloße Benennen einen Sinn hat. Und wenn wir davon reden, daß einer dem Schmerz einen Namen gibt, so ist die Grammatik des Wortes ‚Schmerz' hier das Vorbereitete; sie zeigt den Posten an, an den das neue Wort gestellt wird". Wittgenstein (2001: §257b).

> sensation. To this end I associate it with the sign 'S' and write it in a calendar for every day when I have this kind of sensation. – I first want to remark that a definition of the sign cannot be expressed (*aussprechen*). – But I can give myself a kind of ostensive definition! – How? Can I point to the sensation? – Not in the usual sense.[64]

The original reads "nicht aussprechen läßt", which in Anscombe's translation is 'cannot be formulated'. This would be, however, a contradiction: if one could write down the sign 'S', then clearly the private sensation *can* be formulated. It is just cannot be expressed either by private language or by any ostensive definition. Thus, the written sign 'S' is a *formulation* that stands for a private sensation but does not *express* anything meaningful.

Wittgenstein here had a choice between '*formulieren*' and '*aussprechen*'. In the PI I, the first is used only once (PI §1), while the second 36 times. Both editions of the PI use 'formulate' and 'express' for 'aussprechen' without making any distinction nor making clear by which criteria the choice was made.[65]

As to why an ostensive definition cannot capture one's private sensation was explained in the previous sections of the PI, starting with his criticisms of the Augustinian/*Tractarian* picture of language. He then proceeds with what is usually considered to be the central argument for refuting the mere notion of private language for a private sensation:

> But I speak or write the sign, while I focus my attention on the sensation – to point out as if it was inside me. – But what is this ceremony? because that is all it seems! A definition surely serves to establish the meaning of a sign. – Well, that happens just by concentrating the attention, because this way I impress on myself the connection between the character and the sensation.[66]

The rule for assigning a private name to private sensation must be private as well. If it is private, then the private linguist him/herself cannot *know* what this rule is because it is not expressed by any language, be it private or public. This passage is parallel to the previous section where the child invents a name for a sensation but cannot explain its meaning even to her/himself. Another way to explain the impassé of the private language is what Wittgenstein is doing in §258a, viz. to say that it is impossible to express

---

[64] „Stellen wir uns diesen Fall vor. Ich will über das Wiederkehren einer gewissen Empfindung ein Tagebuch führen. Dazu assoziiere ich sie mit dem Zeichen ‚E' und schreibe in einem Kalender zu jedem Tag, an dem ich die Empfindung habe, dieses Zeichen. – Ich will zuerst bemerken, daß sich eine Definition des Zeichens nicht aussprechen läßt. – Aber ich kann sie doch mir selbst als eine Art hinweisende Definition geben! – Wie? kann ich auf die Empfindung zeigen? – Nicht im gewöhnlichen Sinne". Wittgenstein (2001: §258a).

[65] This obvious inconsistency in the translation was pointed out by Candlish and Wriskey (2011), but it occurs in Hellett (1977: 339), Hacker (1990: 339), Wittgenstein (2009: §98[e]).

[66] „Aber ich spreche, oder schreibe das Zeichen, und dabei konzentriere ich meine Aufmerksamkeit auf die Empfindung – zeige also gleichsam im Innern auf sie. – Aber wozu diese Zeremonie? denn nur eine solche scheint es zu sein! Eine Definition dient doch dazu, die Bedeutung eines Zeichens festzulegen. – Nun, das geschieht eben durch das Konzentrieren der Aufmerksamkeit; denn dadurch präge ich mir die Verbindung des Zeichens mit der Empfindung ein". Wittgenstein (2001: §258b).

a private rule because the ground for naming is absent. If there is no ground for naming, then the rule, even if it can be expressed, is empty. Empty rules are used in those pseudo language–games Wittgenstein methodologically employs to show the absurdity of some views, i.e. that of private language. The term 'ground' can be exchanged for the Wittgensteinian 'criteria', as in the remaining sentences of this section: "–'I impress it on myself' can only mean: this process brings it about that I remember the connection *correctly* in the future. But in the present case I have no criterion of correctness. One would like to say: correct is whatever appears to me to be correct. And that only means that here we cannot properly speak of 'correct'".[67]

Contrary to the received view, memory here has no importance. For let us assume that I have a perfect memory and I do remember without fail that particular sensation about which I had written 'S' in my diary. But what does this 'S' stand for? In my vocabulary it might stand for the sensation as much as for anything else. The rule, as demanded by Carnap, must be explained as a syntactical rule, and further it must be explained and shown by its application. The sign alone, like an empty rule, has no meaning. Nor has meaning ostension, in which case ostension is just a slightly more linguistically elaborate version of the empty rule.

3.2.3 *§265, §270: The Language–Game of 'S'*

If neither ostensive definition (which is no definition at all) nor empty rule–following can explain the meaning of the private sensation, the only two candidates remaining as semantic designators are either ordinary language or language–games. The first candidate will be eliminated in §293, but already in §270 Wittgenstein suggests language–games as the only semantic designator for private sensations. This is the central point in §270. In it he suggests another physiological language–game in which the rising of blood pressure coincides with someone's having the sensation of 'S'. The reverse is true as well, so much so that whenever I have a sensation of 'S' I have no need of a gauge for I know that my blood pressure is high. Thus, the private sensation of 'S' has a very much public representation in the rising of the blood pressure: "Let us now imagine a use of the registration of the sign 'S' in my diary. I have the following experience: Whenever I have a certain sensation, a pressure gauge shows me that my blood pressure rises. So I am in a position to indicate the rise of my blood pressure without the aid of an apparatus. This is a useful result".[68]

At the time of his service during World War II at Guy's Hospital in London,

---

[67] „– ‚Ich präge sie mir ein' kann doch nur heißen: dieser Vorgang bewirkt, daß ich mich in Zukunft *richtig* an die Verbindung erinnere. Aber in unserm Falle habe ich ja kein Kriterium für die Richtigkeit. Man möchte hier sagen: richtig ist, was immer mir als richtig erscheinen wird. Und das heißt nur, daß hier von ‚richtig' nicht geredet werden kann". Wittgenstein (2001: §258c).

[68] „Denken wir uns nun eine Verwendung des Eintragens des Zeichens ‚E' in mein Tagebuch. Ich mache folgende Erfahrung: Wenn immer ich eine bestimmte Empfindung habe, zeigt mir ein Manometer, daß mein Blutdruck steigt. So werde ich in den Stand gesetzt, ein Steigen meines Blutdrucks ohne Zuhilfenahme eines Apparats anzusagen. Dies ist ein nützliches Ergebnis". Wittgenstein (2001: §270a[1–4]).

Wittgenstein invented a similar devise for recording pulse pressure. The invention must had been made in 1943 or in 1944, shortly before §270 was written. It was used for the treatment of badly injured patients in investigating "the relationship between breathing (depth and rate) and pulse (volume and rate)".[69] The details of that apparatus are not known since the apparatus itself has been lost. But here we can imagine the purpose of the apparatus in measuring multiple particular events by one and the same devise. The result would be similar to an application of the public framework for a private state or an event.

Wittgenstein takes up the question of 'correct' interpretation from §258 and remarks that the right or wrong identification of 'S' is of no importance once we have the public representation of it: "And now it seems to be quite indifferent to know if I recognized the sensation *correctly* or not. Let's say I'm always wrong in identifying it: it makes no difference. And this shows already that the assumption about the error was only an illusion".[70]

What was "only an illusion" is not the sensation itself but the assumption of a possible identification of it by some private notation, ostension or a sign. The sensation is real regardless if I am identifying it right or wrong. The process of verification must not be against the grammar of the private language (which is empty even if real) but against some public association.

Five sections before Wittgenstein rejects the idea that the rules of language can only exist in one's imagination and that the verification process of correctness of use is that between the private word and the likewise private rules of use:

> Let us imagine a table that exists only in our imagination, something like a dictionary. A dictionary can be used to justify the translation of a word X with word Y. Should we also call it a justification if the table is looked up only in the imagination? – 'Well, in that case it is just a subjective justification'. – But one can only provide justification by appealing to an independent body. – 'But I can also appeal from one memory to another. I do not (for example) know if I have correctly remembered the time of the train's departure and to check it I call to mind the image of the page of the train schedule. Don't we have the same case here?' – No; this process has to invoke indeed the *right* memory. If the mental image of the train schedule does not *examine* itself for accuracy, how could it confirm the correctness of the first memory? (As if someone bought several copies of the newspaper this morning to make sure that they write the truth.)[71]

---

[69] Monk (1990: 453).

[70] „Und nun scheint es hier ganz gleichgültig zu sein, ob ich die Empfindung *richtig* wiedererkannt habe oder nicht. Nehmen wir an, ich irre mich beständig bei ihrer Identifizierung, so macht es garnichts. Und das zeigt schon, daß die Annahme dieses Irrtums nur ein Schein war". Wittgenstein (2001: §270a[5–7]).

[71] „Denken wir uns eine Tabelle, die nur in unsrer Vorstellung existiert; etwa ein Wörterbuch. Mittels eines Wörterbuchs kann man die Übersetzung eines Wortes X durch ein Wort Y rechtfertigen. Sollen wir es aber auch eine Rechtfertigung nennen, wenn diese Tabelle nur in der Vorstellung nachgeschlagen wird? – ‚Nun, es ist dann eben eine subjektive Rechtfertigung.' – Aber die Rechtfertigung besteht doch darin, daß man an eine unabhängige Stelle appelliert. – ‚Aber ich kann doch auch von einer Erinnerung an eine andre appellieren. Ich weiß (z.B.) nicht, ob ich mir

The problem with testing against one's memory is that we cannot be sure that our page of the train schedule is correct or not. This process *should* invoke the right memory but there is no way of saying *that* it will be the right memory. And "if the mental image of the train schedule does not *examine* itself for accuracy" we do not have any criterion of justification. In other words, there is no internal criterion of justification but only external ones. The mental image of the train schedule is an internal one.

Again, one's memory might be perfect and the imaginary page might be a perfect copy from the official train schedule (i.e., photographic memory). No matter how perfect one's memory is, the imaginary page is not a copy of the official train schedule; it is ontologically different: the mental image and the page of the schedule. Our mental images are private and always changing, just as any sensation is private and changing. But these private images and sensations cannot be the source of a self–verification process.

J. Hintikka suggests an alternative translation for the German '*prüfen*', '*examine*' which in Asncombe's version is translated as '*tested*'.[72] Hintikka reminds us of the term '*Prüfbarkeitsthese*' used by the logical positivists as the criterion of the meaningfulness of propositions and their verifiability. The last part of §265 confirms the same idea in so many words: "Looking up the table in the imagination does as little for looking up a [real] table as looking up the idea of the result of the imaginary experiment does for the result of a [real] experiment".[73]

Returning to §270 it should be said that Wittgenstein rejected the idea of self–referential images. This needs not, however, undermine the fact that such images are real nor that memory here is of little help. In the background we have the exorcised Cartesian demon whose trick is to deceive us of the certainty of our private sensations. What is more, in the background of these paragraphs is Wittgenstein's rejection of any self–referential *cogito* that alone can prove the correctness of the *res cogitans*. In fact, not even Descartes was able to establish that absolute self–referentiality of the *cogito* without the necessity of postulating a non–deceiving God.

The final paragraph of §270 describes a pseudo language–game in which the self–referentiality of the private images and sensations requires private rules. It also presupposes the use of a private phenomenological language. This game turns out to be impossible and absurd: "(We turned, as it were, to a knob that looked as if one could set something with it on the machine, but it was a mere ornament, not at all connected with the

---

die Abfahrzeit des Zuges richtig gemerkt habe und rufe mir zur Kontrolle das Bild der Seite des Fahrplans ins Gedächtnis. Haben wir hier nicht den gleichen Fall?' – Nein; denn dieser Vorgang muß nun wirklich die *richtige* Erinnerung hervorrufen. Wäre das Vorstellungsbild des Fahrplans nicht selbst auf seine Richtigkeit *zu prüfen*, wie könnte es die Richtigkeit der ersten Erinnerung bestätigen? (Als kaufte Einer mehrere Exemplare der heutigen Morgenzeitung, um sich zu vergewissern, daß sie die Wahrheit schreibt.)" Wittgenstein (2001: §265a).

[72] Hintikka (1969b: 425). Although Hintikka objected to it, this same translation by Anscombe is implemented without change by Hacker and Schulte in Wittgenstein (2009: §100[e]).

[73] „In der Vorstellung eine Tabelle nachschlagen, ist so wenig ein Nachschlagen einer Tabelle, wie die Vorstellung des Ergebnisses eines vorgestellten Experiments das Ergebnis eines Experiments ist". Wittgenstein (2001: §265b).

mechanism)".[74]

In this last sentence in parenthesis we are invited to consider a case in which private images are described by the private rules of this self–referential game. The result is that such image has no connection with anything public. The knob designed for the mechanism turns out to be just an ornament; the image without public reference is an illusion; the game defined by private rules is impossible to play. In this short passage Wittgenstein invites us to imagine private language. Even if it could exist, it would have not only no connection with anything real in the physical world, it would be unintelligible even to its author.

3.2.4 *§282, §289: Sensus ex machina*

Passages from §281 to §289 seem to indicate that only in human behavior can we speak of sensations and of conscious states. This is clear from the §281b: "–It comes to this: only about living human beings, and what is similar to a human being (behaves like) can one say: it has feelings; it can see, is blind; can hear, is deaf; is conscious or unconscious".[75]

The important was clause "*ihm ähnlich ist, (sich ähnlich benimmt)*" sometimes was ignored, which generated interpretations favoring only human beings in being able to exhibit sensations. Other similar creatures would include animals but certainly nothing inanimate, such as computers. From the previous examinations of Wittgenstein's passages from the RFM and RPP I, it should be clear that its author was not partial to computing machines as long as computation was not identified *in toto* with thinking. Computation, as thinking, is a rule–following process. But unlike computation, thinking requires awareness and judgment. At the same time computation is part of thinking and as such it can perfectly be attributed to computers. Or better, to non–human computing machines, since 'computer' is no other than a computing person and Turing machines are humans who calculate.

In §281b Wittgenstein does not specify the creatures who behave like us. But in the following sections we are given examples from the fairy tales with personages who behave just like human beings: they talk, they suffer, they think and they are fully conscious. The point of these remarks is to repeat what he was already saying in §245, viz. that pain and pain behavior can be completely separate facts. A playing child attributes pain behavior to a doll and says "But this use of the concept pain is secondary".[76] When child plays with her dolls and attributes to them pain–sensations, this is an example of the complete language–game which is composed of the primitive language–games in which pain behavior is real. This is clear from the remaining remarks of §282, put in

---

[74] „(Wir drehten, gleichsam, an einem Knopf, der aussah, als könnte man mit ihm etwas an der Maschine einstellen; aber er war ein bloßes Zierat, mit dem Mechanismus garnicht verbunden.)" Wittgenstein (2001: §270b).

[75] „– Es kommt darauf hinaus: man könne nur vom lebenden Menschen, und was ihm ähnlich ist, (sich ähnlich benimmt) sagen, es habe Empfindungen; es sähe; sei blind; höre; sei taub; sei bei Bewußtsein, oder bewußtlos". Wittgenstein (2001: §281b).

[76] „Aber diese Verwendung des Schmerzbegriffs ist eine sekundäre". Wittgenstein (2001: §282c).

parenthesis:

> (If children are playing railroad, their playing is related with their knowledge of the railroad. It could, however, be possible for children of a tribe, to whom the railroad is unknown, to be acquainted with this game from others, and play it without knowing that something is being imitated. Someone could say that the game does not have the same meaning for them as it has for us.)[77]

Children who are not familiar with the notion of trains and the railroad are imitating the behavior of those children who are. For them the game is completely different even though the rules and their manifestations can be exactly the same. An observer can say that these children not knowing anything about the railroad are actually not playing the game; but these observers would be wrong, for all children are playing, just not the same game. What differentiates these games is the secondary use of the objects of play; what unites these games is the fact that their primitive or constituent games are the same.

Let us imagine, visualize that game from §282d. Imagine that the movements of the children who know and who do not know what the railroad is are exactly the same. In their different complete games they all use the same primitive games. This can be reiterated in the way that from the point of view of their complete games they use different strategies but the same rules. One set of rules can, in fact, generate different games under one domain of the game. This may sound as a truism given the fact that, i.e. in the game of chess one has an infinite number of strategies but a limited number of rules, or a limited number of notes can generate an infinite number of melodies, etc. However, the obvious remark sheds some light on how little rule–materials one needs to construct a potentially infinite number of language–games.

There is an important difference between these two activities: applying different strategies to the same rules and following the rules without any strategies. The children in §282c are indeed playing two different games following similar rules. Machines follow the rules without any strategies; the strategies which people ascribe to them are treated by the machines just as a set of rules.

This last point is held in common between Wittgenstein and Peirce in their criticisms of the computational theories of mind. It will be more fully explored in the next chapter, but at this juncture it is important to notice that Peirce compared simple rule–following to formal deduction.[78] On the other hand, strategies imply decisions and finality of actions, which rules, or rule–following, cannot explain.

The important notion of §282 is to make a clear distinction between sensation–behavior and sensation itself. Wittgensteinian scholarship is notorious for ignoring this distinction. So much so that it gave a bad name to Wittgenstein among the AI researchers, neuroscientists and cognitive scientists. Both camps, the Wittgensteinian scholars and the scientists, question whether his philosophy could be used in neuroscience. The

---

[77] „(Wenn Kinder Eisenbahn spielen, hängt ihr Spiel mit ihrer Kenntnis der Eisenbahn zusammen. Es könnten aber Kinder eines Volksstammes, dem die Eisenbahn unbekannt ist, dies Spiel von andern übernommen haben, und es spielen, ohne zu wissen, daß damit etwas nachgeahmt wird. Man könnte sagen, das Spiel habe für sie nicht den gleichen *Sinn* wie für uns.)“ Wittgenstein (2001: §282d).

[78] See Peirce (1887).

answer often was a resounding 'no', unless a serious conceptual reform on the part of neuroscience was to be made.[79]

One of the major sources of inspiration to these works is Peter Hacker's thesis about conceptual confusion within cognitive science in attributing mental states to brain states. This thesis goes by the name of 'mereological fallacy' and can be summed up as follows:

> Although neurological complexity (crudely speaking) is empirically requisite for possession of perceptual, volitional, and cognitive faculties, the kinds of features and the nature of their 'complexity' (if any) that underlie, and constitute criteria for attributing such faculties and their exercise to a being are quite different from this. Psychological concepts are not concepts of ethereal properties or processes, and the presuppositions and conditions of their application concern issues logically independent of neurological complexity, or indeed of the 'computational' complexity or power of a machine.[80]

Hacker's refutation of the mereological fallacy (mental states are *mere* physical states) is based on the logical separation of the mental and physical states. However, as we have seen in the previous chapter, both elimination and separation of the mental and physical events compromise on the possibility of an explanation of either of them. Hacker's attempts to avoid epiphenomenalism come with the price of a strong version of property dualism.

His criticism of mereological fallacy is based on two false assumptions: first, identifying sensations and sensation–behavior (psychological concepts are manifested in human behavior), and second, in attributing epistemic designators, such as knowledge, understanding, meaning, etc. to the primitive language–games (or basic behavior manifestations, since Hacker does not acknowledge any primitive language–games in Wittgenstein).[81]

However, these points are in much mutual unity. Primitive language–games (i.e. sensation of pain and its manifestation) do not use the criteria or justification of use. By criteria Wittgenstein means that which provides an evidence for or justification for being something. This evidence or justification must be provided by experience and serves as the conceptual foundation of knowledge. Criteria are important for the complete language–games because they provide the necessary framework of the game. For instance, facial expressions can be mistaken as pain while they are actually expressing pleasure. By these expressions *alone* one cannot infer *what* they signify. Criteria do the work of such an identification network in the complete language–games. But they are completely irrelevant in the primitive language–games where rules are followed without any reflection, or justification. § 289b speaks precisely about that: "To use a word without justification does not mean to use it unjustly".[82]

The context of §288 shows more evidently the point in §289b: we can be mistaken

---

[79] Holborow (1973); Malcolm (1986); Budd (1989); Klage (1989); Hark (1995); Nadelhoffer (2011).

[80] Hacker (1990: 162). See also Bennett and Hacker (2003: 68–107).

[81] Hacker (1990: 224–253).

[82] „Ein Wort ohne Rechtfertigung gebrauchen, heißt nicht, es zu Unrecht gebrauchen". Wittgenstein (2001: §289b).

about complete language–games but never about the primitive ones. Children might not know what trains are or what a railroad is but they can join in the game of moving toys of those children who know the entire game at play. If epistemic designators can only refer to the complete language–games then the error of understanding, interpreting, and the like belong to the complete language–games as well. Consequently, if such epistemic designators do not belong to the primitive language–games then a mistake is not possible in them. This simply means that, i.e. I can be mistaken or being deceived whether the other person is in pain, but I can never be wrong about being in pain myself. This is exactly what §289a says (to get the complete point of §289, I will add its second part again): "When I say 'I'm in pain', I am certainly justified *before myself*". – What does that mean? Does it say: "If someone else might know what I call 'pain', would he admit that I use the word correctly"? To use a word without justification does not mean to use it unjustly".[83]

This is how two mistakes come together: if we attribute knowledge and understanding to every experience then we are implying also that sensations and their manifestations must come together. The same mistakes were pointed out in the case of Searle's identification of intentionality and reference, or as it was called, 'intentionality of knowledge' and 'intentionality of perception'. The suggested solution then was 'linguistic intentionality' where a neural basis was interpreted as the rule–foundation for cognitive activity. While neural brain activity requires no epistemic designators they are required by that to which the activity is directed: goals, intentions, expectations, etc.

Perhaps Hacker's own fallacy of attributing mental states to behavioral states comes as the result of a more basic misinterpretation of the nature of the language–games. It must be pointed out again that language–games are larger than any behavior they manifest. This is why we can speak of language–games as the links between language and reality without any fear of a mistake. Mistakes are possible only in complete language–games where epistemic designators are present and where different interpretations are possible. In the primitive language–games, where reality and language come together, such mistakes are not present and are not even possible.[84]

In the light of these distinctions it must be said clearly that when Wittgenstein insists on not attributing cognition to brain activity he is in fact applying these very distinctions and not insinuating that the brain has nothing to do with the mind.[85] It is rather the mistake of attributing too much to the rules of the game and to the primitive games themselves, and attributing too little to the unity between rules and strategies, between primitive and complete games. Rules without strategies make no sense and strategies without rules are unthinkable. Primitive games are constitutive to the complete games

---

[83] „'Wenn ich sage ›Ich habe Schmerzen‹, bin ich jedenfalls *vor mir selbst* gerechtfertigt.' – Was heißt das? Heißt es: ‚Wenn ein Anderer wissen könnte, was ich ›Schmerzen‹ nenne, würde er zugeben, daß ich das Wort richtig verwende?'

Ein Wort ohne Rechtfertigung gebrauchen, heißt nicht, es zu Unrecht gebrauchen". Wittgenstein (2001: §289 a–b).

[84] Except in the cases of i.e., hallucinations, but in these cases, we deal with the oddity of one language–game having two sets of rules (real and imagined). Instead of my behavior (strategies) following the real sets of rules, it follows the imaginary ones.

[85] Wittgenstein (1980: §903ff); Wittgenstein (1969: §§608ff).

without which the later would not be possible. Cognition is that set of strategies composed of the neural activity of the brain defining the rules of cognition.

In short, cognition cannot be just about rules or just about strategies. Physicalism in its most general description was rejected by Wittgenstein at the same period when he realized that the primacy of rules over the strategies in the game was his own misconception. This transition occurred sometime between the BB and the PI I, which is clear from his discussions on rule–following in PI §§185–242. Since §§243–315 were written much later, the passage on private experiences presupposes the primacy of the language–games over their rules and of the use over rule–following.

These remarks are in harmony with Wittgenstein's understanding of intentionality and his resistance against an internalist as well as externalist understanding of it. Intentional states, unlike sensations, are not processes with a beginning, some timely duration, and an end. I can say that I began to feel pain this morning but it stopped after taking an analgesic; I cannot say the same about having a thought, beginning to understand a concept from time t_1 to time t_2.

Instead, Wittgenstein insists that our intentional states and consciousness are imbedded in the grammar of our lives, in the rules of our uses of words and concepts. Unlike sensations, intentional states do not have some logically independent status from the way they are expressed.

This treatment of intentionality is closely related to Wittgenstein's general notion of intentionality expressible in the grammar of language. By demythologizing Brentano's intentionality Wittgenstein shows us that intentionality of knowledge entirely depends on the intelligibility of the grammatical structures of our languages. Sentences with broken grammar fail to generate not only meaning but also intentionality. This works the other way around: the knowledge of an object as something, the feeling directed toward some event must be semantically intelligible and expressible in grammar: "What we cannot think, that we cannot think; thus we also cannot *say*, what we cannot think".[86] And what we cannot say we cannot know either.

But notice that his, what I called in the previous chapter 'linguistic' or 'grammatical' intentionality, is radically different from the reductive treatment of intentionality and grammar by Chomsky. Chomsky reduces meaning and intentionality to the grammar of our ordinary language; Wittgenstein shows through grammatical analysis the workings of both meaning and intentionality. Furthermore, in PI I 'grammar' is extended to the rules of our practices, thus becoming the rules of the language–games. There grammar is not the mere syntactical rules of our language but also the rules of the language–games. Ordinary language is not enough.

3.2.5 *§293: The Little Worlds of the Private Beetles*

When and if one does speak of his/her private sensations, what becomes of our understanding? So far from Wittgenstein's discussion it appears to be clear that private

[86] „Was wir nicht denken können, das können wir nicht denken; wir können also auch nicht *sagen*, was wir nicht denken können". Wittgenstein (1974a: 5.61).

language is self–contradictory and that ostensive definitions do not express meaning. But what about ordinary language; can one rely on a simple linguistic description of one's sensations, feelings and experiences? Is it not a complete language–game, an elaborated primitive one? Perhaps the best well–known passage from the PI I, the beetle–in–the–box thought experiment, answers precisely that question: "If I say of myself that I know only from my own case, what the word 'pain' means, – must I not say *the same* of the others? And how can I generalize that *one* case in such an irresponsible manner?"[87]

In §293a Wittgenstein invites us to imagine the case in which we presume no public framework of how others assign the meaning of their private sensations. Certainly not having such a framework would not cancel out the realism of one's private sensations. Nor does it cancel out the language in which words stand for sensations. The point in §293 is not really about language in general but its ability to convey meaning. In the end, if this thought experiment would require us to cancel out all talk of sensation from our language, what sort of language would that be? The main point here is rather to examine whether the public, ordinary language of private sensations is enough.

The thought experiment invites us to imagine a language in which the direct relationship between 'object and designation' ('*Gegenstand und Bezeichnung*') is established. Indeed, it is a common–sense understanding how meaning of the words is conveyed by the use of words alone, by correct grammatical constructions of our everyday language. Wittgenstein insists, however, that the language of 'object and designation' is unsuitable for communicating one's private experiences, for when it attempts to make such communication, the private sensation not only remains hidden as a beetle in a box but inexpressible in principle:

> Now, everyone tells me about himself, that he knew only from himself, what pain is! – Assume that each had a box, with something in it, which we call "beetle". No one can ever gaze into the box of the another, and everyone says he knows only by looking at *his* beetle, what a beetle is. – In such a situation it would be possible that each would have another thing in his box. One could imagine that such a thing is constantly changing. – But what if the word "beetle" of these people still would have a use? – If so, it would not be the name of a thing. The thing in the box does not belong in the language game at all, not even as a *something*: for the box might even be empty. – No, [one] can 'divide through' this thing in the box; it cancels out, whatever it is.[88]

---

[87] „Wenn ich von mir selbst sage, ich wisse nur vom eigenen Fall, was das Wort ‚Schmerz' bedeutet, – muß ich *das* nicht auch von den Andern sagen? Und wie kann ich denn den *einen* Fall in so unverantwortlicher Weise verallgemeinern?" Wittgenstein (2001: §293a).

[88] „Nun, ein Jeder sagt es mir von sich, er wisse nur von sich selbst, was Schmerzen seien! – Angenommen, es hätte Jeder eine Schachtel, darin wäre etwas, was wir ‚Käfer' nennen. Niemand kann je in die Schachtel des Andern schaun; und Jeder sagt, er wisse nur vom Anblick *seines* Käfers, was ein Käfer ist. – Da könnte es ja sein, daß Jeder ein anderes Ding in seiner Schachtel hätte. Ja, man könnte sich vorstellen, daß sich ein solches Ding fortwährend veränderte. – Aber wenn nun das Wort ‚Käfer' dieser Leute doch einen Gebrauch hätte? – So wäre er nicht der der Bezeichnung eines Dings. Das Ding in der Schachtel gehört überhaupt nicht zum Sprachspiel; auch nicht einmal als ein *Etwas*: denn die Schachtel könnte auch leer sein. – Nein, durch dieses Ding in der Schachtel kann ›gekürzt werden‹; es hebt sich weg, was immer es ist". Wittgenstein (2001: §293b).

The beetle is real if and only if it has place in the language–game and if it does not, then the box is empty. Unlike in idealism, strategies do not exist independently from their rules. If strategies are rightly identified with their rules, by inference from the strategy one can know the rules. One can make a false inference but it would still be a valid one (i.e. I might infer a wrong bug in the box). If the strategies are intelligible they are tight to some sets of rules. Then the box cannot be empty. If, on the other hand, all that is manifested in behavior is blind rule–following, the box must be empty and we have no strategies to even begin our inference with.

Likewise, looking at another person's box I do not know whether the box contains the beetle or whether it is empty. The thought experiment is designed to give us the only possible tool to know whether there is a beetle in the box. The linguistic description utilizing the model of '*Gegenstand und Bezeichnung*' will not do the job of testing the contents of the box. The person may be lying, that is, may be playing a whole different language–game from what I may be expecting, and his/her mere words resound to me as rules of at least two possible games. Language alone cannot tell me which game is being played.

If ordinary language is the only method we can rely on, then we might very well accept the idea of the private 'language' while at the same time give up hope of meaning acquisition. However, such private 'language' would lack semantics while retaining some strange relation between word and designation of grammar in a context. Notions of innate grammar and Language of Thought come to mind when one tries to imagine such private 'language'.[89]

In §243 Wittgenstein envisioned a private language, i.e. a language spoken only by one person. The reason why it is not possible for him, however, is that such private language would have no public access, that it would not be understandable to anyone else, not even to the 'private linguist'.

Just as with the formal logical system, I can invent an entire language with a completely new grammar, vocabulary, alphabet or characters, with new unheard before phonetics. All that (and more) is possible, but if this is indeed a language, it must be accessible, learnable *in principle* by others. And if the 'private linguist' is capable of using it, then it is indeed a publicly accessible language.

The wrong model of the name–object relationship is explained very clearly in §293c: "It means: If one constructs the grammar of the expression of sensation on the model of 'object and designation', then the object falls out of consideration as irrelevant".[90] It is important to repeat that the '*Gegenstand und Bezeichnung*' model not only extends to ostensive definitions (as Wittgenstein already explained in §256), but also to the ordinary language as a whole. In the end, ostensive and descriptive languages differ more in degree than in kind. They are both describing, picturing, referring to objects and sensations. Ultimately it is of little difference if the person holds a box and says: "Beetle", or holds the box and goes on describing in detail its contents. In §293c Wittgenstein not only

[89] Fodor, however, denies the application of Wittgenstein private language argument to his Language of Thought theory. Fodor (1975: 68–73).

[90] „Das heißt: Wenn man die Grammatik des Ausdrucks der Empfindung nach dem Muster von ‚Gegenstand und Bezeichnung' konstruiert, dann fällt der Gegenstand als irrelevant aus der Betrachtung heraus". Wittgenstein (2001: §293c).

dismisses ostensive definitions as irrelevant to our knowing about the beetle but also limits the possibilities of the ordinary language to convey such knowledge to us. The only candidate remaining for the possibility of meaning to be intelligible and public is the language–game.

3.3 Against Mental Causation

The following passages from *Zettel*, one of the last note collections written by Wittgenstein between 1945 and 1948, often have been cited to show Wittgenstein's disavowal of psycho–physical causation and to demonstrate his alliance with behaviorism. What the passage says with all clarity is that type–token identity theory cannot account for psychological events, and that there is not any strict identity between brain states and mental states.[91]

The following passage, Z §§608–611, must be read within the larger context of *Zettel* in which Wittgenstein discusses two topics not at all novel to his philosophical treatment: rule–following and thinking as calculation. In Chapter II it was already explained that mere rule–following cannot be accountable for thinking. Computing, calculating, writing, erasing, copying, etc. can be considered as mechanical rule–following and indeed, are compatible with the computer's functions (i.e. Turing Machine's functions). But thinking is more than computation and more than mere rule–following. It requires strategies, decisions, goal–making and goal–achieving actions. These cannot be reduced to or accounted for solely by rules.

Wittgenstein proceeds to show that if we consider brain states as a very complex set of rules, even if taken together, they cannot be identified with the mental states of thinking:

> No supposition seems to me more natural than that there is no process in the brain correlated to association (*Assoziieren*), or thinking, so much so that it would be impossible to read from the brain processes thought processes. I mean this: When I speak, or write, there is, I suppose, a correlated (*zugeordnetes*) system of impulses coming from my brain with my spoken or written thoughts. But why should the *system* continue further in the direction of the center? Why should not, as it were, this order originate out of chaos?[92]

---

[91] Compare with other less, however, straightforward criticisms of physicalism and dualism in the earlier writings: Wittgenstein (1958: 47), (against dualism and computationalism); Wittgenstein (1958: 118–121), (against computationalism: reading as rule–following vs. reading as "a par-ticular conscious mental act"); Wittgenstein (1980: §1063), (against physicalism); Wittgenstein (2001: §§149–150; §§153–154; §158), (against dualism and physicalism being foremost grammati-cal differences).

[92] „Keine Annahme scheint mir natürlicher, als daß dem Assoziieren, oder Denken, kein Prozeß im Gehirn zugeordnet ist; so zwar, daß es also unmöglich wäre[(1)], aus Gehirnprozessen Denkprozesse abzulesen. Ich meine das so: Wenn ich rede, oder schreibe, so geht, nehme ich an, ein meinem gesprochenen oder geschriebenen Gedanken zugeordnetes System von Impulsen von meinem Gehirn aus. Aber warum sollte[2] das *System* sich weiter in zentraler Richtung fortsetzen?

Wittgenstein "supposes" that "a system of impulses" in the brain is "correlated" with his thoughts. It is important to keep in mind that the importance of the rules for the strategies was never undermined by Wittgenstein. This criticism of physicalism is not his denial of the indispensible role that brain states have for our thinking. But thinking *is not* the set of the brain states.

The word '*zugeordnetes*' can be translated as 'correlated'. The meaning here is that of a continuous action of a directed relation between the brain states and the mental states. Correlation does not mean relationship of causation in which one state implies or follows the other. This is why Wittgenstein, among other reasons that will be explained at the end of this chapter, refused the notion of identity either as sameness (tautology) or as implication (nonsense). However, causation implies some notion of identity.

Of course, there is no talk of any identity theory here, quite the contrary. The brain impulses are neither causally associated nor identified with our thoughts and speech, but correlated with these, just as rules and strategies are correlated and organized into one indivisible language–game.

The rest of the paragraph demonstrates the above point by an analogy to the seed of a plant and the plant itself. Again, here the idea of association must be read in two possible ways: the seed completely determines what the plant will be in every aspect (then by examining the seed we could picture the structure and the properties of the future plant to some exactness), and the correlation between the seed and the plant in terms of the sameness of species. Hence, the idea of association is either that of determination or that of correlation. Wittgenstein denies the first and affirms the second:

> The case would be similar to this, – that certain kinds of plants reproduced by a seed, so that a seed always produces the same species of plant from which it was created, – but *nothing* in the seed of the plant from which it comes, corresponds with the seed; it is therefore impossible to conclude from the seed, from its properties or from its structure, for this can be done only from the *history* of the seed. So an organism could be generated from something quite amorphous, uncaused so to speak, and there is no reason why the same should not be true with our thoughts, our words, or writing, etc.[93]

The two terms: "the same species" and "properties or structure" are opposed in identifying the seed and the plant grown from the seed. Here Wittgenstein does not imply that out of the lemon seed an apple tree might grow. The kind of the plant corresponds to the kind of the seed, while properties and structure of the plant does not correspond

---

Warum soll[3] nicht, sozusagen, diese Ordnung aus dem Chaos entspringen?" Wittgenstein (1969: §608a).

[93] „Der Fall wäre ähnlich dem – daß sich gewisse Pflanzenarten durch Samen vermehrten, so daß ein Same immer dieselbe Pflanzenart erzeugt, von der er erzeugt wurde, – daß aber *nichts* in dem Samen der Pflanze, die aus ihm wird, entspricht; so daß es unmöglich ist[(4)], aus den Eigenschaften, oder der Struktur des Samens auf die der Pflanze, die aus ihm wird, zu schließen, – daß man dies nur aus seiner *Geschichte* tun kann[5]. So könnte[6] also aus etwas ganz Amorphem ein Organismus, sozusagen ursachelos, werden; und es ist kein Grund, warum sich dies nicht mit unserem Gedanken, also mit unserem Reden oder Schreiben etc. wirklich so verhalten sollte[7]". Wittgenstein (1969: §608b).

to the seed. First, because these properties are not in the seed, second, we need to look into the history of the seed, viz. into its genetics. Similarly, the tone of the suggestion that some organism might emerge from something amorphous is derisive of the identity between the seed and the plant, brain processes and thoughts. The conclusion, therefore, is clear: "It is therefore quite possible that certain psychological phenomena *cannot* be investigated physiologically, because physiologically nothing corresponds to them".[94]

A comparison to the third dimension of the language–games, to the family resemblances and forms of life, can be easily made in view of the second part of the last sentence. Our thoughts, speech and writing depend more on our own history than on the brain states that make these acts possible. This is why Wittgenstein denies the brain–mind causation. Mental acts cannot be caused by the brain states because they are necessary but not sufficient for these mental acts. Peirce adds the importance of the way our thoughts are expressed (see the inkstand vs. the brain lobe example discussed in the next chapter, CP 7.366), but Wittgenstein seems to touch here an even more important aspect of our thinking: its genealogy from different language–games and different forms of life, compared to the genetics of the seed as being more relevant to the structure of the plant than its physical composition.

In the next section Wittgenstein states his anti–causality position clearly. Without one–to–one correspondence, even without his associative isomorphism of the *Bildtheorie,* any talk of the brain causing minds is meaningless:

> I've seen this man a few years ago, and now I see him again, recognize him, remember his name. And why must there be a cause for this remembering in my nervous system? Why does anything whatsoever have to be stored up *in some form* in there? Why must it have left a trace? Why should there be any psychological regularity that corresponds to *any* physiological [one]? If that overturns our notions of causality, then it is time that they [should] be overturned.[95]

The order of rules and strategies cannot account for the correspondence of psychological and physiological regularities. Again, Wittgenstein is not saying that my recognition of the man and the remembering of his name does not involve the brain state of a particular kind. But he is saying that the stimulation of the brain area that reacts to the face of that man would not produce the memory of his face and name. This alone is not enough. Language–games are more than the working of the brain states, because their strategies are always goal–oriented in solving problems that are posed outside of the self. Causality demands a closer relationship between brain and mind than that which Wittgenstein allows and it is not surprising that causality is rejected.

---

[94] „Es ist also wohl möglich[8], daß gewisse psychologische Phänomene physiologisch nicht untersucht werden *können*[9], weil ihnen physiologisch nichts entspricht". Wittgenstein (1969: §609).

[95] „Ich habe diesen Mann vor Jahren gesechen; nun sehe ich ihn wieder, erkenne ihn, erinnere mich seines Namens. Und warum muß[10] es nun für dies Erinnern eine Ursache in meinem Nervensystem geben? Warum muß[11] irgend etwas, was immer, *in irgendeiner Form* dort aufgespeichert worden sein? Warum *muß*[12] er eine Spur hinterlassen haben? Warum soll[13] es keine psychologische Gesetzmäßigkeit geben, der *keine* physiologische entspricht? Wenn das unsere Begriffe von der Kausalität umstößt, dann ist es Zeit, daß sie umgestoßen werden". Wittgenstein (1969: §610).

At this point it is very important to refer to the original German text that uses in these three selected sections (§§608 – 611) four modal verbs thirteen times. Two of them ("*können*" in §609 and "*müßen*" in §610 are emphasized by the author). I underline these verbs (and two modal formations of "*umöglich wäre*$^{(1)}$" in 608a "*unmöglich ist*$^{(4)}$" in 608b) to show their importance to the style of these sections. This play with the modal verbs by Wittgenstein indicates this message: physicalism tells us that there *should* and *must* be some identity between brain and mental states, but it *cannot* be the case, for there *can* be *no* identity of logical or mathematical equality between one and the other except the well–orderedness and correlation of rules and strategies.

If reductive physicalism is rejected, then all we have remaining is dualism. Likewise, the well–orderedness of rules and strategies must be something similar to preestablished harmony or parallelism between the two. The next section shows the ridiculousness of such choice between physicalism and dualism as the only possibilities of treating human cognition: "The prejudice in favor of psychophysical parallelism is a fruit of primitive views of our concepts. For if you allow causality between psychological phenomena that is not mediated physiologically, you imply conceding to the existence of a soul *apart from* the body, a ghostly soul-creature".[96]

The "primitive views of our concepts" ("primitive interpretations" in Anscombe's translation) is an echo of Wittgenstein's concept of philosophy as a therapy of language. Dualism allows for such primitive interpretation, just as preestablished harmony allows the identity of indiscernibles. But so does reductive physicalism. The choice between these two is as false as the profession of one or the other. Rejecting primitive interpretations for Wittgenstein makes it impossible for him to side with either of the two most common positions in philosophy.

4. Psychoanalysis through the Language–Games

If ostensive definition and ordinary language are ousted as candidates for defining the world–language relationship, we seem to have only two possibilities left to choose from: solipsism or public framework. Strong solipsism was denied by Wittgenstein in the TLP for its self–defeating reasons. A weak form of solipsism was adopted in the TLP but was given up later at the same time when the idea of phenomenological language was abandoned. Solipsism of any form would be an unsustainable idea when the notion of a private language is radically denied.

Language–games are this public framework, understood as rule–guided human activities being the main representative link between language and reality.[97] If his earlier picture theory is that same very link then it employs ordinary language. Wittgenstein refutes that idea in the later sections of the Private Language Argument and insists that if

[96] „Das Vorurteil zugunsten des psycho–physischen Parallelismus ist eine Frucht primitiver Auffassungen unserer Begriffe. Denn wenn man Kausalität zwischen psychologischen Erscheinungen zuläßt, die nicht physiologisch vermittelt ist, so meint man damit ein Zugestehen, es existiere eine Seele *neben* dem Körper, ein geisterhaftes Seelenwesen". Wittgenstein (1969: §611).

[97] Hintikka (1973: 63–66); Hintikka (1969a: 3–19).

we use picturing in describing our sensations then no information, no facts are actually conveyed (PI §§295–298). Picture theory is useful only when pictures are composed of many interrelated language–games.[98] Picture theory is then somewhat synonymous to the family resemblance. One family of language–games is one picture of some part of reality. In this way, two semantic 'theories' of Wittgenstein, picture theory of language and language–games, or 'meaning as use', come together in perfect harmony by the time PI I was completed.

But what is being pictured by the language–games? What is resembled by the family of interrelated language–games? Certainly these are sensations, objects, facts, thoughts, concepts and ideas, but of what kind, of what nature?

At the beginning of this chapter it was indicated that objects of the TLP are phenomenological and that they come under the description of phenomenological language. By the early 1930s Wittgenstein gives up the second but retains the first. Now, physical language describes phenomenological objects such as sense–data of sensations, colors and objects of acquaintance. Introduction of the language–games only reinforced Wittgenstein's use of the physical language. But did it treat in any way the existence of phenomenological objects?

Not in the first part of the PI, or at least not until §§243–315 were introduced. As was mentioned, these sections were the last of the Part I written in 1944. Between 1945 and his death in 1951 Wittgenstein changes his attitude toward phenomenological objects, especially that of the sense–data.[99]

Merrill and Jaakko Hintikka insist that Wittgenstein never really gave up his phenomenological objects and consistently used the notion of the sense–data.[100] The alleged use of the sense–data in Wittgenstein's later epistemology would also confirm Wittgensteinian essentially Cartesian world–view of private sensations even if expressed only in public language.[101] Paul Snowdon explains that Wittgenstein opposed the notion of private sense–data but retained public or general sense–data. This would be, according to Snowdon, the way out from skepticism.[102] Examining Wittgenstein's last writings, however,

---

[98] Stenius (1967).

[99] According to Cora Diamond's printed material from Margaret Macdonald's notes on Wittgenstein's lectures this change occurred already in 1935. See the "Appendix" to Diamond and Gerrard (1999). Notes on Wittgenstein's lecture from 1935–1936 are on pages 130–134.

[100] "For the Wittgenstein of the *Philosophical Investigations* no less than for the Wittgenstein lecturing in Cambridge in 1930–1932, the world we live in is the world of sense–data. His problem is to show how I can speak of them in the only language I understand, which is the language of phys-ical objects", in Hintikka and Hintikka (1986: 251).

[101] "Wittgenstein's statement [PI II, xi, 217] does not belie his analysis of sensations in which he did not depart radically from Descartes as far as metaphysics is concerned, only as far as semantics is concerned. According to Wittgenstein's rightly understood view, if God had looked into your mind he would have been able to see there what you sense or feel. Now it can be seen, however, that Wittgenstein's closet Cartesianism does not extend from sensations and other episodic private experiences to thinking or to propositional attitudes. As far as propositional attitudes are concerned, Wittgenstein is after all an anti–Cartesian". Hintikka and Hintikka (1986: 292–293).

[102] Snowdon (2011).

tells a different story. Eliminating sense–data completely is a much more radical remedy against skepticism than taking on the private sense–data.

The old Kantian problem of a mismatch between phenomenological objects and physical language is solved by means of language–games. The (ordinary) physical language for phenomenological objects is indistinguishable from that which is used for physical objects. I will use the exact same vocabulary and grammar if I speak of the Lost Island as much as if I speak of, i.e., Ireland. But the same is impossible if our physical language includes language–games, because unlike in a simple 'object and designation' model of description, or picturing, in the physical language–game language, or model, I must necessarily implement within my actions the use of the term. While I can do that in the case of the name 'Ireland', i.e. by planning my trip to it, buying ticket to it, calling somebody living in it, I cannot do the same with the 'Lost Island'. The 'Lost Island' has no place in that language–game at all; not even as *something*: it might not even exist. In fact, outside of the minds of Gaunilo and all of us who have been imagining the Lost Island with him since the 11th century, *it does not* exist.

At this point further elaboration on Anselm's ontological argument and Gaunilo's objection of the Lost Island might be in order. For Wittgenstein undoubtedly stands on the side of the great prior of Bec. As we know from replies of Anselm to Gaunilo's objections, Gaunilo confused the argument for a thought experiment. By thinking about any concept, existent or not, one cannot prove the existence of it. But Anslem never implied anything otherwise. The difference between God and the Lost Island is in the place they both occupy in different language–games. The Lost Island exists only within the limited 'language–game' of speaking or thinking about it. On the other hand, "a being than which a greater cannot be conceived", is part of the multitude of family–resembled language–games and forms of life of those who not only think but live their lives according to what "a being than which a greater cannot be conceived" represents.

The argument (and not the thought experiment as that of the Lost Island) from *Proslogion* was designed so that the quality of the prayer life of the monks of the priory of Bec would improve. This is clear from the opening lines of *Proslogion*.[103] The activity of thinking about "a being than which a greater cannot be conceived" is *only part* of the larger activities of the monks in their everyday service to God through work and prayer. The *only* activity that comes from thinking and speaking of the Lost Island is that same activity of thinking and speaking about it.

I suspect that Wittgenstein would not agree on considering our discourse on the Lost Island as some language–game, given the fact that the concept does not correspond to anything real in the world. In this he was a follower of Frege, for whom these concepts would not even belong to the third (Platonic) world. This third world is for the concepts that are referential to the real things in the first world, including numbers. For some, the discourse on the Lost Island can be part of the language–game as long as the rules of such discourse are specified clearly.[104] Here, however, language–game would not be considered as the representative link between language and reality due to the lack of 'reality' of the Lost Island outside of thought and language to begin with:

---

[103] Anselm of Canterbury (1969: 110–112).

[104] See Chakrabarti (1997).

> The child does not learn that there are books, that there are chairs, etc. etc., but he is learning to fetch books, (to) sit on chairs, etc.
>
> Questions about existence certainly also come later: "Is there a unicorn?" etc. But such a question is only possible because as the rule such questions have no appropriate occurrence. For how do you know, how you have to convince yourself of the existence of unicorns? How do we learn the method for determining whether something exists or not?[105]

The concept of 'the language–game of the Lost Island' is a mixture of Wittgenstein's later philosophy in which language–games are the sole link between language and reality and his earlier philosophy in which phenomenological objects (i.e., the Lost Island) and language about them were admissible. It is, however, no coincidence that phenomenological language was abandoned in October 1929, while the very first record of the notion of the language–game comes in the summer of 1930.

There was, however, one mistake on the part of Anselm to consider the whole argument as a matter of objective certainty while the only certainty one can have here is a subjective one. The conceptual necessity is the objective certainty, denying which would necessarily imply denying the entire grammatical foundation of our language, which is a contradiction. As long as God is not a conceptual necessity and thus not an objective certainty, the Fool will always have sound reasons to object in reply.[106]

Unlike Anselm, Wittgenstein makes a distinction between objective and subjective certainties clear and associates religious belief with subjective certainty. The necessity, therefore, is bound to the grammar of one's language–games and their families, not to the grammar of our language. This is why "a being than which a greater cannot be conceived" rooted in our actions, practices and cultures is always going to be different than the Lost Island rooted only in linguistic discourse and fictional imagery about itself. This difference is quintessential for Wittgenstein's rich philosophy of religion, ethics and aesthetics.

Our religious belief does have a solid foundation in our forms of life, while non–

---

[105] „Das Kind lernt nicht, daß es Bücher gibt, daß es Sessel gibt, etc. etc., sondern es lernt Bücher holen, sich auf Sessel (zu) setzen, etc.

Es kommen freilich später auch Fragen nach der Existenz auf: ‚Gibt es ein Einhorn?' usw. Aber so eine Frage ist nur möglich, weil in der Regel keine ihr entsprechende auftritt. Denn wie weiß man, wie man sich von der Existenz des Einhorns zu überzeugen hat? Wie hat man die Methode gelernt zu bestimmen, ob etwas existiere oder nicht?" Wittgenstein (1969: §476).

This text was written on 6 April 1951, three weeks prior to his death. Compare the contrast of B. Russell's recollections of Wittgenstein at Cambrige in February of 1911: "My German engineer very argumentative & tiresome. He wouldn't admit that it was certain that there was not a rhinoceros in the room ... [He] came back and argued all the time I was dressing. [1. 2. 11]. My German engineer, I think, is a fool. He thinks nothing empirical is knowable – I asked him to admit that there was not a rhinoceros in the room, but he wouldn't. [2.11.11]" Monk (1990: 39).

[106] "It seems to me that in so far as one can build a satisfactory theory of (conceptual) necessity, it will be in the relevant respects sufficiently similar to the logic of knowledge to enable us to say essentially the same things about our chances of reconstructing the ontological argument in terms of ordinary modal logic as we already said about these chances in epistemic logic (the logic of knowledge). Gaunilo, Aquinas, and Kant thus appear to have been shrewder – or perhaps merely sounder – logicians than St. Anselm and Descartes". Hintikka (1969a: 52).

circularly these forms of life are referential to "a being than which a greater cannot be conceived". Again, while the Lost Island can only be confirmed through and by our language, belief in "a being than which a greater cannot be conceived" is confirmed by the language–games (activities) of our faith. (Hence Anselm's appeal to Gaunilo's faith at the beginning of his reply).[107]

Objective certainty is bound with the language–games and with grammar of our language. This is a kind of certainty that cannot be put to doubt, for doubting it would doubt the ground of language in which such doubt is formulated:

> Why am I so sure that this is my hand? Isn't the whole language–game based on this kind of certainty?
>
> Or: isn't this 'certainty' (already) put before in the language–game? Namely in this way, that *who* does not play, or plays it incorrectly, is the one who does not recognize the object with certainty.[108]

But certainty with knowledge and belief come only in a complete language–game. Primitive language–games are knowledge–neutral: "On the basis of well–founded belief lies the unfounded belief".[109] They are still in need of interpretation and of the context within some complete language–games. There is always a choice which interpretation to take, viz. in which complete language–game that primitive one should be used.

This last point was already made in the previous chapters, when a necessary distinction was made between an action and the consciousness of it. Consciousness of, i.e., movement, comes after the movement was performed.

Objective certainty makes inquiries about the existence of objects not only meaningless but also such inquiries undermine the foundations of one's language and knowledge: "'I cannot doubt this proposition without giving up all judgment.' But what proposition is that? (It is reminiscent of what Frege said about the law of identity[110]). It is certainly not an empirical proposition. It does not belong to psychology. It is more like a rule".[111]

---

[107] Anselm of Canterbury (1969: 168–172).

[108] „Warum bin ich denn so sicher, daß das meine Hand ist? Beruht nicht auf dieser Art Sicherheit das ganze Sprachspiel?

Oder: Ist in dem Sprachspiel diese ‚Sicherheit' nicht (schon) vorausgesetzt? Dadurch nämlich, daß *der* es nicht spielt, oder falsch spielt, der Gegenstände nicht mit Sicherheit erkennt". Wittgenstein (1969: §446).

[109] „Am Grunde des begründeten Glaubens liegt der unbegründete Glaube". Wittgenstein (1969: §204). See also Wittgenstein (1969 §624) and Wittgenstein (1980: §151).

[110] "Since the number one, is the same for everyone, confronting everyone in the same way, it can be explained through psychological observation as little as the Moon. After all, there may be different ideas of the number one in the individual minds, but they must be distinguished from the number one the same way that the idea of the Moon must be distinguished from the Moon itself".

„Da die Eins, als dieselbe für Alle, Allen in gleicher Weise gegenübersteht, kann sie ebensowe-nig wie der Mond durch psychologische Beobachtung erforscht werden. Mag es immerhin Vor-stellungen von der Eins in den einzelnen Seelen geben, so sind diese doch von der Eins ebenso zu unterscheiden wie die Vorstellungen des Mondes von dem Monde selbst". Frege (1966: xviii, §8).

[111] „'An diesem Satz kann ich nicht zweifeln, ohne alles Urteilen aufzugeben.'

Just as for Frege, our knowledge about number one, just as our observations of the Moon, have no place in doubt. This is because our attitudes, feelings and thoughts about numbers and the Moon has no bearing on their existence and nature. Objective certainty does not belong to one's feeling or belief about something. Denial, according to Wittgenstein, is denial of the grammatical, rule–like structure not only of one's language but the world as we know it:

> There are cases where doubt is unreasonable, but others where it logically seems impossible. And between them there seems to be no clear boundary.[112]
>
> With the word 'certain' we express the utter conviction the absence of any doubt, and by doing this we seek to convince others. That's *subjective* certainty.
>
> But when is something objectively certain? – If an error is not possible. But what kind possibility is it? Mustn't the mistake be *logically* excluded?[113]

Knowledge does not apply to either subjective or to objective certainty. Objective certainty is beyond doubt; subjective certainty is our own conviction about which we try to convince others. Knowledge is the information coming at the end of the inquiry process. Objective certainty requires no process, just as primitive language–games do not. It is the outcome of the primitive language–game, it is what the primitive language–game *implies*. In many texts one of the favorite examples of the primitive language–games in Wittgenstein is seeing colors:

> Can a man doubt whether what he sees is red or green? [Elaborate this].
>
> "Surely if he knows anything he must know that he sees!" – It is true that the game of "showing or telling what one sees" is one of the most fundamental language games, which means that what we in ordinary life call using language mostly presupposes this game.[114]

To recapitulate, *Tractarian* semantics insisted on using phenomenological language to describe likewise phenomenological objects. Because such language fails in its job of description, it was substituted by physical language. Still, neither physical language can make a distinction between real and imaginary things, and without language–games the notions of truth and even existence are not possible to be conveyed in it. Once Wittgenstein introduced language–games as the main link between reality and language, physical language of the language–games could not serve anymore as a description of phenomenological objects.

It was not a simple mismatch, for mismatches can be handled as long as there are good criteria for distinguishing truth and falsity, necessity and contingency. If language–games are activities, they cannot be in a business of describing something as unreal as

---

Aber was für ein Satz ist das? (Er erinnert an das, was Frege über das Gesetz der Identität gesagt hat.) Er ist sicher kein Erfahrungssatz. Er gehört nicht in die Psychologie.

Er hat eher den Charakter einer Regel". Wittgenstein (1969: §494).

[112] „Es gibt Fälle, in denen der Zweifel unvernünftig ist, andre aber, in denen er logisch unmöglich scheint. Und zwischen ihnen scheint es keine klare Grenze zu geben". Wittgenstein (1969: §454).

[113] Wittgenstein (1969: §194).

[114] Wittgenstein (2000: MS 149:1, 1935).

the Lost Island, inexistent sensation, or unconceivable and non–communicable meaning.[115] For all these reasons giving up Cartesian semantics is not enough; one must also give up Cartesian metaphysics and ontology, then Cartesian epistemology is dropped out as well.

In his last writings the concept of sense–data or of any other name for phenomenological objects appears to be less and less relevant:

> When we were young, we were taught generic names of physical objects. That a human being has the concept of a physical object is shown by the fact that he brings chairs when asked, etc. It is said that the propositions about physical objects to be analysed in terms of sense data. But the fact is we have been able to learn to obey. "Bring a chair". We have mastered this technique regardless of how a chair looks; we have learned this without first learning another game. If it makes any sense to say that we have a concept of physical object, that is what it is. This is the technique we learn, not the other one.[116]

In this passage it appears that obeying an order takes care entirely of the problem of learning names of objects and understanding situations. To learn to obey means learning the meaning of the word 'chair'. What does this learning to obey mean? It simply means learning the rules of the game of bringing the chair when asked. When are the propositions about physical objects analyzed? When does one acquire the sense–datum of a particular object? When does one know what the chair means? It is rather when one learns the meaning of the word through its use, viz. through learning how to play the game, "without first learning another game".

This has an important consequence for our notion of knowledge and its justification: the old epistemological questions. Idealism and realism always presupposed something intermediary between object and name on the one hand, and name and understanding on the other. Be it an idea, form, sense–datum, or any other phenomenological realm. In the last years of his life Wittgenstein returns to these hard questions of epistemology. He *returns* because these questions were already answered in the TLP and in his early answers phenomenological objects played as much an important role as his picture theory of language.

His answers this time are radically different. Instead of representational objects, ideas, forms and sense–data Wittgenstein insists that use and activity in which meaning manifests itself is enough. Now language–games can take care of themselves and of the meaning: "But how can we *show* somebody that we not only *know* truths about sense data, but also about things? Because it certainly cannot be enough for someone to assure us that he knows it. Where do we start to show it?"[117]

---

[115] In fairytales the ordinary language description also implements, of course, language–games of both kinds. In these cases, however, the unreality or some parallel world of fairytales must be specified in the rules of these language–games. The universal character of language–games by no means excludes the possibility of ordinary language description of surreal objects.

[116] Wittgenstein (1988: 177–178).

[117] „Wie aber ist es Einem zu *zeigen*, daß wir nicht nur Wahrheiten über Sinnesdaten, sondern auch solche über Dinge *wissen?* Denn es kann doch nicht genug sein, daß jemand uns versichert, *er* wisse dies.

The answer begins by indicating the conduct of a person in which the meaning as well as knowledge is manifested. The idea is then carried further by analyzing the language–games in which such conduct appears: "We must show that, even if he never used the words 'I know...' his conduct shows what we are concerned about".[118]

If one's beliefs, intentions, knowledge and understanding are one's conduct, there is no need of referring to sense–data or any other kind of phenomenological objectry. As Danièle Moyal–Sharrock points out, the relation between world and mind, or mind and action, exists but it is not located in some Platonic (or Fregean) third realm. It is explained in and by the language–games and their analysis.[119] Edward Minar and Marie McGinn speak of the end of the 'mythological understanding' of the mind as some 'place' in which private objects are to be found.[120]

How would this analysis proceed? I already indicated the important distinction Wittgenstein used himself between primitive and complete language–games, as well as the usual division within each game between its defining rules and strategies. But there is another important analysis that goes beyond the internal constitution of language–games. It is the analysis of one game by another. It can be described as an analysis of the primitive language–game by its complete language–game.

This can be shown in a rather complex example of how one can figure out the complete language–game of joy (the inner smile) from the primitive language–game of a smile on someone's face:

> First question: How do you know, how do you judge, whether his inner face smiles? Second question: How important is it? – But the two are related. And one might ask *another*, albeit related, question like this: What degree of importance does his – external – smile have? After all, if the inner is of importance, then so also – in a (slightly) different way – must be the outer.
>
> (It is not easy to see that my manipulations are justified.)[121]

It was mentioned before that complete language–games are constituted by the primitive ones and the complete language–games can be analyzed by analyzing their constitutive primitive language–games. But in the example above, the order is reversed. There is a potentially complex complete language–game whose manifestation is a physiognomic primitive language–game. Since the purpose of primitive language–games is to

---

Wovon muß man denn ausgehen, um das zu zeigen?" Wittgenstein (1969: §426).

[118] „Man muß zeigen, daß, auch wenn er nie die Worte gebraucht ‚Ich weiß, . . .,' sein Gebaren das zeigt, worauf es uns ankommt". Wittgenstein (1969: §427).

[119] Moyal–Sharrock (2004: 204).

[120] Minar (2010); McGinn (1997: 35).

[121] „Erste Frage: Wie weiß man wie beurteilt man, ob sein inneres Gesicht lächelt? Zweite Frage: Welche Wichtigkeit hat es? – Aber die beiden hängen zusammen. Und man könnte eine andere, wenngleich verwandte, Frage so stellen: Welche Wichtigkeit hat sein – äußeres – Lächeln? Denn, ist das innere von Wichtigkeit, so muß es – in einer (etwas) andern Art – auch das äußere sein.

(Einzusehen, daß meine Manipulationen gerechtfertigt sind, ist nicht leicht)". Wittgenstein (1992: II, 86); Wittgenstein (2000: MS 174:4v., 24.4.1950). See also Wittgenstein (2000: MS 174:8v–9v.).

be the immediate link between language and reality, according to Wittgenstein they cannot be corrected or analyzed into something more basic. Merrill and Jaakko Hintikka stress that if such correction would be the case, then they are no more the representative links between language and reality.[122]

But we see that such further analysis is going on in LW II, 86. Does Wittgenstein contradict himself? Not really. In that passage the primitive language–game of smiling serves as the complete language–game of the inner smile. It can, therefore, be analyzed but *only as the complete* language–game standing in a particular relationship to its constitutive primitive game of inner smiling.

What happens here and in many other places in Wittgenstein's last writings is something of psychoanalysis. We have visible manifestations of hidden private states, like the perceptible behavior of an invisible and often subconscious state. If we consider these visible manifestations as a complete language–game and the subconscious state as its constitutive primitive game then we might be able to apply an analysis of the complete games to uncover the hidden reality of its primitive game.

Wittgenstein sometimes called himself a 'disciple of Freud'[123] in that he, like Freud, used metaphors to get to the meaning of hidden psychological states. Philosophy as a therapy of language fits within this analytical context as well. Wittgenstein was also compared to the Russian developmental psychologist Lev S. Vygotsky whose methods included placing language as an essential element in cognitive development and considering inner thinking as a result from public speech.[124]

Notwithstanding these important similarities, Wittgenstein offers an unique methodological tool in the form of language–games. When we analyze them we analyze that which stands between us and the world and between the world and language. One might say that what stands between us and the world and between the world and language are our actions. As true as it is, these actions are in need of explanations. Language–games are open to such explanations and analysis.

In light of these remarks one final note regarding §293 is in order. There are two opposite but equally false common interpretations of the text. The first is given by the

---

[122] Hintikka and Hintikka (1986: 279, Thesis 4).

[123] Comparisons between Wittgenstein's philosophy and Freudian psychoanalysis must be taken with caution. As Ray Monk remarks: "Wittgenstein's feeling that he would have made a good psychiatrist seems to rest on a belief that his style of philosophizing and Freudian psychoanalysis required a similar gift. Not, of course, that they are the same technique. Wittgenstein reacted angrily when his philosophical method was dubbed 'therapeutic positivism' and compared with psychoanalysis. When, for example, A. J. Ayer drew the comparison in an article in the *Listener*, he received from Wittgenstein a strongly worded letter of rebuke. However, Wittgenstein was inclined to see some sort of connection between his work and Freud's. He once described himself to Rhees as a 'disciple of Freud', and at various times summed up the achievements of both himself and Freud in strikingly similar phrases. 'It's all excellent similes', he said in a lecture of Freud's work; and of his own contribution to philosophy: 'What I invent are new *similes*.' This ability to form a synoptic view by constructing illuminating similes and metaphors was, it appears, what he wished to contribute to psychiatric medicine". Monk (356–357). Among other studies on Wittgenstein and Freud are Lazerowitz (1977), Bouveresse (1991) and Cioffi (1998).

[124] Williams (1999: 260–281).

defenders of ordinary language, who say that Wittgenstein's beetle–in–the box thought experiment is a failure because of its assumption that ordinary language is not able to transmit the meaning of the event, sensation or fact. The second is a consequence of the denial that picture theory, ostensive definition or ordinary language can indeed successfully convey the meaning, and that only human activity is the place for such 'definitions'. In this interpretation, it seems, whatever sensation we might have cannot be private in any way. A human person (her face and behavior) is thus a mirror of her sensations.

Both interpretations are exclusive of some aspect of meaning expression. In fact, Wittgenstein never denied the role that ordinary language can play within a particular language–game. Not every pain behavior must be moaning and screaming. It can also include language use in defining one's sensation. The beetle–in–the box experiment shows only that *mere* linguistic description, just as *mere* ostensive definition is not enough for meaning acquisition/communication. Language–games can include ostensive behavior and picturing as long as they are incorporated within the rules of the game but are not exclusive means of communication. Whenever ostensive behavior or ordinary linguistic description of some private event is applied, there is always the possibility of ambiguity or idleness.

While the first misconception of the thought experiment is widespread in non–professional literature,[125] the second one is common in the professional philosophical texts. If neither *Tractarian* (ostensive) isomorphic showing nor the ordinary language definitions are admissible, then all that remains is our ordinary behavior. This interpretation would then deny the existence of any private states that are not somehow manifested externally. The problem is that some private experiences might not be shown, while it can easily be the case that showing is not connected with any particular behavior but rather results from several instances of conduct.

Private sensations are real regardless if they are part of some manifested language–game or not, but if we want to communicate them, then the only way to do so is by language–games. These games can, of course, include ostensive behavior, isomorphic interplay between different language–games resulting in a family–resemblance between them. The main point is that neither language alone nor behavior alone is enough.

For many it is a sad fact that Wittgenstein was very critical of Shakespeare's dramaturgy.[126] He did not like aphorisms, witty metaphors and long–worded analogies. Unfortunately, some have misinterpreted many language–games in the PI for metaphors and short aphorisms. The reason why Wittgenstein was critical toward Shakespearian and Socratic uses of language is that both, each in his own way, believed that long linguistic analysis would transmit the meaning and clarify it. In Shakespeare's time a theatrical play would have to include lengthy dialogues and soliloquies in order to convey the character's internal world of emotions, intentions and feelings.

The Russian theatrical reformation in the late–19th century ('Stanislavky's System') by K. S. Stanislavsky (1863 – 1938) and A. P. Chekhov (1860 – 1904) consisted of going from

---

[125] See, i.e., Cohen (2005: 106–119).

[126] Wittgenstein's references to Shakespeare alone are too many to give a full bibliographical account here. See, i.e., Wittgenstein (1998: 95–98) and Wittgenstein (2000: MS 131, 162, 168).

language of emotions to the playing of them out on the stage.[127] The same would go for Wittgenstein's favorite author F. M. Dostoyevsky (1821 – 1881) whose style was capable of transmitting the feelings of the personages in his novels by narrating seemingly unrelated things, such as the architecture of the summer city, stiffness of the air, a description of the objects in the room, even their positions and relations to the people (i.e. the door in *Crime and Punishment*, but also accentuation of the role of small objects and their relation to the personages in Chekhov's plays, i.e. chairs in *The Seagull*, knitted stockings and pencil in *Uncle Vanya*, and cupboard in *The Cherry Orchard*).

Any use of the ordinary language, no matter how elaborate it can be in its expressions, will always presuppose incompleteness and ambiguity in conveying human psychological states. Linguistic analysis likewise would always be incomplete in understanding private states. Psychoanalysis through language–games in their three–dimensional application can serve as an alternative and, at the same time, linguistically–inclusive method of study.

5. Naturalizing the Language–Games

5.1 Epistemic and Non–Epistemic Perception

It has been said that only in the case of a complete language–game can we speak of knowledge, interpretation, and conscious states. Primitive language–games are knowledge–blind, interpretation– and consciousness–free activities yet to be composed in complete and more complex language–games. The idea that in the case of relative criteria change, a primitive language–game can be considered as a complete one should not be seen as some sort of exception from the rule. One's physiognomic expression is a primitive language–game in relation to the person's expression of joy, fear, pain, pleasure, etc., but it is also, at the same time, a complete language–game in relation to the internal state of that person. The analysis of someone's behavior, the examination of some situation, the search for a strategy as a solution to some specific problem can be done by first identifying correctly complete language–games then discerning within these complete language–games their constitutive primitive language–games, and so on, until one runs out of the criteria for any further analysis.

Applications of the above sketched analysis can be extended to many areas of investigations. At this point, however, it would be more fitting to focus on one important faculty, perception. Without claiming to give a new theory of perception I shall indicate how language–game analysis could be applied to the study of perceptual processes.

Fred Dretske in his 1969 *Seeing and Knowing*[128] introduced set of important distinctions relevant to the notion of perception. He begins by asserting some common area of agreement among philosophers that perception is indispensible in knowledge acquisition and that sense–perception is the source of our information about objects and facts.

---

[127] Moore, Gielgud and Logan (1984).

[128] The following presentation will be mostly focused on the first chapter in Dretske (1969: 4–77) and Dretske (2000: 97–112).

Dretske makes a rather common–sensical distinction between 'perceptual knowledge of facts' and 'perceptual knowledge of objects'. The perception *of* a cat and perception *that* there is a cat are very different kinds of perceptions and different kinds of information. If someone, say a visitor from a distant planet, has never seen a cat but has the very same perceptual apparatus as we do, seeing it for the first time will not be enough for her to know that this object is a cat.

The acquired perception without knowledge is non–epistemic. The object is known for its physical properties but not known *as* an object due to the lack of the public identification of the object. Contrary to object–perception, fact–perception is always epistemic. I can only know that it is a cat on the mat if and only if I had some previous knowledge about cats.

Dretske concludes that most of our perceptual knowledge is indirect and dependent. The learning process consists in our ability to connect different links of already acquired knowledge about many facts in the world. I might learn, for instance, about new cat–species, about some cat–behavior, etc. only by connecting this new information with my previous knowledge about cats. Unlike me, the recently landed extraterrestrial would have to start first by learning what cats are.

Derived knowledge is acquired not only in vision but in every other sensory faculty as well. Perceiving that *a* is *F* and knowing that *a* is *F* depends on perceiving a different condition, that *b* is *G*. My knowledge that *a* is *F* derives from a more basic perceptual knowledge that *b* is *G*: I *know* that she is in pain because I *see* her facial expressions (*b* is *G*) in some particular context (*a* is *F*).

There is, however, another kind of derived knowledge, when we know about the object by perceiving only the object itself: we know that *a* is *F* by perceiving it, not because *b* is *G*, but because *a* itself is *G*. But this kind of knowledge is also the knowledge of facts about objects, not just the mere perception of it. Therefore, knowing that someone is in pain comes from observation of the physiognomic *facts* of the subject's face.

In the later case, the inferential character of knowledge does not involve logical inference, viz. when we pass from premises to conclusion. It is rather a psychologically immediate kind of knowledge: from seeing someone's pain behavior I *infer* that she is in pain. This immediate kind of perceptual knowledge does not involve any conscious mental process but rather an instinctive reaction and recognition.

However, even this psychologically immediate perceptual knowledge too requires training and learning. There is no innate knowledge of pain–behavior. In this context Dretske distinguishes between 'beginners' and 'experts' of learning. Whenever a beginner perceives some new object or a fact (that *a* is *F* without knowing or perceiving that *b* is *G*), she makes an inference (true or false) on the basis of her perception. In this way perception is a learning process. The expert, on the other hand, has no need of going through the process of recognition, as long as all of the necessary facts for the identification of the object's features are given in the single act of perception.

Knowing that *a* is *F* by perceiving that *b* is *G* would presuppose some background. Internalism and externalism differ, however, to what extent the background assumption is required. Internalists (as we had seen in the case of the intentionality of perception in Searle) would require that background assumption would be present not only in epistemic perception but in non–epistemic as well. Indeed, non–epistemic perception would

always presume some background of knowledge by the perceiver. Externalism, on the other hand, assumes that a simple act of belief would suffice in the case of non–epistemic perception.

Dretske inquires whether derived knowledge can in fact be called knowledge at all. If I know that *a* is *F* on the basis of knowing that *b* is *G*, do I really *know* that *a* is *F*? Perception seems to be a mere part of the more important process of knowledge acquisition in which background assumption is indispensable. In this case, all perceptual knowledge is theory–laden and appears to be an ability of connecting new pieces of information with the previously acquired ones.

If this is the case, then naturalism is inconsistent. Naturalism would imply that sense perception is directly responsible for our knowledge acquisition. Nothing intermediary can stand between the senses and the brain. It should not presuppose anything like background assumptions or any other knowledge that is not based on the sensory experience itself. If, however, perception as a process in any way depends on some intermediary background assumptions then naturalism is inconsistent or at the least compromised.

A naturalist solution would be either direct realism or representational realism. Direct realism (i.e. in Searle) insists that knowledge is not dependent on any other knowledge or belief. It is justified within the experience itself and not by some other epistemic assumptions. Direct realism would, of course, presuppose learned skills of identifying objects and facts, but that skill is not based on some other more basic knowledge. Instead, these skills were acquired directly from the experience.

Representational realism considers perceptual knowledge to be of objects such as ideas, impressions and sense–data. These are the objects in the mind of the observer. Sense–datum is a subjective appearance in the mind of the perceiver which can, nonetheless, be studied objectively. Knowledge of these objects, realists claim, is certain. One cannot be mistaken about these objects because they are the way things appear to be. Sensory perception is then an indirect perception of the physical objects. Unlike direct realism, representationalism accepts that our knowledge is indirect and relies on some background assumption.[129]

Within representative theories one must distinguish between, what William Seager and David Bourget, call classical and modern representationalism.[130] The above description is true of the classical representationalism. It relies on the sense–data as that representative link between mind and reality. Modern representationalism, on the other hand, does not postulate any phenomenal objects. Instead, according to Robert Cummins' reading, their adherents interpret informational content as a covariation between the representing state and the state represented.[131]

Dretske formulates his version of representationalism for a philosophy of mind in terms that all mental facts are representational and these facts are about acquired information. This way he accepts private, qualitative experiences but suggests that they are studied objectively and without considering consciousness in any phenomenal way.[132]

---

[129] Dretske (1993).

[130] Seage and Bourget (2007).

[131] Cummins (1989: 67–75).

[132] Dretske (1995: lectures 1, 3, and 5).

Although the representative realism solution is not as radical as that of the direct realism, it does restrict background assumptions as auxiliary to the perceptual knowledge and not as knowledge on its own. What is, however, most interesting about both solutions is how closely they both consider the mind's directness toward the world. In this, of course, sensory perception plays an essential role.

Representationalism insists that mental states represent the physical reality. It does so in a naturalistic way (usually by means of neural communication) without any recurrence to semantics. Both realisms insist that the starting point in explaining the mind's relation to the world (direct or represented) must be with sensory perception. Whether it is representationally or directly, the brain gets its information about the world through the senses. This, of course, presumes that sensory transmitting of information must be veridical even if some fallibility is always possible (insufficient conditions for perception, illness, injury, hallucinations are among the usual exceptions to the rule).[133]

In her well–known study Kathleen Akins takes on mainly the representative realism/naturalism study of sensory perception.[134] She identifies the main point of representationalism about sensory perception as the necessary mirroring between sensory signals and relevant relations in what is sensed. The relation should be isomorphic: the warmth transmitted by the neurons in the body must correspond somehow to the temperature in the environment. This way the sensory systems serve the brain in identifying the situation 'out there'. Through top–down processing the brain uses stored information to match the new upcoming data and interprets it accordingly. Sensory systems communicate that which is required by the brain without 'knowing' nor deciding about the information they transmit.

One might recall the Aristotelian theory of perception in which the perceiving organ is in some way like the perceived object. The correspondence between these two is the condition for perception to occur. It is no surprise then that for many contemporary naturalists the Aristotelian theory of perception is open for consideration.[135]

Akins uses studies on the functioning of thermal receptors in the skin (but mentions also similar functions in the visual perception) to show that sensory systems are 'narcissistic' more than veridical. Each part of the system gives very different accounts to the brain about the environment. The example of the thermal receptors in the skin is less complex than visual sensory systems, but they are complex enough to show four different types of receptors: two thermo–receptors that identify warm and cold spots and two pain receptors that react only in the extreme conditions of temperature change. While we feel the gradual temperature change in the body it is not the result of the continuous response of one thermo system.

The system is narcissistic in the sense that it informs of the changes in temperature only relative to itself, not to the overall situation in the body. Signals separated from the whole picture are ambiguous and in need of interpretation. Scientific studies that measure each thermoreceptor's response to the temperature changes show that each system exaggerates the situation and reacts according its own 'needs and interests'. Akins

---

[133] Dretske (1981: ch. 6, esp. pp. 153–168).

[134] Akins (1996).

[135] See, i.e., Damasio (1994).

postulates that the naturalists' question about the sensory systems "What is the system detecting?" should really be "What is the system doing?"[136] Every sensory system is a part of the set of a specific behavioral task. Human (and animal in general) behavior is not a simple stimulus response based behavior. It includes thinking, projecting, discerning the action based on the sensory information after it was processed by the brain.

Perhaps the necessity of the corresponding features between the sensory system and the world as it is was result of type theories of 1950s. The correspondence between pain and C–fiber firing was also seen in terms of identity. Since then we know from neurobiology that pain is a much more complex set of states usually within the central nervous system rather than a local physical event.

Akins reports similar–in–structure but superior–in–complexity vision systems in mammals. Different visual percepts report a specific but distinct piece of information. Each has a specific connection with some motor site while all visual areas have connections with subcortical sites. These then have their own connections with motor areas. Akins concludes that "Even as intentional, conscious perceivers, we are equally in need of narcissistic sensory strategies, indeed, perhaps more so given the complexity of our behavior and bodies".[137]

Akins herself claims to be a naturalist in understanding brain functions and in considering sense perception. The main point of her study was not to refute naturalism but rather indicate that the isomorphic relation between sensory perception and physical states of the environment is far from being the usual picture. Her analysis undermines the received naturalistic view of how the brain acquires information about the environment.

Akins' study in some respects improves Dretske's representationalism without threatening his important distinctions between epistemic and non–epistemic perception. In fact, this distinction goes in hand with Akins' insistence on the narcissistic character of sensory systems.[138]

Non–epistemic values of these systems and non–epistemic perception go also in hand with Wittgenstein's notion of the primary language–games as the most basic representative links between language and reality. In terms of perception, primitive language–games operate the way non–epistemic perception does. The main role of the primitive language–game is to establish the relation between word and object; it is not to convey any information or knowledge of any sort. But just as non–epistemic perceptions are constitutive of epistemic perceptions, so the language–games are constitutive in complete language–games where epistemic designators are present.

The notion of the background assumptions was already briefly explained by the family–resemblance notion and the use of the picture theory as the third dimension of the language–games. Both notions must include the networks of the language–games in the explanation of the background assumptions. What unites this Wittgensteinian in nature

---

[136] Akins (1996: 352).

[137] Akins (1996: 354). A very similar account of visual perception can be found in Kandel and Wurtz (2000: 492–506).

[138] By Dretske's own recognition, his obvious distinctions between different kinds of perception received more criticisms by naturalists than any other part of his philosophy. See Heil (1991).

vision with that of the realist concept of background, is that background is not considered as knowledge on its own but rather as a unifying strength between past and present perceptions.

Wittgenstein did not have a theory of perception. In the secondary literature, explanations of Wittgenstein's remarks on perception are usually centered around the well known passage from PI II, xi, where he presents the notion of 'seeing an aspect'. It would be wrong, however, to say that 'seeing an aspect' could be extended to all perception. Instead, it refers only to those perceptive experiences that are ambiguous in perceiving the same shape, object or picture.

In this work I will not consider Wittgenstein's treatment of 'seeing an aspect', since its purpose is centered on language–games. There is, however, one important application of the language–game notion to the overall treatment of perception in Wittgenstein.

While 'seeing an aspect' is an action of the will of the perceiver (i.e. seeing the figure from specific angle or focusing one's eyes on a part of the picture generates a certain image), perceiving in general is apprehending the visual or any other sensory experience as it is. In terms of the language–games nomenclature suggested before, perception in general works according to the *strategies* of a given language–game, while 'seeing an aspect' is a change in *tactics* implemented in a particular strategy of the game of perception. By a tactic I simply mean the way a particular strategy is implemented and as such, the notion of tactics plays a very limited, auxiliary role within the larger (epistemic) notion of strategy.[139]

Perception can be both analytic and synthetic. Most visual perceptions are analytical: we perceive different parts, colors, shapes but we are conscious of the object as a whole (strategies without tactical implementation). Most olfactory and auditory perceptions are synthetic: we perceive the content as a whole, even if later we might discern some different scents or sounds (strategies and tactics of discernment, focus and attention). The common point, however, is that both analytic and synthetic perceptions operate by the common rules which are applied in different strategies. Some strategies are analytical, others are synthetic, but they are both strategies of perceptions.

---

[139] In the military there is a very clear hierarchy of interrelated levels: strategic, operational and tactical. "Strategy involves establishing goals, assigning forces, providing assets, and imposing conditions on the use of force in theaters of war". "Tactics refers to the concepts and methods used to accomplish a particular mission in either combat or other military operations. (…) We normally think of tactics in terms of combat, and in this context tactics can be thought of as the art and science of winning engagements and battles". Finally, "the operational level of war links the strategic and tactical levels. It is the use of tactical results to attain strategic objectives". U.S. Marine Corps (1997: 28–30).

J. von Neumann and O. Morgenstern likewise associate the notion of strategy with that of planning: "Imagine now that each player $k = 1, \ldots, n$, instead of making each decision as the necessity for it arises, makes up his mind in advance for all possible contingencies; i.e. that the player k begins to play with a complete plan: a plan which specifies what choices he will make in every possible situation, for every possible actual information which he may have at that moment in conformity with the pattern of information which the rules of the game provide for him for that case. We call such a plan a *strategy*". Emphasis is by the authors. von Neumann and Morgenstern (1953: 11.1.1 §79).

5.2 Non–Epistemic Proprioception

While one must clearly distinguish between epistemic and non–epistemic perception in the case of sense–perception, such a distinction is not so obvious in the case of proprioceptions. Wittgenstein is credited with being the first to point out, against empiricist accounts of proprioceptions of his time, that the sense–experiences from the external world and the feelings of the movements and positions of one's own body (proprioception) cannot be considered as the same communicative ways our body processes information. The usual adversaries of Wittgenstein who supported the empirical approach to proprioception as well as to sensory perception, were William James[140] and Wolfgang Köhler[141], but their account was widely used by the sense–datum British philosophy, i.e., by Bertrand Russell's so called 'neutral monism' ("neither mental nor material").[142]

Wittgenstein's starting point for his criticism of the contemporary theories of proprioception is that the awareness of one's own body cannot be reduced to the sentiment of bodily sensations and that they cannot be put on the same conceptual level as sense–perception. His attention was mostly directed to kinaesthesia (feeling of one's limb movements). His writings on kinaesthesia began in PI I and were elaborated in the post–*Investigation* period (1946–51); they are found in PI I §§611–628, PI II[143], RPPI and LPP.

What Wittgenstein distinguishes sharply from sense–perception are two events: our immediate proprioception and emotions. It is on these topics that his major disagreements come with James, Köhler and Russell. In what follows I shall only concentrate on Wittgenstein's treatment of the first event, proprioception and kinaesthesis, although from the language–games perspective the same could be said of perception of our emotions.

When Wittgenstein compares kinaesthesis with sense–perception he mentions that the difference between them is a grammatical one: "Now I am looking for the grammatical difference".[144] But what is this grammatical difference?

The sense of touch, i.e., supplies the subject with information about the touched object: its shape, position, movement, weight, etc. However, by the feelings in our bodies we cannot judge their position or their movements. The feeling of the position and movement of our limbs, i.e., which we certainly do feel, is not *caused by* the sensation in the limbs. The reason being that these feelings are simply insufficient for me to *know of* the position and movement of the limbs. Can they be somehow increased to the point when

---

[140] James (1950: chs. 20–23).

[141] Köhler (1930: 127–129).

[142] Russell's 'neutral monism', however, was really just another version of physicalism, since it interpreted all psychological phenomena as "build up out of sensations and images alone". See Russell (1921: 279).

[143] Properly speaking, *Philosophical Investigations* Part II is *not* the same book as *Philosophical Investigations* Part I, as has been argued by von Wright (1980: 111–136). In Hacker–Schulte's edition this part is titled as "Philosophie der Psychologie – Ein Fragment". Thus, by the term 'post–*Investigations* period' is meant Wittgenstein's writings after Part I was completed in 1945. Some scholars call this period for the last writings of Wittgenstein as the 'Third Wittgenstein'. See Moyal–Sharrock (2004).

[144] „Ich suche nun nach dem grammatischen Unterschied". Wittgenstein (2001: II, viii).

feeling them I *know of* my body's position? Wittgenstein claims that it cannot, unless the feeling is very specific and independent of other sensations.[145]

He illustrates this point by a simple experiment of moving his index finger in a pendulum movement, without looking at it. All one can feel, he claims, is "a slight tension" but not in the joint, on the tip of the finger. At this low intensity of feeling and by this feeling alone one cannot *know of* the position of one's finger. What is required here for one to *know of* the position and movement of the finger is being aware of it even without seeing the movement.[146]

Here Wittgenstein makes a very important distinction. Being aware of the movement is not the same as being aware of the feeling that is caused by the movement. Again, the feeling caused by the movement is not enough to cause awareness. Here, awareness of the movement and the position of the finger is an awareness of feeling, not of sensation. Awareness of sensation is immediate in sense–perception, not in kinaesthesis.

This last point Wittgenstein proves by considering cases in which I might be very tired or that the movement of my limbs also causes great pain and takes away the possibility for me to be aware of the directionality of movements and positions of my limbs. In other words, being aware of the feelings in such cases would not let me know of the position of my limbs.[147]

The conclusion that Wittgenstein draws from these and other similar examples is that it would be incorrect to suppose that my awareness of the feelings in my body would give me *knowledge of* the movements of my body. Instead, it is the quality and strength of the sensation that gives me such knowledge.[148]

Thus, awareness is something additional to what might or might not accompany the feeling. Wittgenstein compares kinaesthetic feeling *in need of* awareness to the sensation of pain *in need of* the external sign pointing to the place where the painful sensation occurs. But in both cases the awareness and the external sign does not belong either to the feeling or to the sensation: they can occur perfectly without any awareness.[149]

Malcolm Budd points out important criteria to distinguish between kinaesthetic feelings and bodily sensations. Based on RPP II, 63, he indicates four such general criteria of sensation: intensity, quality, duration and location. Under these characteristics, kinaesthetic feeling is in no way similar to sensations. Bodily feelings cannot be considered in any distinctive degree. It has no qualitative nature either. Feelings in my body, as it was mentioned already, cannot let me know its location, nor can feeling be measured by duration of time.

Feelings are of the objective certainty when brought to awareness, while sensations require subjective experience. Feelings are based on my immediate conviction of having them; sensations must have some qualitative grounds rooted in experience. I am certain that I move my arm because I feel it to be moved: "'How do you know that you've raised your arm?' – 'I feel it'. So what you recognize, is that the feeling? And are you sure that

145 In Wittgenstein (1980: §400, §790).

146 Wittgenstein (2001: II, viii).

147 Wittgenstein (1980: 382).

148 Wittgenstein (1980: 400–407); Wittgenstein (1988: 78).

149 Wittgenstein (2001: II, viii).

you properly recognize it? – You are sure you have your arm raised? Is this not the criterion, the degree of recognition?"[150] I know about the position of my arm through the direct awareness without observation. While observation is needed for sensations, awareness is necessary for feelings to generate knowledge.

In English it is difficult to make a clear difference between 'feeling' and 'sensation', given the fact that most dictionaries consider them synonymous (see Joseph Devlin's *A Dictionary of Synonyms and Antonyms* and *Oxford English Dictionary* definition of 'feeling' as "a physical sensation"). In Wittgenstein, however, these are different terms, such as, i.e., 'feeling fear' and 'sensing pain'. The first is time–irrelevant (i.e., fear of heights) while the second has clear time–dependent characteristic (i.e., sensing headache).

In German the difference is as difficult to notice as it is in English. The word '*Empfindung*' means feeling but so does '*Gefül*'. Wittgenstein uses both words in PI II, viii and until §63 it is not clear that there is any difference. However, the entire chapter viii proceeds in a crescendo style. In the first section of the chapter (§56) he uses '*kinästhetischen Emfindungen*' within a quotation, just as in the above §625, of some interlocutor who, indeed, does not make any distinction between 'feelings' and 'sensations'. Wittgenstein then proceeds by giving more examples of misuse, asks questions, shows inconsistences and, in §60 he inquires about the criterion for the sensory impression "*Sinneseindruck*". Then, from §61 the vocabulary changes radically. We do not encounter the word '*Empfindungen*' but instead '*Gefül*' at a consistent rate in every remaining section of the chapter viii. Thus, the correct word for feeling is '*Gefül*' and sensation is '*Empfindungen*'.

Anscombe, Hacker and Schulte use the same translations. When Wittgenstein uses in a quotation "*kinästhetischen Empfindungen*" in §56a, it is translated as "kinaesthetic sensations", and when Wittgenstein uses "kinästhetischen Gefühl" in §63a[1] both translations says "kinaesthetic feeling". We must, however, keep in mind that there are no kinaesthetic sensations ('*Empfindungen*') but only feelings ('*Gefül*').

We can now return to Wittgenstein's remark in PI II, viii §62 about the grammatical difference: it is necessary to keep in mind the relative complexity of sensations versus feelings. Sensations require experience; feelings are experience–neutral. The grammatical difference consists in that while both feelings and sensations are primitive language–games, sensations in context can be considered as complete language–games in themselves, relative to the physical states of the body that causes them. Feelings can only be primitive language–games and only when they are brought to awareness. There is nothing more basic or more primitive than feelings.

Another important difference has particular application within the philosophy of language. Feelings, unlike sensations, cannot express judgments. Feelings are expressions of psychological and physical states, in the same way in which groaning is an expression of pain. The linguistic expressions of feelings, i.e., 'I feel dazed', differ from the linguistic expression of sensations, i.e., 'I see a red patch'.[151] Recall also that certainty of the

---

[150] „'Wie weißt du, daß du deinen Arm gehoben hast?' – ‚Ich fühle es.' Was du also wiedererkennst, ist die Empfindung? Und bist du sicher, daß du sie richtig wiedererkennst? – Du bist sicher, daß du deinen Arm gehoben hast; ist nicht dies das Kriterium, das Maß des Wiedererkennens?" Wittgenstein (2001: §625).

[151] This point is further elaborated by Romdenh–Romluc (2010).

expression of feeling is different from the expression of sensation. In fact, to this difference Wittgenstein dedicates most of his last work *On Certainty*, against George E. Moore's assumptions that the expressions of feelings and kinaesthetic affirmations need as much proof as the expressions of sensations and that they have epistemic relevance.

Stewart Candlish mentions that any treatment of feelings vs. sensations would rely on a kind of identity theory of mind/body we postulate.[152] Free from the need of experience and observation, feelings neither postulate presence of an active mind nor can be fully explained by the activities of the dorsal root ganglion neurons responsible for proprioception.[153] While sensations fall to an easy prey of Cartesian dualism and physicalism alike, kinaesthetic feelings are immune to any exclusive theoretical treatment. They do not *postulate* unity of mind and body, for most cases they are involuntary actions, yet they are capable of calling our attention and being guided by our conscious will. Their status will depend on the theory of identity we choose to implement.

6. Wittgenstein on Identity

Wittgenstein's remarks on identity are in the form of a criticism against Frege's and Russell's theories of identity, against treating logical identity as a case of a mathematical one. In the middle 1920's this criticism extended also to Frank Ramsey's attempts to change Russell's notion of identity to accommodate infinite sets. According to Wittgenstein, the treatment of identity requires a radical revision of logic and the foundations of mathematics by reconsidering their place in relation to the world. Failure to perform such revisions will result in identity remaining nonsense (*Unsinn*).

Frege's logicist project in part consisted in reducing arithmetical equations to logical propositions. To accomplish this task, Frege introduced '=' into his logic to express truth–functions. It is rather unproblematic in arithmetic when it serves to denote binary equations, but it seems to imply ambiguity in ordinary language.[154] Consider this proposition and its constitute parts:

(1) 'The morning star is the evening star.'

If the identity between (1a) 'the morning star' and (1b) 'the evening star' expressed by the 'is' (or by '=') signifies identity between signs (as it does in arithmetic), then it would not convey any information (meaning) because in logic, just as in arithmetic, one can ascribe any meaning to the sign. In logic it would mean the relationship between an object and its name is random. If, on the other hand, the identity between (1a) and (1b) is that of objects, then (1) seems to be the same as (1*) 'The morning star is the morning star' or (1') 'The evening star is the evening star'. In this later case we have the 'law of identity' which states that every object is identical to itself, but just as in the case of (1*) and (1'), it conveys no information.

Frege's solution in "Über Sinn und Bedeutung" is to distinguish between *Sinn* (sense of the proposition) and *Bedeutung* (reference to the object, or the meaning of the

[152] Candlish (1996).

[153] Gardner, Martin and Jessell (2000).

[154] Frege (1988: §§8. 20–21).

proposition).[155] Thus, the morning star and the evening star mean the same object, viz. Venus. The (1a) and (1b) are two senses of the same object, or the same reference. The sense/reference distinction was already present in the background of the *Grundlagen* (1884), where a very similar distinction was made between number and numeral. But it was not until the "Über Sinn und Bedeutung" (1892) where such a distinction was applied to logic.[156] Although this, Frege believed, does not solve the ambiguity of the 'is', the distinction allows for meaningful denotation of the sentences.

Russell in *Principia Mathematica* (1910[1] and 1912[2]) rejected Frege's sense/reference distinction but largely adopts Frege's notion of identity with one adjustment: names are the shortcuts of definite descriptions.[157] For instance, (1a) is the name of 'The planet visible in the morning sky'. Russell's notion of identity 'a = b' comes as

(2) $(x = y) = (\varphi x \supset \varphi y)$[158]

In (2) it is implied that every class is defined as a propositional *function* (φ) and that 'x = y' if and only if there is φx which is also φy. For instance, (2) can be described by

(3) "Scott is the author of *Waverley*".

In the TLP, Wittgenstein argued against the definitions of identity of Frege and that modified by Russell. Both theories of identity, according Wittgenstein, were the products of the same confusion. The confusion is to consider the arguments of functions as the indices of names. The difference is that symbols and functions can be arbitrary, while names reflect their relationship with the objects and depend on how things are in the world (TLP 5.02). The result of this clarification of confusion is Wittgenstein's elimination of the identity in the TLP: "Russell's definition of '=' is not enough, because you cannot say with it that two objects have all their properties in common. (Even if this statement is not correct, it does make sense)".[159]

With this criticism of Russell's (and Frege's) definition of identity Wittgenstein found an ally in Frank Ramsey. For Ramsey, however, the problem with this definition of identity was different. Ramsey wanted to include infinite indefinable classes that the logic of *Principia Mathematica* would not allow. Ramsey's solution was to introduce a different definition of identity without '='. His suggestion would be that 'x = y' can be rewritten as

(4) $(\varphi_e) \cdot \varphi_e x \equiv \varphi_e y$

This definition of identity treats it as a function between variables. Wittgenstein in his letter to Ramsey rewrites (4) in the following way:

> You define 'x = y' by

---

[155] Frege (1967: 143–162).

[156] In fact, according to M. Beaney, this distinction was already present in *Begriffschrift* (1879) under the terms of 'content' (i.e. geometrical point) and 'modes of determination' (i.e. symbols denoting geometrical point). Beaney (1997: 21).

[157] Russell and Whitehead (1957: **66–71).

[158] Russell and Whitehead (1957: *13.01).

[159] „Russells Definition von ‚=' genügt nicht; weil man nach ihr nicht sagen kann, daß zwei Gegenstände alle Eigenschaften gemeinsam haben. (Selbst wenn dieser Satz nie richtig ist, hat er doch *Sinn*)". Wittgenstein 1974a: 5.5302).

$(\varphi_e) \cdot \varphi_e x \equiv \varphi_e y.$

$Q(x, y)$

The ground of this definition should be that 'Q(x, y)' is a tautology whenever 'x' and 'y' have the same meaning, and a contradiction, when they have different meanings.[160]

In his response to Wittgenstein's objection from 1927 Ramsey explains his amendment to the theory of identity and defends his view: "I never really meant to suggest that Q(x, y) was a way of saying that x and y were identical. I imagined that Wittgenstein had shown that it was impossible to say any such thing. I only proposed Q(x, y) as a substitute for the symbol x = y, used in general propositions and in defining classes".[161]

Among many philosophical notions on which Wittgenstein wrote, the notion of identity, perhaps, is the only one on which he never changed his mind from the TLP to OC (i.e., in PI §216). But like most of his ideas, his notion of identity was and remains to be considered controversial. The controversy consists in the fact that Wittgenstein makes a sharp distinction between logical identity and arithmetical equality.

In the *Begriffsschrift* Frege introduced the sign '=' to denote identity as a logical constant. With Frege and Russell logicians assumed that arithmetical equality is a special case of logical identity. As we will see further, in the TLP Wittgenstein would eliminate logical identity but hold that mathematics consists of equations. Mathieu Marion noticed that this view goes against the received view in logic which continues to treat mathematical equality as a case of logical identity.[162] The exception, he indicates, is that of W. E. Johnson who supported Wittgenstein's distinction.[163]

For Frege, identity is a relation neither between names nor between objects, but between a name and the object for which the name stands for. As Roger White indicates, for Wittgenstein's notion of truth, this notion of identity is unacceptable.[164] In the TLP, every proposition is a truth–function of elementary propositions. These propositions describe the actual states of affairs in the physical world. So, identity then is not just a function between names and signs but also, and most importantly, between objects. But stating identity between objects (or names) is either nonsense (i.e., x = y, or $\forall F (F(x. \rightarrow F(y)))$ or tautology (i.e. x = x, or $\exists x (x=x)$, $\forall x (x=x)$).

The existence of an object can only be shown by the use of the ostensive definition of that object. This is Wittgenstein's picture theory which is part and parcel of his ineffability of semantics thesis. In this regard, the identity between objects is nonsense and the identity of objects is tautology (TLP 4.1272; 5.5303).

In his later Cambridge lectures Wittgenstein would consistently criticize both Frege's and Russell's notions of identity. Especially in the case of Russell's use of the '=' in his use of equinumerosity.[165] His point is the same as in the NL, viz. that it is not the

---

[160] Wittgenstein (1979b: 189).

[161] Wittgenstein (1979b: 191).

[162] Marion (1995: 343–371). This view is shared, i.e., by Gödel in his "Die Vollständigkeit der Axiome des logischen Funktionenkalküls", in Gödel (1986: 102–123), and by Quine (1982: 268–273.

[163] Johnson (1922: ch. VI, §9).

[164] White (1979: 157–174).

[165] Wittgenstein (1979b: Lecture XI, 146–50).

business of logic to state the identity between objects but that of physics.[166]

Wittgenstein, considering Ramsey's rewritten 'x = y' as 'Q(x, y)', remarked that if x and y in 'Q(x, y)' have the same meaning, then 'Q(x, y)' is a tautology. If, instead, they have a different meaning, then it is a contradiction. Wittgenstein's point here is that rewriting identity in any way would not solve anything in logical terms. Identity is perfectly valid in mathematics, but it is not the job of logic to state it. It is neither the job of logic or mathematics to say whether the identity holds in the physical world or not. Any statements about the world are posterior to the facts in the world.[167]

In his letter to Russell from 1913 Wittgenstein states this point which will not be altered by his later philosophical developments:

> I want to repeat once again in another way what I wrote in my last letter about logic: All the propositions of logic are generalizations of tautologies and all generalizations of tautologies are propositions of logic. There are no other logical propositions but these. (I consider this to be definitive). A proposition like "(∃x). x = x", for example, is actually a proposition of physics. The proposition "(x):x = x.⊃. (∃y). y = y" is a proposition of logic; it is now matter of physics to say *whether anything exists*. The same is true of the infinity axiom; whether there are $\aleph_0$ things is to be determined by the experience (and it cannot determine it).[168]

For Wittgenstein, early and later, our language on infinity is subject to the same rules of language and use as any other concept. The problem of our language on infinity is that it does not correspond to any use: experience cannot determine it.[169]

In his later philosophy Wittgenstein continued to reject logical identity as mathematical equality. In this passage from PG philosopher's work must bring clarity to the problems of mathematics, but with the clarity comes pruning: "Philosophical clarity will have the same influence on the growth of mathematics as the sunlight on the growth of potato shoots. (In the dark cellar they grow several meters long.)"[170]

For later–Wittgenstein, empirical facts about objects come *before* the rules of grammar,

---

[166] Wittgenstein (1979a: 128).

[167] This view seems to be initiated by Carnap in the quoted before Principle of Tolerance (§17). While Carnap would not share Wittgenstein's distinction between mathematical and logical identity (§5, §16a), he did acknowledge the priority of the world over logic and language in his *Logische Syntax der Sprache* (§1–§2). Carnap also shared Wittgenstein's view that all propositions of mathematics are tautologies, that they say nothing about how the objects in the world are, but concern only the way in which we predicate of them (§14). Carnap (1968).

[168] „Ich will dasjenige, was ich in meinem letzten Brief über Logik schrieb, noch einmal in anderer Weise wiederholen: Alle Sätze der Logik sind Verallgemeinerungen von Tautologien and alle Verallgemeinerungen von Tautologien sind Sätze der Logik. Andere logische Sätze gibt es nicht. (Dies halte ich für definitiv). Ein Satz wie ‚(∃x).x = x' zum Beispiel ist eigentlich ein Satz der Physik. Der Satz ‚(x):x = x. ⊃.(y).y = y' ist ein Satz der Logik; es ist nun Sache der Physik zu sagen, *ob es ein Ding gibt*. Dasselbe gilt vom infinity axiom; ob es $\aleph_0$ Dinge gibt, das zu bestimmen ist Sache der Erfahrung (und die kann es nicht entscheiden). Wittgenstein (1979a: 127).

[169] On different interpretations of infinity in Wittgenstein see Moore (2011: 105–121).

[170] „Die philosophische Klarheit wird auf das Wachstum der Mathematik den gleichen Einfluß haben, wie das Sonnenlicht auf das Wachsen der Kartoffeltriebe. (Im dunkeln Keller wachsen sie meterlang)". Wittgenstein (1974b: 381).

logic and mathematics. The world as it is must be in place, so to speak, before we make sense of it in our language, logic and mathematics. Changes in Wittgenstein's semantics from picture theory to language–games did not change this presupposition of the priority of the world to language and logic. In fact, it only strengthened this imperative. In his later philosophy the relationship between objects and their names was 'formulated' by the language–games which could also include ostensive definitions of the isomorphism. However, language–games come *after* the world of physical objects and their uses show the inexpressible. In fact, the ineffability thesis of the TLP seems to be radicalized by the language–games, with more stress on the priority of the world over language, logic and mathematics.

This priority also constitutes the fundamental difference between Wittgenstein's notion of 'language–games' and von Neumann's 'game theory':

> Is mathematics a game? Argument against it. "The Theory of the game is <u>not</u> arbitrary although the game is". The theory of the game as pure mathematics and physics. Can we say that the fact that you can't mate with … rests on certain physical and certain mathematical facts? Can we say that the possibility of proving so and so in such and such a way rests on a mathematical, logical fact? Great temptations. This of course, restating our old problem.
>
> Suppose we said: it never happens that A mates B with …. This should the more modest propositions. But what does it mean? But couldn't we say: it never happens that we <u>say</u> A … B …? Certainly and this is a very important <u>fact</u> but based on what? So … is the theory of the game arbitrary?
>
> "I believe that Godlbach's theorem will come true". How is this belief in the end verified? By a proof. By any proof? No. By this particular proof? No. By something we shall recognize as a proof. But isn't the fact that such and such a proof is <u>possible</u> based on a mathematical fact, a mathematical reality? I mean the fact that there is a proof at least somewhere in the region we still recognize as that of proofs? The mathematical fact being that such and such a structure is possible. That it is imaginable. How do we imagine this possibility? What is a structure like which is impossible. Possible = describable.[171]

Pietarinen indicates that Wittgenstein was well aware of the game theory as a mathematical theory,[172] but he also made a clear distinction between von Neumann's project and his own: "Formalists speak of mathematics as a game. Frege remarked that the formalists confuse the game and the theory of game".[173]

In von Neumann and Morgenstern the notion of a 'game' comes analogically in its use in mathematics. There is no, strictly speaking, any game in mathematics nor could there be because of the determining factors of the rules over strategies. Instead, in the game

---

[171] Wittgenstein (2000: MS 161: 15r–16v, 1 Jan. 1939). All emphases are the author's.

[172] Pietarinen (2006: ch. 8).

[173] Rhees (2002: 9). Wittgenstein's reference to Frege most likely comes from his quotation of Thomae, in Frege (1950: II §95): "For the formalist, arithmetic is a game with signs which are called empty. That means they have no other content (in the calculating game) than they are assigned by their behaviour with respect to certain rules of combination (rules of the game)". Translation by Black (1960: 190).

theory especially applied to economics, social behavior is compared to the analogical games in mathematics: "The typical problems of economic behavior become strictly identical with the mathematical notions of suitable games of strategy".[174] However, the greatest problem in the game theory remains the obvious mismatch (disanalogy) between determining the nature of mathematical rules over the modality of human behavior and decision–making. Language–games, on the other hand, do not have this problem of mismatch: limited rules generate infinite strategies and the games do not have "to cover all these possibilities, all intermediary stages, and all their combinations".[175] This comes, however, at the price of not being a 'theory'.

The same as of the game theory can be said about the very new interdisciplinary field called 'neuroeconomics' where game theory is used in applications with neuroscience and psychology in an attempt to explain human decision making and, hopefully, to elucidate the influence of the economic behavior on the brain: "The goal of neuroeconomics is to produce a single unified model of human decision making that spans the economic, psychological, and neuroscientific levels of analysis".[176]

However, these theories treat the notion of a game differently from the way it was treated by Wittgenstein. Unlike game theory, Wittgenstein never even attempted to devise a general solution to the problem solution in human behavior, to "the common behavior of humankind" (Anscombe). The notion of language–games, in fact, essentially resists such an approach.

Returning to Wittgenstein's notions of sensations, we can ask the legitimate question from the point of view of the philosophy of mind, whether sensations are physical or mental in nature. Ultimately, it is the metaphysical question on the nature of the mind and body.

Assuming that sensations are physical events leads to the problems of qualia. These problems indicate that the physical states' corresponding mental states cannot alone explain the qualitative characteristics of a mental event. Assuming that sensations are mental events we would have to suppose that sensation must have some mental image corresponding to it. The problem here is that very often sensations (and feelings) do not generate any mental images. Pain, perceiving a red spot, raising a hand, understanding a concept, intending something, – presume no presence of any mental image. But even if they would, mental images would not explain the sensations that we have.

To speak of the mental events as if they are physical ones is a conceptual mistake. But so is identifying mental with physical. Using Wittgenstein's own analysis of the language–games, mental and physical are two parts of the same structure of the game and neither by itself is sufficient in characterizing the game as a whole.

Physical events serve as the foundational basis for the entire game. These are the rules of the game. They constitute the strategy and define the game as a whole. Mental events, on the other hand, are the strategies of the game and are defined by their rules. As it was explained before, rules are often composed by the strategies and in view of the strategies. Because of that we cannot speak of one, linear causation of the strategies by the rules.

---

[174] von Neumann and Morgenstern (1953: 1.1.2).

[175] von Neumann and Morgenstern (1953: 2.2.3).

[176] Glimcher (2011: 393).

The process is a two–way relationship between rules and strategies.

Sensations are processes, but intentions and beliefs are not. Intentions and beliefs and other conscious states cannot be defined by their (physical) rules, unlike sensations. Although they have their physical counterparts, these counterparts are entirely dependent on the strategies, on the conscious events. The difference between sensations and conscious events consists in changing the aspect of our cognitive attention from strategies being defined by the rules (sensations) to rules being defined by the strategies (consciousness).

By this conceptual picture, both physicalism and dualism must be refused. To accentuate the rules alone is meaningless. First, because rules are never formed in isolation from the strategies; whether in sensations or in conscious states. Second, because rules are not properties of strategies nor are strategies properties or functions of rules. Strategies are implementations of rules in infinitely many ways. Rules, on the other hand, are finite, syntactical steps in defining the *modus operandi* of the strategies. Rules are not descriptions of how strategies function. Whether rules are prior or posterior to strategies makes no difference to the rules themselves.

In this regard, to speak of rules and strategies in physical and mental terms has no meaning. They are two sides of the same coin. We can speak of one or the other side of the coin but we cannot speak (at least not meaningfully) of a one–sided coin.

Wittgenstein never went so far as to abandon dualistic language in his philosophy. But it is clear that he refused both physicalism and dualism of substance and of property. His notion of identity as tautology or as nonsense can be applied this way to his philosophy of mind. Identity is a matter of physics, viz. a matter of how things are in the world and not a matter of logic or of mathematics. How things are in the world are shown by the use of terms in the language–games we play continuously.

In the list of different senses of identity at the beginning of the previous chapter, all formulations of identity are that of mathematical equality, not of logical identity. But let us assume that univocal identity, written as 'x = y', will be rewritten by Ramsey and Wittgenstein as logical identity with Q(x, y). This means that x and y share equally in some property Q. Defining x as some brain states and y as corresponding mental states, Q will be defined as some state in general, i.e. pain, thinking, perceiving red, etc. Property Q is that which is shared equally by x and y. But 'Q(x, y)' cannot tell us of the relationship between x and y. We still must refer to the world of actions in which their relationship (between x and y) and their property Q is realized. In other words, the analysis of the 'Q(x, y)' must be done *outside* of the 'Q(x, y)'. The inconsistency of the so called 'neutral monism' in the works of E. Mach[177], W. James and B. Russell[178] is expressed in their inability to explain (among other problems) what is that relationship between physical states and the neutral reality, be it 'pure experience'[179], information[180], or 'protophenomenal property'[181].

---

[177] Mach (1918).

[178] Russell (1954).

[179] James (1996: 1–38. 39–91).

[180] Sayre (1976: 160).

[181] Chalmers (1996: 154).

If we consider that any mental and physical act is identified with the rules and strategies of the language–game, then we can assume that Q is that language–game, while x is the finite set of its rules and y is a potentially infinite set of its strategies. But even in this case 'Q(x, y)' by itself tells us nothing about how rules are related to their strategies (in a given game) and vice–versa. An identity considered as an identity *between* x and y still remains a contradiction. While an identity considered as that *of* x is a tautology.

Reconsidering physicalism or any kind of dualism under the Wittgensteinian notion of identity is either a tautology or a contradiction. And if this is how things are with identity, nature, existence of mind and body, and our ways of knowing and speaking of them, then the four questions defined at the beginning of Chapter I and the notion of identity defined at the beginning of Chapter II are meaningless. So are the theoretical foundations of the philosophy of mind. I will explain this point in the following conclusion to this chapter.

How should we speak of the dynamism in the language–games? Wittgenstein will be of no help in answering this question. There is very little one *can* say about the language–games according to the ineffability belief of our author. His philosophy is a therapy whose goal is to free us from the 'bewitchment'[182] of false analogies, metaphors, identities and tautologies. But if we insist on a more theoretical approach to the questions of cognition, we must look to a different philosophical tradition which allows us to express the analysis of games. In the introduction to this chapter I suggested to look into Peirce's philosophy of mind for such an approach. This will be the topic of the next chapter.

Conclusion: The Five Paradigms in the Three Philosophical Dimensions

Leiber,[183] Proudfoot[184] and Copeland[185] suggest considering the entire history of philosophical psychology and the philosophy of mind as a set of five paradigms. They insist that the continuity between the two sets is much closer than the contemporary philosophers of mind are aware of (thus, the authors of the second set of paradigms).

In this conclusion I will follow their method of investigation of these paradigms and Wittgenstein's refutation of them. The paradigms are of ontology, metaphysics, epistemology, semantics and identity. The 'old' paradigms refer to the philosophical psychology covered in Chapter I, the 'new' paradigms refer to philosophy of mind covered in Chapter II. While the methods are borrowed, much of the conclusions and interpretation of the paradigms and their refutation by Wittgenstein will be often my own. Of course, the presentation of the paradigms and Wittgenstein's responses to them are too superficial to call them an analysis. It serves rather to summarize many points given in the preceding three chapters.

1. The ontological question according to the old paradigm asks whether the other minds exist. Being a thinking thing without a clear relation to one's own body, one

[182] Hintikka and Hintikka (2002).

[183] Leiber (1991).

[184] Proudfoot (1997).

[185] Proudfoot and Copeland (1994).

cannot rely on the physical manifestations of other's minds. This substance dualism necessitates some sort of analogy between oneself and others.

According to the new paradigms, the question of the other minds transforms into the question of whether machines can think. While the old paradigm applies analogy, the new paradigm implements the Turing Test. If a computer is capable of imitating some human behavior, then it passes the Test, at least to some degree. But both answers are essentially resting on the same notion of analogy. If my behavioral attitudes are similar to other's people behavioral attitudes, then other people have minds too. If a man can imitate a woman, then a computer can (at least in some features) imitate a human being.

Wittgenstein refused both paradigms' uses of analogies by insisting that human behavior and thinking cannot be reduced to rule–following alone. As to the old paradigms' skepticisms of the other minds, such doubt for Wittgenstein is excluded as a grammatical mistake not only by the grammar of our language, but by the grammar (set of rules) of our entire existence. If one doubts indeed the existence of other minds, one doubts the foundation of the entire knowledge. To these arguments Wittgenstein dedicated his last book, OC.

Edward Witherspoon considers Wittgenstein's criteria as what other people say and do. These are the criteria of proof (allegedly for Wittgenstein) that other minds exist.[186] I believe that Wittgenstein's approach was much more radical than that. If 'other minds' are rooted in the grammar of the world, there is no need for criteria of any proof in their existence. Was this not the whole point in Wittgenstein's debate with G. E. Moore in the OC? Again, the point is not to answer philosophical questions and solve philosophical problems but to show that these questions and problems originate from our pathological misuse of language. Once we apply Wittgenstein's therapy we will see the world rightly (TLP 6.54).

2. Both new and old paradigms consider mind as a thinking thing. What for the old paradigm is *psuchê*, *anima*, *soul* and *mens*, for the new paradigm, for most part, is a thinking computer, machine. What unites these paradigms is that both are distinct from their physical realizations. Philosophical psychology treats the soul as separated from the body; functionalist theories of mind speak of multiple realizability of the mental states in different, sometimes unrelated to each other physical states.

Wittgenstein's response is a denial that mind is some phenomenological entity either (somehow) linked to or separated from the physical body. Mental states for Wittgenstein are inseparable from their manifestations in the actions of human beings. His treatment of the mind/body dichotomy is based on his overall criticism of metaphysics.

In the TLP the idea of metaphysical inquiry contained the relation between language or thought, and the world. Wittgenstein's main criticism against metaphysics in his earlier philosophy consisted mainly in rejecting the idea that thought and language can be about the world. He thought it to be part of the major philosophical confusion of treating formal properties as if they were ordinary properties. This is what he mains by our attempts to say that which can only be shown (4.121).

But TLP itself is a collection of nonsensical propositions since Wittgenstein by means of semantic isomorphism is *saying* that (ordinary properties of objects) which can only

[186] Witherspoon (2011).

be *shown*. The main point of the TLP was to provide a general logical description of the world. Such (phenomenological) description would accommodate any proposition, actual or potential in our language. Since 1929 Wittgenstein came to understand that such logical generality was preventing us from seeing the complexity of physical language and the multifarious richness of our ordinary experiences.

In this later period, as noted by Cora Diamond,[187] his critique of metaphysics shifts into the direction of critique of philosophical language. We have seen that the passages from PI §§243–315 are excellent examples of such critique. Metaphysical questions arise from misunderstanding the ways the grammar of our experiences operates. Just as learning the rules of grammar allows us to use our language in infinitely possible ways, understanding the rules of our language–games (activities) permit us to comprehend the meaning of our ordinary language, share our experiences, sensations and intentions. Sharing a common grammar is a condition for sharing common meaning. This translates directly to the uses of language–games. Once we learn the grammar of our experiences, metaphysical questions fade as something redundant.

3. Both paradigms (with the exception of direct realism) assume some kind of representational theory of perception and of knowledge of the external world. In the cases of the old paradigm it consists in speaking of ideas, impressions and forms. In the cases of the new paradigms it includes sense–data, neurally instantiated internal representations, cortical maps and the like. Both paradigms are very close to each other by the fact that both instantiate skepticism of the representation of the external world, or at the least begging the question of how reality is to be related to these representations.

Wittgenstein claimed that the logic of perception and the structure of the world are in harmony. His major criticism was directed against internal imaging as perception. The Private Language Argument refuses any sort of phenomenological objects as well as any phenomenological language about them. These points were strengthened further in his last writings where sense–data and other phenomenological objects were eliminated entirely. Epistemological and semantic inquiries come as unified one in Wittgenstein's treatment of language–games.

4. Contemporary theories of meaning were initiated by the 17th century interpretation of language as the medium of expressing the internal representations. Ostensive definitions were already implemented in a naïve theory of language and meaning by St. Augustine (PI §1) and extensively applied until the early 20th century. Contemporary computational theories are divided between internalism and externalism. Both have the same question: Where is meaning situated? Both have opposite answers: cognitive acts for the earlier, world and social constructions for the later. Theories of intentionality according to the new paradigms are parallel to the theories of meaning according to internalism or externalism. Neither paradigm can offer certainty of correspondence between language and reality. Thus, the old paradigms postulate an omniscient God who guarantees reality according to our minds and language; the new paradigms refer to the cognitive neuroscience's brain maps that show the hard drives of our cognition.

Wittgenstein's approach to semantics is radically opposite to the accounts of both paradigms. His refutation of his own *Tractarian* semantics came on par with refuting that

[187] Diamond (1991: 13–38. 73–94); Diamond (2009).

meaning is a mental act or a hard–wired process in the brain. Unlike sensations, thinking, understanding, intending, meaning are not processes. His notion of the language–games as representative links between language and reality provides a dynamic approach to every human action and the possibility of analysis of it in terms of analysis of rules and strategies of the language–games.

5. The old paradigms could not have the notion of identity attached to the mind/body dichotomy due to its essentially dualistic understanding of human nature. Furthermore, the notion of identity is missing in propositional logic prior to Frege (with the unique exception of Leibniz). The new paradigms applied every notion of identity available in logic and mathematics. Identity of equality in mathematics became a part of logical identity. Distinctions between these two can hardly be made in contemporary quantificational logic and are presumed to be the same in the philosophy of mind. Alas, knowledge of identity in logical and mathematical languages tells us little about the identity of body and mind. What makes them identical and in what consists their identity are the main unsolved problems within the personal identity theories.

The distinction between logical and mathematical identities was strongly emphasized by Wittgenstein in his critique of Frege's, Russell's and partially of Ramsey's treatments of identity. Such a distinction is important in logic and mathematics, but no notion of 'identity–of' or 'identity–between' will help in understanding the relationship between mental and physical states. Wittgenstein insists that such relationship is provided by the use of language and by the activities we undertake. Instead of the theory of identity establishing an identity between mental and physical states, we are asked to see their implementations in the language–games in which mind and body, language and reality are inseparable constitutive parts of the meaning–generating grammar. Such meaning renders past and present to Wittgenstein's philosophy paradigms resting on the old inquiries and assumptions meaningless and redundant.

CHAPTER IV

C. S. Peirce's Semeiosis of Cognitive Continuum

Introduction

C. S. Peirce's interest in theoretical and experimental science brought him to revise the foundations of logic to make it the appropriate method of scientific inquiry. Logic serves to facilitate scientific discovery, and unlike in Frege's and Russell's logicism, the foundations of logic are mathematical. Peirce's revisions of logic begun in 1865[1] and inspired him to provide the theory of signs, called semeiotic in 1868.[2] Consequently, his theory of signs prompted him to think about some mathematical questions concerning infinity and continuum. Apart from the questions of scientific discovery, the Cartesian problem of mind/body and the emergence of a new physicalistic approach in the philosophy of science was the major reason for Peirce's theory of signs to develop.

Peirce had four periods of semeiotic and five periods of the development of his theory of continuum. In this chapter I shall argue that the developments and changes in semeiotics influenced changes and revisions in the theory of continuum. However, both topics must always be read within Peirce's desire to design a new logic for science and radically reform Cartesian dualism in metaphysics, ontology and epistemology.

This chapter will be divided into three sections. The first section will treat Peirce's four periods of semeiotics, the second section will treat his five periods of the development of the theory of continuum. The last section will focus on how these two theories were of fundamental foundational importance for his non–dualistic and non–monistic philosophy of mind. As was done in the previous chapter, this chapter's conclusion will sum up Peirce's approaches to the five questions of ontology, metaphysics, semantics, epistemology and identity.

1. Semeiotics and Semeiosis: Theory of Signs and their Uses

C. S. Peirce was notoriously inconsistent in the uses of his own terminology. This is particularly clear in his semeiotics. Since he held to a triadic doctrine of the signs, Peirce

---

[1] "On the Logic of Science". 1865–1866 in Peirce (1982: 162–504).

[2] "On the New List of Categories" in Peirce (1984: 49–58).

would create and change more than once the names for signs, indices, symbols, etc. In this chapter I will not give much attention to these terminological nuances but rather focus on the development of his semeiotics.

However, one note is necessary to clear the issue of what his theory of signs should really be called. Here too Peirce was anything but consistent. Throughout his philosophical writings he would use the names of 'semeiotic', 'semeiotics', 'semiotic', and 'semeotic' interchangeably. Some authors, i.e., Max H. Fisch[3] and Thomas L. Short[4] would insist that 'semiotic' is the most accurate way of naming Peirce's theory of signs. First, it clearly distinguishes Peirce from de Saussure's 'semiology', a dyadic theory of 'signifier' and 'signified', widely used in constructivist linguistics. Second, it distinguishes Peirce from contemporary 'semiotics' much founded on de Saussure's 'semiology', a discipline that studies signs in communication. The major difference between Peirce's theory of signs on the one hand and 'semiology' or contemporary 'semiotics' is Peirce's theory being a cognition–oriented theory of logic with important implementations for scientific inquiry.

In this chapter I use 'semeiotics' as a noun form and 'semeiotic' as an adjective to make sure that no association with 'semiotics' or 'semiology' is made. Peirce himself borrowed the term from J. Locke's last pages of the *Essay*, quoted at the end of the chapter. Thus, 'semeiotic' is from *σημειωτική*, 'theory of signs', from the Greek *σημείωσις*, 'observation of signs'. Likewise, I will use 'semeiosis' instead of 'semiosis' throughout this work. Despite Peirce's own inconsistency in the use of these terms, the difference between contemporary 'semiotics' and de Saussure's 'semiology' must be kept in mind.

Another inconsistency in Peirce as well as in the secondary literature is between 'pragmatism' and 'pragmaticism'. Peirce wanted to distance himself from William James' 'pragmatism' with his own term 'pragmaticism'. In this paper I use the later.

1.1 The First Account (1865 – 1877)

1.1.1 *Origins of Semeiotics*

Peirce's earlier semeiotics derived from Kantian epistemology and metaphysics, but the use of signs goes back to Locke's understanding of semeiotics in which words were treated as signs of ideas.[5] However, as it was explained in the first chapter, Locke did not properly distinguish between thoughts and ideas and, although this distinction was introduced by Hume, the empiricist treatment of both relied too much on the strength with which ideas and impressions come to our mind and on the controversial theory of causation. Ultimately, both explanations rely on the habits and patterns of our experiences (mind as a set of *bundles*).

Thus, Peirce's semeiotics was Kantian in the sense that his treatment of thoughts and ideas was conceptual and categorical. Thoughts come as an internalized discourse or an inner language, being *conceptually* different and posterior to impressions. Thoughts, like

[3] Fisch (1978).

[4] Short (2007: ix. 16–21).

[5] Locke (1975: IV, 21 §4).

language, are signs standing in interpretative relation to the objects and their impressions in our minds.

For Kant there are two kinds of objects: those that are constituted by thought and those that are not. The first are knowable objects: phenomenological (or mental) and physical. The second are not. These are the things–in–themselves.

The difference between Kant and Peirce is that for the later there are in principle no unknowable objects. From the late–1860s Peirce endorsed an essentially idealistic approach by claiming against Kant that there are, in fact, no such objects not constituted by thought.

Peirce was never fully satisfied with, what he called, 'objective idealism'. Changes and transformations in his philosophy brought him further from Kant and idealism toward pragmaticism as well as toward new forms of his semeiotics.

1.1.2 *"On the Logic of Science". Harvard Lectures of 1865*

Peirce's semeiotics comes as a result of Peirce's earliest attempts to establish an objective foundation for logic freed from mentalism and subjectivism. This project would be also evident in the later works of Frege, especially in his *Grundlagen*. In order for such a project to succeed, Peirce believed that we must delineate clearly what logical reasoning, inference making, is about. It is about thought that comes from observation. However, to limit thoughts and ideas to an exclusively mentalist operation means to fall into the same mistake as we have seen in Locke and Hume. It was Kant who began the process of liberating logic from the purely mental domain but it was Peirce who insisted that ideas and thoughts are about the mind as much as they are about the external objects.

Peirce begins by making the first triadic distinction between thing, representation and form as epistemological foundations of any logical inference. The central term here is 'representation' which Peirce elaborated from Kant's use of the term.

In his first *Critique*, Kant used the word *Vorstellung* (idea) in relation to any mental content. In his *Harvard Lectures* of 1865 Peirce translated Kant's *Vorstellung* as 'representation' and 'performance'. Perhaps the origin of this secondary dictionary use of the word that was picked up by Peirce is related to the verbal form *vorstellen*, 'to put forward', 'to represent', 'to mean'. One cannot limit representation or performance,[6] nor one can put forward only ideas but any object in general. Thus, Peirce refused to limit the act of representation only to mental contents.

In "Lecture VIII: Forms of Induction and Hypothesis" (April–May 1865) Peirce introduced for the first time his triadic notion of representations in terms of signs, copies and symbols:

> By a *copy*, I mean a representation whose agreement with its object depends merely upon a sameness of predicates. By a *sign*, I mean a representation whose reference to its object is fixed by convention. By a *symbol* I mean one which upon being presented to the mind – without any resemblance to its object and without any reference to a

[6] Recall the preference of translating *Sprachspiel* as a language–performance rather than language–game. The verb *spielen* has also the meaning of acting and performing.

> previous convention – calls up a concept. I consider concepts, themselves, as a species of symbols.[7]

The triadic distinction between signs, copies and symbols reflects not only the world of reasoning, thinking and argumentation but, most importantly, the world in which we live. It is a world of representations and its similarity with thinking allows us to grasp the truth about the mind–independent world. This correspondence is essential for Peirce's epistemology and semantics, and will not change throughout the years of his philosophy.

1.1.3 *"The Logic of Science; Or, Induction and Hypothesis". Lowell Lectures of 1866*

One year later Peirce indicated the essential feature of semeiotics: the relations between the thoughts as signs. The thought or an idea is the representation of something *as* something:

> We are, then, sufficiently familiar with the fact that many words have much implication; but I think we need to reflect upon the circumstance that every word implies some proposition or, what is the same thing, every word, concept, symbol has an equivalent term – or one which has become identified with it, – in short, has an *interpretant*. (…) Now a representation is something which stands for something.[8]

The interpretant is not another person to whom the object is represented, but one representation is presented as another to the same person. One thought is interpreted (represented) by another thought, is the semeiotic process of objects being interpreted as signs and signs being represented by the interpretant.

As we had seen, Frege's distinction between *Sinn* and *Bedeutung* rests on the direct signification or on the direct relationship between the sign being the word (*Sinn*) and the object or the reference (*Bedeutung*). Wittgenstein's criticism of this directness brought him to his earlier picture theory and later, much more complex relationship between the sign and the object, known as the language–games. For Peirce, like for Wittgenstein and unlike for Frege, the relationship is never direct but mediated by the interpretant. The meaning of the sign is its use in a language–game; the meaning of the sign is found in the interpretant.

Meaning is a triadic relation between sign, object and the interpretant. But the interpretant itself is a sign in relation to the idea, the thought, represented. Peirce's earlier notion of the infinitum rests on his earlier semeiotics.

1.1.4 *"On the New List of Categories". 1867*

In 1867 Peirce presented a paper to the American Academy of Arts and Sciences entitled "On the New List of Categories". There Peirce follows Kantian metaphysical

---

[7] Peirce (1982: 257–258).

[8] Peirce (1982: 466).

deduction which says that the form of a judgment corresponds with the form of its concept which is formed from the sensory data in experience. The transcendental deduction further affirms that such judgment formation is the only way to acquire any knowledge. However, this agreement with Kant is limited to this acceptance of the metaphysical and transcendental deductions. The novelty of Peirce in that paper consisted in explaining through the five steps how from the senses judgments are formed. His triadic relation of the sign played an essential role in that process of judgment formation.

Peirce suggests three new categories (which later will be called Firstness, Secondness and Thirdness) as an alternative to the Kantian categories. The first category is that of likeness "whose relation to their objects is a mere community in some quality". the second is index or sign "whose relation to their objects consists in a correspondence in fact". and the third is called general sign or symbol "whose relation to their objects is an imputed character".[9]

The twelve Kantian categories are transformed, rather than reduced, into three: quality, relation and representation. These new categories have likewise three modes of representations: likeness, signs and symbols.

The most interesting part of this semeiotics is how Peirce treated the relation between the signifying element and the interpretant. The later comes as a sign itself, as the product of the interpretation of the object or as a consequence of translation of the object into a sign. Thus, interpretant appears to be a sign in need of further interpretation that will inevitably generate new signs (new interpretants) themselves in need of further interpretations, and so on *ad infinitum.*

There are three possible ways in which this infinite generation of signs can occur. First, qualitatively by generating more information on the object. These are called 'icons' or, in the earlier account, 'likenesses'. Icons are representations of their object by virtue of some shared physical quality, by some similarity between the two. Peirce gives the example of the similarity between the letters 'p' and 'b'.

The Second is by correspondences to the facts. These are called 'indices'. An interpretation resulting in indices would come as a result of some sort of physical correspondence between two objects. Peirce's examples include the relationship between a murderer and a victim.

The Third is by some assignment of the character to the object. These were called 'symbols'. The last interpretation comes as a result of some convention between the sign and its object. An example is given between 'homme' and 'man' having the same reference.

Since every interpretant is itself a sign of its object, the main focus of the earlier semeiosis is on cognitive activity. Peirce calls these sorts of signs 'thought–signs', for which symbols have primary importance. In fact, in this earlier account Peirce was rather dismissive of icons and indices. The reason being that only symbols, being general and conventional, were fitting candidates for his new system of logic.[10]

But there were more general reasons to prefer symbols to icons and indices in that period. Just as Peirce was eager to reform Kantian metaphysics, he was even more inclined to challenge the Cartesian method of philosophy. The foundationlist account of

---

[9] Peirce (1984: 56).

[10] Peirce (1984: 213), from 1868; Peirce (1984: 439), from 1870.

the Cartesian *cogito* that was supposed to serve as the basis of all knowledge (and itself only found in the idea of a non–deceiving God) was the target of Peirce's efforts.

Symbols were meant to provide the infinite interpretation of signs and this made not only logic but semeiotics as well as any scientific investigation infinitely open to further discovery and amendment. Not only there was no need for any foundationalist epistemology, the goal was not having it at all.

If a sign (like *cogito*) has some final interpretation which itself is not subject to any further interpretation, then a non–interpretative sign is a contradiction in terms. The infinite semeiosis was not seen as faulty by Peirce but as a necessary goal; it is not a weakness of his epistemological foundation of logic but its strength.

As T. L. Short indicates, in the 1890s Peirce would radically reform his own conception of categories by denying that there can be any *a priori* demonstration of the judgment formation: "each category has to justify itself by an inductive examination which will result in assigning to it only a limited and approximate validity".[11] However, it is important to state that from 1867 Peirce considered his semeiotics as part and parcel of his metaphysics. The notions of semeiotics and categories will change but their mutual relationship will persist.

1.1.5 *Articles in the Journal of Speculative Philosophy, 1868 – 1869*

Peirce's earlier account of semeiotics culminated in his three articles for the Journal of Speculative Philosophy written in 1868 and in 1869. In them his criticisms of Cartesian foundationalism continued to be the central theme.

The first article "Questions Concerning Certain Faculties Claimed for Man" begins with Peirce's criticism of the Cartesian notion of intuition as the sole foundation for affirming the subjective certainty of the mind. Such certainty would not depend on the external world but on the indubitable intuition that is in an immediate referential relationship with the object.

Peirce denies that cognition is intuition at all. Instead, cognition is a series of thoughts interpreting previous thoughts. These thoughts, as we have already seen earlier, are in signs and thoughts are interpreted as signs:

> From this proposition that every *thought* is a *sign* it follows that every thought must address itself to some other, must determine some other, since that is the essence of a sign. And yet this after all is but another form of the old axiom, that in intuition, i.e. in the immediate present, there is no thought. Or to put the thing in another familiar form all that is reflected upon has past.[12]

Here Peirce puts his finger on the real problem with the Cartesian notion of intuition: sharp individuation of the thought. In *cogito* as in the act of thinking, there is no *cogito* without my thinking that was before and without my thinking that follows the *cogito*.

---

[11] Peirce (1931: 303), from 1894; See Short (2007: 31).

[12] Peirce (1984: 173).

Cartesian '*cogito*' results in being an impossible exercise, an abstraction by individualization from what Wittgenstein called "the stage–setting in the language" (PI §257).

Peirce further insists that this chain of thoughts–signs is an infinite process of one thought interpreting the other:

> The paradox here is similar to that of motion. The Zenonian may say no thought can be accomplished if there must have been a thought since every thought. But the contradiction here is a merely formal and not a real one. Since any time in the past there have been an infinite series of times. It is only at a date that there has not been an infinite series of times since that date. Now what is here said is that thought cannot happen in a date, but requires a time. That is only another way of saying that every thought must have been interpreted in another thought.[13]

Thus, every thought is determined by the previous thought and in this way he affirms that each thought is both a sign and an interpretant. The infinitum extends toward past thoughts–interpretants and toward the future thoughts–signs.

In the second article "Some Consequences of Four Incapacities" Peirce extends his theory of thought–signs to the theory of meaning as translation. What gives meaning to a thought is the capacity of every thought to be translated into a different thought. The meaning is not tied to its immediate referent (as it was in Descartes and as it will be in Frege) but in it being connected by the further representations: "No present actual thought (which is a mere feeling) has any meaning, any intellectual value; for this lies not in what is actually thought, but in what this thought may be connected with in representation by subsequent thoughts; so that the meaning of a thought is altogether something virtual".[14]

The notion of virtuality will be explained later in this chapter when Peirce's theory of virtual identity will be considered. At this point it is important to point out the connection between virtuality of meaning and continuum. This becomes clear from the remaining part of the paragraph quoted above:

> At no one instant in my state of mind is there cognition or representation, but in the relation of my states of mind at different instants there is. In short, the Immediate (and therefore in itself unsusceptible of mediation – the Unanalyzable, the Inexplicable, the Unintellectual) runs in a continuous stream through our lives; it is the sum total of consciousness, whose mediation, which is the continuity of it, is brought about by a real effective force behind consciousness.[15]

One thought is different from the other (previous or future), but if each is codependent on the other there must be some kind of identity of continuity between them. This identity is virtual, not individualized, and it is this virtual identity that guarantees that all thinking is a continuum.

This notion of identity and continuity brings Peirce to his complex criticism of nominalism. Nominalism here is understood as affirming that only individuals exist while

---

[13] Peirce (1984: 173).

[14] Peirce (1984: 227).

[15] Peirce (1984: 227).

universals (concepts) exist only in name. Although Peirce's refutation of nominalism is consistent throughout his philosophy, it comes clear in relation to the continuous infinitum of thoughts in the third article "Grounds of Validity of the Laws of Logic: Further Consequences of Four Incapacities":

> All the arguments of Zeno depend on supposing that a *continuum* has ultimate parts. But a *continuum* is precisely that, every part of which has parts, in the same sense. Hence, he makes out his contradictions only by making a self–contradictory supposition. In ordinary and mathematical language, we allow ourselves to speak of such parts – *points* – and whenever we are led into contradiction thereby, we have simply to express ourselves more accurately to resolve the difficulty.[16]

The conclusion that Peirce draws here is that there are no individual objects, Kantian things–in–themselves, but only things–as–we–know–them, presented to our cognition over time through representations of signs. If there is a correspondence between the world and our knowledge of it, just as there are no individual things, there is no individual consciousness to which things are represented. All three, therefore, things, consciousness and representation are in an infinitely extended continuum:

> But even if there were such an orderliness in things, it never could be discovered. For it would belong to things either collectively or distributively. If it belonged to things collectively, that is to say, if things formed a system, the difficulty would be that a system can only be known by seeing some considerable proportion of the whole. Now we never can know how great a part of the whole of nature we have discovered. If the order were distributive, that is, belonged to all things only by belonging to each thing, the difficulty would be that a character can only be known by comparing something which has it with something which has it not. *Being*, *quality*, *relation*, and other universals are not known except as characters of words or other signs, attributed by a figure of speech to things. Thus, in neither case could the order of things be known. But the order of things would not help the validity of our reasoning – that is, would not help us to reason correctly – unless we knew what the order of things required the relation between the known reason *from* to the unknown reasoned *to*, to be.[17]

From this passage it is clear that the author desires to bring to our attention the remarkable unity of the world and cognition. This idealism of Peirce sometimes is misinterpreted as our thoughts being about something unreal, since the thing–in–itself is unreal. [18] From the above passage, however, it is clear that our thoughts are about real things, but these things are not things–in–themselves. Our thoughts are in a continuum because they are about the world being likewise continuum. Peirce's refutation of nominalism is radical in that it amounts to no admission of any individual, not that an individual is considered to be general: "The upshot is that the individual that is unreal is the absolute individual, the singular or ideal boundary of cognition. Whereas, the individual

[16] Peirce (1984: 256).

[17] Peirce (1984: 265); Peirce (1934: 343).

[18] Short (2007: 38).

that we want to count as real, for example, Philip [of Macedonia], is retained; but it is general".[19]

If Peirce indeed considered individuals to be generals then not only was he clearly contradicting himself, he fails short of refuting nominalism in its most general description as an affirmation of the existence of individuals and universals as merely names.[20] On the contrary, for Peirce there are no individuals because of the infinite continuum and the virtual identity of thoughts and objects. Our concepts and our language are reflections of that kind of world.

1.1.6 *Further Elaborations, 1869 – 1877*

Between 1869 and 1877 Peirce made many attempts to elaborate on the unity between the world and cognition. However, his conviction that first, there is a strong correspondence between the two and, second, that individualization of cognition is false, remained unchanged. There are several texts that can testify of this conviction of which I shall present two: one from a 1872 manuscript and the second from 1877.

The unity between the world and cognition is affirmed at the beginning of his note "On Reality" from the Fall of 1872:

> The question is, "Whether corresponding to our thoughts and sensations, and represented in some sense by them, there are realities, which are not only independent of the thought of you, and me, and any number of men, but which are absolutely independent of thought altogether". The objective final opinion is independent of the thoughts of any particular men, but is not independent of thought in *general*.[21]

What makes reality dependent upon our thoughts? The answer is, the signs' correspondence to that which the signs represent. If the sign gives an impulse to generate the thought as an interpretant of the object for which the sign stands, then "the objective final opinion" is the goal toward which the semeiotic process is headed. Thus, there are two kinds of dependencies here. On the one hand we depend on the objects as they are presented to us upon the signs for which they stand, on the other hand we depend on our objective final opinion upon thoughts in general.

If the above inference is right, then individualization of objects as well as individualization of cognition is false. This is affirmed in Peirce's piece from the Summer of 1877 entitled "On a New Class of Observations, suggested by the principles of Logic". The text begins by an historical examination of Duns Scotus' theory of individualization and Occam's denial of existence of any generality of the objects of thought:

> It is usually admitted that there are two classes of mental representation, Immediate Representations or Sensations and Mediate Representations or Conceptions. The former are completely determinate or individual objects of thought; the latter are partially indeterminate or general objects. Granting that both these classes of objects exist, the

[19] Short (2007: 39).

[20] Forster (2011: ch. 3).

[21] Peirce (1986: 28–29).

> question of the principle of Individuation or of the respect in which the individual differs from the general becomes one of extreme difficulty.[22]

Peirce puts forward a different solution. He denies the individuality of an object over time and that even the most concrete objects are indeterminable in thought. Their determinacy can only be pointed out in "an indivisible instant of time":

> I on the other hand have undertaken to show that just the reverse of this is the case. That no object is individual but that the things the most concrete have still a certain amount of indeterminacy. Take, Phillip of Macedon for example. This object is logically divisible into Phillip drunk and Phillip sober; and so on; and you do not get down to anything completely determinate till you specify an indivisible instant of time, which is an ideal limit not attained in thought or in *re*.[23]

Notice two opposite adjectives here: "logically divisible" object and "an indivisible instant of time". Thoughts are time–transcending in that unlike sensations we cannot pinpoint when my thought started and when it ended. Wittgenstein will make the same point in his clear distinction between feelings and thoughts, and sensations. But since thoughts and sensations stand together in the chain of interpretations, the distinction made by Wittgenstein is only a logical one: thoughts and sensations are like objects, they are only "logically divisible". Thus, Peirce concludes: "It follows from this doctrine that we have no pure sensations, but only sensational elements of thought".[24]

In the previous chapter a distinction was made between events (i.e. pain) and facts (i.e. crying). We know that one can be independent from the other or that facts can belong to different events (i.e. my crying is connected with the event of lying, not with the event of being in pain). Although the distinction between events and facts is consistent, generally speaking no fact can be ever separate from *some* event. Facts are individual instances captured by the language manifesting some event, just as signs are interpretations of some objects. We can procure a "logical division" in our ordinary language: 'Phillip is drunk' is opposite from 'Phillip is sober'. But any real division is as meaningless as dividing Philip into two distinct objects. In conclusion, the division of objects is unattainable neither in thought nor in the thing itself but only in language or in logic. Or, using language–game distinctions, rules determining the event are the same while strategies extending toward distinct facts differ. Yet, even so, just as rules and strategies are inseparable, objects and thoughts are inseparable and co–dependent.

1.2. The Second Account (1877–1885)

1.2.1 *"Fixation of Belief". 1877*

Several months later Peirce will elaborate further on the distinction between logical division and actual individuation. T. L. Short suggests seeing it as Peirce's acceptance of

---

[22] Peirce (1986: 235).
[23] Peirce (1986: 235).
[24] Peirce (1986: 235).

the previously strongly rejected nominalism.[25] It would be so if Peirce would admit the existence of individual objects. In the paper from November 1877, however, he only affirmed that logical individuation or division is essential for the method of science, not that the individuation of objects is possible:

> To satisfy our doubts, therefore, it is necessary that a method should be found by which our beliefs may be caused by nothing human, but by some external permanency – by something upon which our thinking has no effect. (…) Our external permanency would not be external, in our sense, if it was restricted in its influence to one individual. It must be something which affects, or might affect, every man.[26]

In this passage, as in the previously examined ones, Peirce stresses the importance of the correspondence between the world and thoughts. Only by severing such correspondence could Peirce embrace nominalism, viz. the belief that only individual objects are real. The "external permanency". an objective reality which is fundamental for everyone's belief, is the world of continuity corresponding to the continuity of our cognition.

What was new in the late–1877, early–1878? It was the importance of logic for science and the need for an objective language to reflect the world and to correspond to the continuity of cognition:

> The method must be such that the ultimate conclusion of every man shall be the same. Such is the method of science. Its fundamental hypothesis, restated in more familiar language, is this: There are real things, whose characters are entirely independent of our opinions about them; those realities affect our senses according to regular laws, and, though our sensations are as different as our relations to the objects, yet, by taking advantage of the laws of perception, we can ascertain by reasoning how things really are, and any man, if he have sufficient experience and reason enough about it, will be led to the one true conclusion.[27]

When on the previous page Peirce says that the method must be not affected by anything human or by "something upon which our thinking has no effect". he explains it further by speaking of "our opinions" not affecting the real things. Rules of reasoning cannot depend on changing opinions but should be based upon some objectivity found in the "external permanency". What we will not find in this text is an affirmation of the things–in–themselves or of the individual objects.

1.2.2 *"On the Algebra of Logic". 1885*

The inspiration behind Peirce's change of focus from semeiotics, from though interpreting thought, to the "external permanency" was logic. The major difference was not so much in his semeiotics as in the necessity to make the language of science more direct and more specific. In the earlier 1880s he and his student O. H. Mitchell would introduce

[25] Short (2006: 225–226).

[26] Peirce (1986: 253).

[27] Peirce (1986: 253–254).

quantifiers into predicate logic, around the same time but independently from G. Frege's same logical innovation. Quantifiers 'demand' individualization in one's logic.

The result was that not only symbols but also icons and indices played an important role in his quantificational logic. Propositions for objects, that could to be picked out as definite descriptions, were treated as particular signs, or indices. Furthermore, Peirce noticed that some symbols have iconic features. This was evident especially in mathematics, i.e. in the treatment of numbers.

Peirce begun to realize that general descriptions are not sufficient in predicating propositions.[28] That very job was meant for indices to give ostensive definitions to single concepts:

> The relation of the sign to its object does not lie in a mental association, there must be a direct dual relation of the sign to its object independent of the mind using the sign. In the second of the three cases just spoken of [the sign, *the object*, and the mind], this dual relation is not degenerate, and the sign signifies its object solely by virtue of being really connected with it. Of this nature are all natural signs and physical symptoms. I call such a sign an *index*, a pointing finger being the type of the class.[29]

If we recall the development of Wittgenstein from the TLP to PI, we would notice the reverse of that of Peirce. Wittgenstein admitted ostensive definitions as part of his picture theory of meaning, inherited from Russell's denotations. While ostensive definitions maintained their importance for Wittgenstein in his later philosophy, the meaning could not be associated with simple pointing at something. For the meaning we must look into performance, actions, and language–games. For Peirce, performances or representations played an essential role since 1865, while ostensive definitions, through indices, acquired their importance only in the 1880s.

However, contrary to T. L. Short's interpretation, the ideas of the generality of signs and that signs interpreted prior signs remained in Peirce's semeiotic after 1877. Two paragraphs later Peirce affirms precisely that:

> I have taken pains to make by distinction [n.1 See *Proceedings of the American Academy of Arts and Sciences*, Vol. VII, p. 294, May 14, 1867] of icons, indices, and tokens clear, in order to enunciate this proposition: in a perfect system of logical notation signs of these several kinds must all be employed. Without tokens there would be no generality in the statements, for they are the only general signs; and generality is essential to reasoning.[30]

Peirce immediately continues by stressing that generality alone is also insufficient for our language and logic. Ostensive defining is needed for indicating, distinguishing and describing:

> But tokens alone do not state what is the subject of discourse; and this can, in fact, not be described in general terms; it can only be indicated. The actual world cannot be

---

[28] "Methods of Reasoning". from 1881 in Peirce (1989: 245–256).

[29] Peirce (1993: 163).

[30] Peirce (1993: 163).

> distinguished from a world of imagination by any description. Hence the need of pronouns and indices, and the more complicated the subject the greater the need of them.[31]

Neither generality alone nor only ostensive definition can be held accountable for expressing meaning. Peirce indicates the difference between the first and the second as that between Euler's circles and the shadings of them by Venn. The first represents relations of terms, the second adds precision. Tokens are responsible for generality, indices bring clarity and distinction.

In the first account of semeiotics the distinction between real and non–real objects could not be made. In this second account this lack is corrected through the introduction of indices. However, as we have seen in the example of the Lost Island, no amount of pronouns and indices in any description of it will make it real. At this stage of semeiotic in Peirce, the notion of meaning remains problematic.

1.2.3. *"Notes on the Categories". 1885*

This paper from 1885 comes as a further elaboration of his earlier article "On the New List of Categories" from 1867. The three categories of likeness, index and symbol were now called firstness, secondness and thirdness. In this paper Peirce particularly stresses the importance of these categories for cognition.

Categories are in monadic, dyadic and triadic relations. Likeness, or icon, is a monadic relation to its object; index is dyadic and symbol is triadic. Meaning and cognition require triadic relations in which object and sign are interpreted. While icon is only the resemblance to the object, index is the call for attention, "a pointing finger" at something. By itself it does not require any act of thinking and, Peirce insists, it does not have to be limited only to human cognition. A deer running away alarmed by a gun shot is an example of an index causing a deer's attention. Cognition, on the other hand, requires the triadic relations of icons, indices and symbols.[32]

1.3. The Third Account (1901 – 1903)

In the period between 1885 and 1901 Peirce wrote very little on his semeiotic. The first text in this period comes from Peirce's entry of 'sign' in the Baldwin's *Dictionary*. Written in 1901, it is usually considered to be his best known definition of the sign: "Anything which determines something else (its *interpretant*) to refer to an object to which itself refers (its *object*) in the same way, the interpretant becoming in turn a sign, and so on *ad infinitum*".[33]

In this passage Peirce continues to insist that interpretation proceeds ad infinitum, and so the interpretant itself becomes a sign. In this third account Peirce realizes this to be

---

[31] Peirce (1993: 164).

[32] "One, Two, Three: Fundamental Categories of Thought and of Nature". from 1885, in Peirce (1993: 242–247).

[33] From 1901 and published in 1902, see in Peirce (1932: 303).

problematic especially for the explanation of meaning. If significance is bounded to interpretation and interpretation in turn depends on the sign, then significance is not explained in this infinite interpretation of interpretation but just postponed further and further. The problem of meaning brought Peirce to reconsider his notion of infinity not only in semeiotics but also, as we will see in the following section, in his mathematics.

This fullest development of Peirce's theory of signs by 1903 (*The Harvard Lectures*[34] and *Lectures at the Lowell Institute*[35]) can be read as further elaboration of his earlier three–partite distinction between the signifying element of sign, its object and its interpretant. These developments also brought Peirce to increase the constitutive number of signs from three to ten.

1.3.1 *The Signifying Element*

The signifying element stands to its object only in some capacity representing or signifying its general features. In this regard Peirce indicated that the signifying element would signify the object either in terms of its qualities existential or physical facts or laws about the object. Thus, in the first instance, we have a signifying element as 'qualisign', in the second as 'sinsign' and in the third as 'legisign'.

Colors in traffic lights can serve as the example of the first signifying element. Their role is to indicate, to signify only the quality of some object; i.e., 'red' for 'stop', etc. For this reason the usual candidates for qualisigns are colors. José F. Vericat speaks of the "iconical force" of colors in Peirce and Wittgenstein by representing "the abstractiveness of abstraction".[36]

Sinsigns as signifying elements are casually related facts, such as cry is the (sin)sign of pain or smoke is of fire. Legisigns depend upon laws and commonly agreed conventions, such as the elementary dictionary meaning of words, or, by Peirce's own example, a type of a diagram.

1.3.2 *Objects*

In terms of the objects, Peirce retained the same classification as before. An icon represents the qualitative feature of its object, an index signifies physical features and a symbol reflects the conventional or nomological element of its object.

In terms of objects, Peirce made a further comment on icons to include also geometrical diagrams; indices to include proper names and indicative gestures; symbols to include judgments. However, unlike symbols, icons and indices are conventional and, mostly symbolic themselves. This meant that in terms of objects, his trichotomy of icon–index–symbol were abstract by nature. A sign would display in some way the combination of all three features.

---

[34] From 1903 in Peirce (1998: 133–241).

[35] From 1903 in Peirce (1998: 258–330).

[36] Vericat (1994: 299).

1.3.3 *The Interpretant*

Whenever sign is interpreted by its qualitative features, Peirce called it rheme. Atkin speaks of "unsaturated predicates". a term borrowed from chemistry often used by Peirce in his logic. These unsaturated predicates are verbs standing between the subject and object in a proposition, i.e., '___ loves ___', '___ gives ___', etc.[37]

A rheme is an interpretation of an object by its qualities. Whenever we interpret an object by its physical feature, it is a dicent. These are 'saturated predicates', complete phrases filling rheme: 'Adam loves Eve', 'Eve gives fruit to Adam', etc. Whenever a sign determines an interpretant by the nomological or conventional feature of an object, it is called delome, or argument.

Rheme, dicent and delome/argument have their most clear use in Peirce's new predicate logic, indeed, as unsaturated and saturated propositions and as a rule of inference. In this later period of further elaborated semeiotic Peirce included in his discourse on the signs the notion of reference. His mature theory of meaning relies on that important later development.

Meaning is explained by the reference of the object to an idea through either shared understanding between the minds of the interpretants of signs, or through private understanding by a single mind. By 'private', however, one must not mean incommunicable. Just as Wittgenstein, Peirce did not consider any possibility of private meanings nor private languages about them.

In this third account of semeiotics, Peirce elaborates significantly on his theory of meaning in respect to his earlier account from 1868. In that first account the relation between word (or symbol) to its object was direct, just as it will be in Frege's 1892 paper "Über Sinn und Bedeutung". However, from this later period Peirce's definition of a sign clearly expresses also his overall notion of meaning from 1908:

> I define a *Sign* as anything which on the one hand is so determined by an Object and on the other hand so determines an idea in a person's mind, that this latter determination, which I term the *Interpretant* of the sign, is thereby mediately determined by that Object. A sign, therefore, has a triadic relation to its Object and to its Interpretant. But it is necessary to distinguish the *Immediate Object,* or the Object as the Sign represents it, from the *Dynamical Object,* or really efficient but not immediately present Object. It is likewise requisite to distinguish the *Immediate Interpretant,* i.e. the Interpretant represented or signified in the Sign, from the *Dynamic Interpretant,* or effect actually produced on the mind by the Sign; and both of these from the *Normal Interpretant,* or effect that would be produced on the mind by the Sign after sufficient development of thought.[38]

The first two sentences in the above quotation are reaffirming Peirce's earlier definitions of sign and the triadic relation was already explained. However, his distinction between the immediate and the dynamical objects are novel, and so is the new triadic relation between the immediate, dynamic and normal interpretants.

---

[37] Atkin (2008: 63–85).

[38] Peirce (1958b: 343).

According to T. L. Short, these modifications came as a result of Peirce' corrections to the flaws of this earlier semeiotics from 1868.[39] The major difficulty for Peirce was to explain how thoughts are related to signs and to objects. If thoughts being signs themselves rely on interpretation and interpretation is consisted of signs one easily can see the regress of thought–signs.

Peirce's corrections came from 1883 to 1906 in a series of further divisions among signs. The first division was that of icon, index and symbol:

> There should be three classes of *signs;* for there is a triple connection of *sign, thing signified, cognition produced in the mind.* There may be a mere relation of reason between the sign and the thing signified; in that case the sign is an *icon.* Or there may be a direct physical connection; in that case, the sign is an *index.* Or there may be a relation which consists in the fact that the mind associates the sign with its object; in that case the sign is a *name* [or *symbol*]. Now consider the difference between a logical *term,* a *proposition,* and an *inference.* A term is a mere general description, and as neither *icon* nor *index* possesses generality, it must be a name; and it is nothing more. A proposition is also a general description, but it differs from a term in that it purports to be in a real relation to the fact, to be really determined by it; thus, a proposition can only be formed of the conjunction of a name and an index. An inference, too, contains a general description.[40]

Thus, we understand sign the way it plays a specific role in reasoning, by the features of the world it signifies. We deduce information, make inductive and abductive inferences based on the way each sign stands in its particular capacity in the world just as in our inferential thinking.

1.3.4 *Semeiotic Ontology and Metaphysics*

Peirce places the sign between experience (firstness and secondness) and habit as a general rule of cognition. Thus, signs are everything but experience and habit; their role is to be that connecting force between the two: "A sign therefore is an object which is in relation to its object on the one hand and to an interpretant on the other, in such a way as to bring the interpretant into a relation to the object, corresponding to its own relation to the object. I might say 'similar to its own' for a correspondence consists in a similarity; but perhaps correspondence is narrower".[41]

The sign is also representamen, representing or standing for something in some capacity. Its nature is relation between the object and interpretant. The object is that which is represented and, just as in the early Wittgenstein, is of phenomenological nature. The interpretant of a sign usually is the mind or certain state of mind. It is something dynamic and by Peirce' own definition is itself a sign. In Peirce's best known definition of a sign, signs are considered to be the same as representaments:

---

[39] Short (2006: 214–240).

[40] From ca. 1885 in Peirce (1933: 372).

[41] Peirce (1931: 372).

> A sign, or *representamen,* is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object.* It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the *ground* of the representamen. "Idea" is here to be understood in a sort of Platonic sense, very familiar in everyday talk.[42]

Here Peirce makes a clear indication as to the phenomenological nature of the 'objects' ("in reference to a sort of [Platonic] idea") and their complete dependence on the mental contents of the interpretants. But the interpretant must also be a sign, the same kind of a sign as the object that it represents. Furthermore, the interpretatant being the (second) kind of sign needs to be interpreted and becomes another (third) kind of sign of an object. This third sign has also mental features, since it is hermeneutical and phenomenological by nature: "Anything which determines something else (its *interpretant*) to refer to an object to which itself refers (its *object*) in the same way, the interpretant becoming in turn a sign, and so on *ad infinitum*".[43]

This implies that if there is any object then there is a sign that stands by that object, and if there is a sign then there is an infinite sequence of signs referred to that one object. Therefore, phaneron is a sign itself and is an infinite sequence of signs. Finally, phaneron is by nature evolutionary. Signs stand to each other in the same 'evolutionary hierarchy' (concept since 1893[44]) as anything mental or physical in the world. Not only species but also objects and ideas evolve with time. Although of a different kind, interpretations of signs presupposes an evolution process as well.

Peirce's logic of scientific discovery (abduction, deduction and induction) must be regarded as an essentially evolutionary process that has no end and cannot put an end to any inquiry. This firm belief shaped not only Peirce' pragmaticism but also his anti–foundationalist and intrinsically fallibilist epistemology but also philosophy in general. Our knowledge does not begin from some private *ego* but from the evolving and constantly changing world. Then, the *ego* is as much part of that world than anything else, and if it is part of the world, it too obeys its laws, subjected to its regularities and habits.

An essential characteristic of Peirce's anti–foundationalist ontology and metaphysics is, indeed, his semeiotic correspondence between the world and its description. The relationship between sign, object and interpretant is never of an exact correspondence. In order to demonstrate this last point I shall briefly present the three elements of signs.

Peirce gave many names to the first element of the sign: 'sign', 'representamen', 'representation', 'ground', these are some among many. In the secondary literature it is often referred to as 'sign–vehicle', although Peirce himself never used that term.[45]

---

[42] From 1903 in Peirce (1932: 228).

[43] From 1903 in Peirce (1932: 303).

[44] From 1893, "Evolutionary Love" in Peirce (1992a: 352–371).

[45] The closest he came to call sign (the signifying element of a sign) as a 'sign–vehicle' is in Peirce (1931: 339), an unidentified fragment: "A sign stands *for* something *to* the idea which it produces, or modifies. Or, it is a vehicle conveying into the mind something from without. That

The reason why sign in general and its signifying element are not exactly treated the same in Peirce is due to the fact that sign never signifies *every* feature of its object. Take, for instance, the road signs signifying objects, road conditions, situations, and so on. While objects, conditions and situations will vary from place to place, road signs are the same in their generalization of what they meant to signify. The signifying element of the sign generalizes to some degree its presentation of its object but does not carry in full all of its features, elements or parts.

The object of a sign *determines* the sign by placing conditions on the sign. Take, i.e. a sign of the right turn on the road. One can easily point that the right turn of the road *determines* directionality of an arrow on the sign. Just as with the signifying element, the object of the sign only in some features determines the sign.

The role of an interpretant is to give a translation between the signifying element and the object of a sign. Since the interpretant, as was mentioned, can be equated with the cognitive process of bringing the signifying element with its object, the interpretant will always depend on our understanding of the object to be signified.

1.4. The Fourth Account (1906 – 1910)

After his *Lectures* in 1903 Peirce began to notice interesting connections between his semeiotics and his logic. Indeed, in the last quoted passage from 1903 this connection already begins to emerge. If logic is that of scientific discovery, the discovery of new information about the world instead of the old epistemological quest to justify previously acquired "true beliefs". and semeiosis as a systematic description of signs in their formation, then use and interpretation is central to that kind of logic and epistemology.

Unlike in his earlier accounts where the chain of signs would proceed infinitely, in this last period Peirce argued for a goal–oriented and end–directed semeiotic process of inquiry. The end of the inquiry is knowledge about the object. A kind of knowledge that completes the inquiry's process, tells us all there is to know about the object. For this reason there would be no further interpretant of the known object, thus, the trichotomy is neither as essential nor constitutive to semeiotic as it was thought before.

Instead of an infinite chain of signs Peirce speaks of a telic *continuum* of signs, terminating in a complete knowledge of its object. But what sort of object is that?

1.4.1 *Dynamic and Immediate Objects*

Peirce distinguished between two kinds of objects: dynamic and immediate. The object of the sign understood at some point of an inquiry is the immediate object. It is changeable and fallible according to our understanding at some particular time of the inquiry. The dynamic object of the sign, on the other hand, is that of a complete understanding, viz. at the very end of the inquiry.

---

for which it stands is called its *object;* that which it conveys, its *meaning;* and the idea to which it gives rise, its *interpretant*".

Christopher Hookway considers the first as the sign of the object used in the inquiry and the second is when we have a scientific knowledge. Thus, the dynamic object as the object we come to know it to be at the end of the inquiry or, as the object really is.[46]

The immediate object, on the contrary, is what we *presume* the object to be at the intermediate points of inquiry, as the inquiry proceeds toward full knowledge. However, we must never think of two separate objects here: the object is the same while our knowledge is changing as the inquiry progresses. Our knowledge of the same object is different and we understand the same object as if objects were indeed different at diverse periods of the inquiry.

Peirce speaks of the chain of signs in connection between the immediate and dynamic objects.[47] The inquiry is the semeiotic process in which the immediate object is our progressive understanding of the object of the sign and the dynamical object is the goal of that process.

1.4.2 *Immediate, Dynamic and Final Interpretants*

Interpretant is the relationship between the object and its sign. In the later period Peirce distinguished three kinds of interpretants: immediate, dynamic and final.

The immediate interpretant is the syntactical recognition of the most basic features of the object. David Savan speaks of the immediate interpretant in terms of "unanalyzed impression" produced by the sign before we acquire critical understanding.[48]

Peirce's earlier notion of the first grade of clarity, also called 'familiarity', is that immediate interpretant. It is our initial familiarization with the object in semeiotic inquiry.[49] The dynamic interpretant is our understanding of the sign and dynamic object relationship within the chain of signs in the semeiotic process. Peirce defines it as the "actual effect which the sign, as a sign, really determines".[50] In relation to the immediate object, the dynamic interpretant is the understanding we have of the immediate object within the inquiry, or at any specific time in the semeiotic process.

The final (or 'normal' by Peirce's own nomenclature of that time) interpretant is the "effect that would be produced on the mind by the sign after sufficient development of thought".[51] The final interpretant is our understanding of the dynamic object at the end of the semeiotic process, which might also be the end of a scientific inquiry. It is the genuine understanding of the dynamic object.[52]

The final interpretant is the stage of the inquiry at which our understanding of the dynamic object is complete. Ransdell speaks of a "coincidence" between the immediate

---

[46] Hookway (1985: 139).

[47] From 1 April 1909 in Peirce (1958b: 315); From 1909 in Peirce (1958b: 185).

[48] Savan (1988: 53).

[49] From 1878 "How To Make Our Ideas Clear" in Peirce (1986: 257–275); From 1909 in Peirce (1958b: 185).

[50] From 1906 in Peirce (1933: 536).

[51] From 1908 in Peirce (1958b: 343).

[52] From 1909 in Peirce (1958b: 184).

and the dynamic object(s) as the final interpretant is reached.[53] It is also the point that sets the norms of understanding the dynamic object, the point toward which all inquiry must strive.

There were further accounts of Perice's semeiotic developments with many alternative nomenclature versions for signs, objects and interpretants. For instance, the six elements of the sign were elaborated into a final classification of sixty–six signs, in addition to further distinctions of interpretants. These accounts are found in his letters and manuscripts in the forms of drafts of a clearly unfinished work. Considerations of space and purpose do not permit here examinations of these materials.

2. Cognition as a Semeiotic Continuum

It has been argued that Peirce had several periods of development of his semeiotics. The same claim was made in relation to his theory of continuity, continuum[54], or synechim, defined as "that tendency of philosophical thought which insists upon the idea of continuity".[55] However, in the secondary literature these two themes are rarely compared together. Likewise, the suggestion that changes in one implied changes in the other are done randomly if at all.

Which theory influenced which in Peirce? If we compare the two in chronological order we will see that changes in semeiotics preceded changes in the theory of continuum, just as his revisions in logic impelled him to revise his semeiotics. Potter and Shields[56] have argued for four periods and Havenel[57] speaks of five periods of development of the mathematico–philosophical theory of continuum. Instead of speaking of distinct periods (some of which cannot be easily determined even to the year), I shall speak of distinct notions of continuum:

1. Infinite Continuum (1868 – 1884)
2. Mathematical Continuum (1884 – 1892)
3. Categorical Continuum (1892 – 1897)
4. Logical Continuum (1897 – 1907)
5. Semeiotic Continuum (1908 – 1913)

In relation to the presented semeiotics, Peirce's Infinite Continuum (1868 – 1884) includes the First Account of Semeiotics (1865 – 1877) and extends to the Second Account (1877 – 1885). The Mathematical (1884 – 1892) and the Categorical (1892 – 1897) Continuum were elaborated at the time when Peirce did not write on semeiotics. The Logical Continuum (1897 – 1907) was elaborated while the Third Account of Semeiotics was given (1901 – 1903). The Fourth Account of Semeiotics (1906 – 1910) influenced his last Semeiotic Continuum (1908 – 1913).

---

[53] Ransdell (1977: 170).

[54] Peirce uses interchangeably terms of 'continuum' and 'continuity'. See Putnam (1955).

[55] From 1902 in Peirce (1935: 169).

[56] Potter and Shields (1977).

[57] Havenel (2008).

Strictly speaking there is no correspondence between Peirce's semeiotics and continuity periodization, especially due to the gap in his writings on semeiotics between 1885 and 1902. While he did not write much in semeiotics at that time, his writings on continuity from the same time are very important. Still, I think that changes in his theory of continuity were effected by the changes in his semeiotics. The process can be seen circular, since changes in semeiotics were influenced by developments in logic and his logic was built on mathematics.

I believe this was not a mere coincidence in time or that the two subjects were related. Peirce's first interest in continuum came from his early theory of cognition as a direct consequence of his idea of the infinite interpretant. Peirce's interest in the theory of continuum from the 1880s and 1890s was mostly of a mathematical character. He was actively participating at that time in discussions on the topic. But by the end of the century and at the beginning of the 1900s his interest gradually switched from mathematics to semeiotics again. In fact, the change from infinity to continuum in the later period of his writings on semeiotics comes just two years before he develops his philosophical account of continuum and abandons most of his previous views on it. This section will explain how this process occurred.

2.1 Infinite Continuum (1868 – 1884)

In Peirce 1868 he gives his first definition of continuum as something that "every part of which has parts, in the same sense".[58] It means that continuum is not a composition of its parts but rather that its parts are of the continuum. As Havenel indicates,[59] in 1868 his interest in continuum developed out of his theory of cognition, when at the same time he rejected nominalism.

A very similar definition of cognition can be read saying that "At no one instant in my state of mind is there cognition or representation, but in the relation of my states of mind at different instants there is".[60] Cognition cannot be localized the way physical objects are. We cannot say that our thoughts are time–dependent, just as we cannot say that our feelings are measured by space and time. Thus, they are not part of a continuum but rather our thoughts and feelings are a continuum themselves: "The mind is a sign developing according to the laws of inference".[61]

The fundamental consequence of this theory of continuum is that there are, strictly speaking, no individuals in a continuum. One can speak of the points on the line (as Cantor did, in fact), but these points lack any individuality on their own.

In 1878 he further defined continuity as "the passage from one form to another by insensible degrees".[62] A continuum is, in fact, the form of forms, and forms within

[58] From 1868 in Peirce (1934: 335).
[59] Havenel (2008).
[60] From 1868 in Peirce (1934: 289).
[61] From 1868 in Peirce (1984: 240–241).
[62] Peirce (1932: 646).

continuity are arranged ad infinitum.[63] In this period, however, Peirce makes no distinction between continuity and infinite divisibility, but by 1881 he began to realize his mistake. However, as Potter and Schields indicated,[64] this mistake will be corrected only after his acquaintance with Georg Cantor's writings.

2.2 Mathematical Continuum (1884 – 1892)

When Peirce revised his previous treatment of the categories in 1884, thirdness then corresponded to continuum and was associated with time. In the same year he read for the first time Cantor's articles in *Acta Mathematica*.[65] These two events were the reasons for Peirce to make the necessary distinction between continuity and infinite divisibility.

In 1889 Peirce thus considers 'continuity' as:

> Uninterrupted connection of parts in space or time; uninterruptedness.
>
> In *mathematics* and *philosophy*, a connection of points (or other elements) as intimate as that of the instants or points of an interval of time: thus, the *continuity* of space consists in this, that a point can move from any one position to any other so that at each instant it shall have a definite and distinct position in space. This statement is not, however, a proper definition of *continuity*, but only an exemplification drawn from time. The old definitions – the fact that adjacent parts have their limits in common (Aristotle), infinite divisibility (Kant), the fact that between any two points there is a third (which is true of the system of rational numbers) – are inadequate. The less unsatisfactory definition is that of G. Cantor, that continuity is the *perfect concatenation* of a system of points – words which must be understood in special senses. Cantor calls a system of points *concatenated* when any two of them being given, and also any finite distance, however small, it is always possible to find a finite number of other points of the system through which by successive steps, each less than the given distance, it would be possible to proceed from one of the given points to the other. He terms a system of points *perfect* when, whatever point not belonging to the system be given, it is possible to find a finite distance so small that there are not an infinite number of points of the system within that distance of the given point. As examples of a concatenated system not perfect, Cantor gives the rational and also the irrational numbers in any interval. As an example of a perfect system not concatenated, he gives all the numbers whose expression in decimals, however far carried out, would contain no figures except 0 and 9.[66]

Among the three definitions of continuity (Aristotle's, Kant's and Cantor's), Peirce considers that of Cantor to be "the less unsatisfactory". The major disagreement between Peirce and Cantor in terms of continuity consists in Cantor's denial that continuity is somehow related to time and space. Compare the above definition with that of Cantor

[63] Hartshorne (1929: 521–534).

[64] Potter and Shields (1977: 21).

[65] Peirce (1976b: 883).

[66] Peirce (2010: 136–139).

from *Grundlagen*: "Under a continuum in the true sense, I understand every perfect set of points, which is continuous in itself".[67]

What Peirce calls "concatenated" is what Cantor uses as "*zusammenhängend*". Havenel translates it as 'connected', and Potter and Shields as 'connectedness' or 'cohesiveness', but these translations are somewhat ambiguous. Ambiguity is eliminated when we consider the verb to be reflexive in view of the reflexive pronoun '*sich*' referred to '*Punktmenge*'. This reflexivity is the essential feature of the set of points being perfect. It is a set whose value coincides with the value of its subset, otherwise it is *reducible.*

If "*hängend*" literally means "hanging". 'concatenated' would be the closest to the German, since it means linked together like in a chain (from Latin *catena*, chain). Furthermore, since it has a direct object it is a transitive (to hang up together) rather than intransitive verb (to be connected). Here, concatenated or "continuous in itself". implies that continuity is distinct and does not imply divisibility. That is why Peirce in 1889 uses Cantor's definition to solve his previous problem of providing such a distinction.

But Cantor's notion of "*zusammenhängend*" does not fit together with Peirce's notion of thirdness as time. Indeed, in 1879 Cantor,[68] defining the notion of continuum as a perfect set of points, criticized his predecessors from Leucippus to Kant for defining continuum in terms of space and time. Continuity is that of points on the line and it must be independent of time and space in order to avoid circularity.

2.3. Categorical Continuum (1892 – 1897)

Between 1892 and 1903 Peirce suggested considering continuity in terms of 'Aristotelicity' and 'Kanticity'. 'Aristotelicity' meant that continuum is a perfect set in which the power set would be equal to the sum of its all subsets. 'Kanticity' implied the infinite divisibility of a continuum. Around 1895, however, he suggested that 'Kanticity' should really mean indivisibility of a continuum into individual points.[69] Starting from the 1890s Peirce begins to think of continuum in less mathematical and more logical and semeiotic terms:

> The reality of continuity appears most clearly in reference to mental phenomena; and it is shown that every general concept is, in reference to its individuals, strictly a continuum. This (though asserted by Kant and others) did not appear quite evident as long as the doctrine of generals was restricted to non–relative terms. But in the light of the logic of relatives, the general is seen to be precisely the continuous. Therefore, the doctrine of the reality of continuity is simply that doctrine the scholastics called realism.[70]

---

[67] „Unter einem Kontinuum im eigentlichen Sinne verstehe ich jede perfekte Punktmenge, die in sich zusammenhängend ist". G. Cantor, „Grundlagen einer allgemeinen Mannigfaltigkeitslehre", §19, from 1883 in Cantor (1966: 236).

[68] „Über einen Satz aus der Theorie der stetigen Mannigfaltigkeiten". §10, from 1879 in Cantor (1966: 190–194).

[69] Peirce (1935: 122–123).

[70] Peirce (1958b: Bibliography General 1893 [G–1893–5]).

The reality of continuity is that of a general, general here being universal of being. Continuity in this period does not only come as thirdness but as Peirce's notion of the category of being, thus his reference to scholastic realism.

In "On Quantity, With Special Reference to Collectional and Mathematical Infinity" (1895), Peirce asked five questions about the nature of mathematics, quantity, continuity and infinity. The fourth question is phrased thus: "Is it possible that there should be two collections not equal in multitude yet of which neither is greater than the other? That is to say, can there be two collections neither of which could in any way be put into a one–to–one correspondence with a part or the whole of the other?"[71]

What Peirce referred here to is, of course, Cantor's notion of continuum according to which for every set there cannot be a one–to–one correspondence between the set's integers and the real numbers. That, according to Cantor, proves the existence of the infinite sets.[72]

In the following pages of the paper Peirce goes on to demonstrate that such proof is not only impossible (indeed this will be shown by Kurt Gödel in 1940[73] and by Paul J. Cohen in 1963[74] – 1964[75]), but that this mathematical problem cannot be even stated in terms of 'quantity' or 'multitude' since these are the notions which the hypothesis is trying to establish. It was the very same fallacy that Cantor was trying to avoid by rejecting the idea of continuum being related to space and time:

> Thus, *quantitative* reasoning, in the narrow sense here given to "quantity". is unable to cope with questions relating to grades of inenumerable multitude, although those questions can perfectly well be mathematically discussed. (…)
>
> Cantor, in effect, defines the continuity of a line as consisting in that line's containing *all* its points. This is a singular *circulus in definiendo*, since the very problem was to state how those points were related. But I should not have noticed it, were it not that the phrase seems to imply that the line contains as many points as it could contain. Now we have seen in the last section that there is no maximum grade of multitude. If, therefore, a line contains all the points there could be, these points must cease to form a multitude.[76]

Peirce grants to Cantor that he had came up with a genuine mathematical problem but that even its mere formulation implying notions of 'quantity' and 'multitude' is circular.[77] Peirce concludes: "We must either hold that there are not as many points upon a

---

[71] Peirce (1976b: 39).

[72] „Ein Beitrag zur Mannigfaltigkeitslehre". from 1878 in Cantor (1878: 242–258).

[73] "The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory". in Gödel (1990: 33–101); see also his "What is Cantor's Continuum Problem?" in Gödel (1990: 176–187).

[74] Cohen (1963).

[75] Cohen (1964: 51).

[76] From 1895 in Peirce (1976b: 58).

[77] In a recent lecture at Harvard, Solomon Feferman goes further and denies that continuum hypothesis *is* a definite mathematical problem. By 'definite' Feferman considers problem that can prove truthfulness or falsity of a proposition. Since continuum hypothesis does not employ truth–

line as there might be, or else we must say that points are in some sense fictions which are freely made up when and where they are wanted".[78]

This critique of Cantor's hypothesis anticipates Hilbert's famous assessment: "From the paradise, created for us by Cantor, no one should be able to expel us".[79] But unlike Hilbert, Peirce was not part of the tradition of mathematical finitism, nor like Brouwer part of intuitionism (despite Ketner's attempts to see Peirce as a forerunner of the mathematical intuitionism of Brouwer[80]). What was, then Peirce's take on continuum and infinitesimals?

Peirce favored referring to Cantor's sets (*Menge*) as *collections*. Collections consisted of individuals. One can think of individuals as numbers or points on the line, just as Cantor did.

Collections are different from generals. For the later, individuals determine generals and generals cannot exist separately from their individuals. Unlike generals, collections form one *continuum*, in which individuals form singularity. It is this singularity of continuum (or singularity of individuals in a collection) that allowed Peirce to have the notion of abstraction.

However, as Pietarinen indicates, abstraction was defined by Peirce not in terms of sets (or collections) but in terms of *signs*.[81] Signs stand to singularity in continuum of future interpretations of the object. Although collection just as general has its identity from its individuals, unlike general, the identities of these individuals within collection are constantly open to change and with that also the identity of collection itself changes.

This way, the mind makes connections between different collections by connecting one collection of individuals with another. In *Elements of Mathematics* Peirce writes: "A *correspondence* is a connection established in the mind between two collections such that, considering either, every object of it is connected with the same number of objects of the other".[82]

Thus, one must understand correspondence as a connection of the signs, not sets. In fact, in the following paragraph Peirce explains further: "It is *not* said that any object of either is connected with the same number of the other collection, since the two numbers

---

value, it is not a definite mathematical problem. Perhaps a somewhat radicalized view of Peirce. Feferman (2011).

[78] Peirce (1976b: 59).

[79] „Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können". Wittgenstein remarks on this quote from Hilbert in Wittgenstein (1983: V, §7).

[80] It is very difficult to properly place Peirce within any mathematical tradition. Constructivism and intuitionism would come to mind, but intuitionism without finitism, thus he was hardly a forerunner of Brouwerian intuitionism. Murphy suggested viewing Peirce's mathematics as that between intuitionism and logicism. See Murphey (1961: 288). However, for Peirce mathematics is independent and could not have been found on logic. Pietarinen sees Peirce more within con-structivist tradition. See Pietarinen (2006: 186–191). As it will be explained in the next chapter, Peirce's own game theory is closer to that of formalists than Wittgenstein and logicism, since games for him were treated analogically in description of social behavior.

[81] Pietarinen (2006: 37).

[82] From 1895 in Peirce (1976a: 34).

may be different".[83] Considering Cantor's diagonal, the fact that numbers are different constitutes the paradox but implies no contradiction to the theory. In fact, mathematics is not preoccupied with the notion of identity between the individuals (points, integers, etc.) of two or more sets. Such problems arise, indeed, for the logician, not for the mathematician.

2.4 Logical Continuum (1897 – 1907)

In the decade between 1897 and 1907, the period which Havenel calls the "Supermultitudinous Period". Peirce becomes more influenced by the Aristotelian notion of potentiality in his treatment of a continua. Such influence clearly shows in his understanding of multitude as potentially greater than any set in Cantor's definition. As a consequence he insisted that continuum cannot be considered in terms of actual but rather potential points.

This last argument is crucial in understanding the difference between sets and collections:

> That a *collection* is a species of abstraction becomes evident as soon as one defines the term *collection*. A *collection* is a substance whose existence consists in the existence of certain other things called its *members.*
>
> An abstraction being a substance whose existence consists in something being true of something else, when this truth is a mere truth of existence the abstraction becomes a collection.[84]

"A mere truth of existence" here must be read not only in relation to Peirce's semeiotics, but also in connection with his notion of modality. Truth is world–dependent because it refers to the collection at its present state of aggregates of its members in a single world. Here the 'world' is meant in a sense of a possible 'world of discourse' and not, like in i.e. David Lewis physical possible world. When the identity of its members changes (due to i.e. different interpretants), predication of that collection and the collection itself changes as well. Notice that none of such changes within and of collections are possible in sets. Once the set is defined it cannot be subjected to any changes.

Elsewhere Peirce speaks of hypostatic abstraction in terms of predication along the same lines as predication of the collections. First Peirce reviews his notion of a collection and multitude:

> A *collection* is an hypostatic abstraction, or *ens rationis,* that *multitude* is the hypostatic abstraction derived from a predicate of a collection, and that a *cardinal number* is an abstraction attached to a multitude. So an *ordinal number* is an abstraction attached to a *place,* which in its turn is a hypostatic abstraction from a relative character of a unit of a *series,* itself an abstraction again.

---

[83] Peirce (1976a: 34); see also Peirce (1976a: 46).

[84] From 1903 in Peirce (1976c: 164).

But then Peirce implies that thinking is nothing more than the same process of a hypostatic abstraction from thinking into thinking about a thought:

> What you mean by a *concept* is a predicate considered by itself, except for its connection with the word or other symbol expressing it, and now regarded as denotative of the concept. Such a concept is not merely prescissively abstracted, but, as being made a subject of thought, is hypostatically abstract. So understood, it is true that it is more removed from the perceptual objects than is the *Vorstellung,* or composite of images. But for all that, its intellectual purport is just the same. It is only the grammatico–logical form that is transmuted.[85]

Thinking about thinking or about a thought can be reiterated as forming a sign of a sign. This ability is what Peirce considered to be cognition. Once again, his semeiotics and metaphysics merge into one idea that essentially cognition is a capacity (potential or actual) to produce signs, but such capacity itself is made possible by what the world actually is. In other words, the world is present to us as continuity of signs open to interpretations.

In this categorical continuum Peirce follows neither Aristotelian nor Kantian understanding of continuum. The Kantian notion was dismissed earlier on the grounds that continuum cannot be infinite divisibility. The Aristotelian continuum does not imply semeiotic nor fallibility. While the Kantian continuum is wrong for Peirce, the Aristotelian is simply not enough to serve as an explanation. Peirce's final rejection of Cantor's notion of continuum as one–to–one correspondence between continuity of space and continuity of numbers on the line (viz. coordinates x, y, z as numbers x, y, z),[86] is rejection of the idea that continuum can be limited. Continuum cannot have numerical or geometrical points.

Rather than providing a purely mathematical account of continuum Peirce's goal was to give a logical account of continuity.[87] However, the secondary literature predominantly treats his writings on continuum as if they were an "intermediate" view within the rich history of set theories and continuum theories of the 19th century.[88] Such a 'mathematical' treatment of Peirce's views on continuum inevitably ends up in confusions over Peirce's use of identity between points on the line and his famous example of

---

[85] Peirce (1934: 534).

[86] „Über unendliche lineare Punktmannigfaltigkeiten. Nr. 3". from 1882 in Cantor (1966: 156). J. Dauben quotes and comments on that article by Cantor: "In general there was no inner constraint for thinking of every point given by coordinates of real numbers x, y, z as actually belonging to three–dimensional space. The assumption, he insisted, "must be regarded as a free act of our mental constructive activity. The hypothesis of the continuity of space is therefore nothing but the assumption, arbitrary in itself, of the complete, one–to–one correspondence between the 3–dimensional purely arithmetic continuum (x, y, z) and the space underlying the world of phenomena [*Erscheinungswelt*]". Dauben (1979: 86).

[87] From 1897 in Peirce (1933: 526ff).

[88] Bell (2006: 208–211); Iordache (2011). Perhaps this is also the reason why some scholars speak of Peirce's misunderstanding of Cantor's ideas on continuum. See Murphey (1961: 241). Was it not difference in approach and method to continuity rather than misunderstanding?

breaking the line in two while the point at which the line was broken becomes two, not one.[89] Likewise confusing is Peirce's seemingly paradoxical view on identity of the line based on the identity of its points, while at the same time the identities of the points are lost in the identity of their line.

The usual solutions are not Peirce's. The identity of points that can be transformed upon division of the line from one to two is said to be monadic, from Leibniz's metaphysics[90] and from Peirce's own logic ('monadic predicates').[91] The line composed of the infinitely many points is said to be potential but determined by actuality to be realized in time, from Aristotelian metaphysics. Similarly, his notion of collections and their multitude is compared to Cantor's sets and their cardinality. The last interpretation is easier to suggest given the fact that both Peirce and Cantor recurred to geometry in their treatments of continuum.[92] One must keep in mind, however, that at least in RTL (1898) Peirce's notion of the 'line' was non–Euclidean.

Furthermore, Cantor (and after him Frege, Russell and Whitehead) belonged to a different mathematical tradition than Peirce. This tradition, often named 'logicism', had as its primary goal establishing a formal system that would be equally compatible with mathematics as well as with quantificational logic. As Burton Dreben and Jean van Heijenoort explain in their *Introductory Note* to Gödel's Completeness Theorem, Peirce was part of the tradition that derived from Boole and had its developments in the works of Schröder and Löwenheim. This approach is "purely model–theoretic, that is, semantic".[93] In the case of Peirce one should also add, 'semeiotic'. Notions of universal truth or validity (so permanent in Frege, Russell, and in TLP) are absent in this approach. Instead, truth and validity are relevant not universally but from one model to another, or from one possible world to another.

Perhaps the major difference between these two traditions consists in the fact that if logicism implies that mathematics must be founded on logic, or that mathematical propositions essentially are logical ones,[94] for Peirce it is the other way about: logic is based

---

[89] Peirce (1992b: 159–160); Peirce (2010: 172–174).

[90] In the sense that hylomorphically composite monad, being the only true substance, can be divided in two. Peirce considered Leibniz's monadology as an extreme nominalism to which he was clearly opposed. See from 1903 in Peirce (1931: 19). Instead, his definition of a monad runs as follows: "a monad will mean an element which, except that it is thought as applying to some subject, has no other characters than those which are complete in it without any reference to anything else". From 1908 Peirce (1931: 292). This is what also is meant by 'monadic predication' in 'monadic calculus.' See „Über Möglichkeiten im Relativkalkül" in Löwenheim (1915). „Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem" in Behmann (1922: 163–229).

[91] See Boler (1963: 74. 76–8. 109).

[92] The first two treatments of Peirce's notion of continuum come from Kenneth Laine Ketner's Introduction to Peirce (1992b: 37–54). The third one comes from Pietarinen (2006: 36–39). On geometry of continuum in both and in Hilary Putnam's Comments on RLT in Peirce (1992a: 94–102). On the other extreme is Matthew E. Moore's assessment that Peirce's "mathematical conception of continuum" is contrasted with the "pseudo–continuum" of Cantor and Dedekind. However, it is not clear of which period Moore is speaking. In Peirce (2010: xviii).

[93] Dreben and van Heijenoort in Gödel (1986: 44–59).

[94] Carnap (1931b).

on mathematics. As Carolyn Eisele writes in her introduction to Peirce's *The New Elements of Mathematics*: "For Peirce, multitudes corresponded to a linear series of objects and one does not use the forms of pure mathematics in investigating them but uses a branch of logic which is directly dependent on mathematics".[95]

The result is that in order to make sense of Peirce's seemingly inconsistencies with mathematical theories one has to suggest readings–into Peirce which, not surprisingly, ends up in creating something that Peirce never wrote himself. The 'secret' of interpreting Peirce right is treating his views on continuum not as a separate topic on its own but as supplementary to his metaphysics and, especially, to his semeiotics, but the method of inquiry is always that of his logic of inquiry and discovery. If his semeiotics itself is supplementary to his metaphysics then the role of continuum cannot be measured or compared to that same role it played in Cantor, Dedekind or Brouwer. But in that case, his views on continuum should not be seen as mathematical or metaphysical but as logical.

The clearest statements of that come from his 1903 *Lowell Lectures*. In the following passage Peirce gives a brief but concise history of the development of his notion of continuum in relation to that of Cantor:

> The doctrine of multitude is not a theory of pure mathematics. It is, rather, an application of the general theory of all numbers to the logical subjects of maniness. Nevertheless, I shall begin by this. The recognized leader of this branch is Dr. Georg Cantor, but I began my studies of it and pushed them to considerable results before Cantor took up the subject, and I had made out the main outlines of the doctrine before I knew anything about his work; and have developed it in my own way quite independently, reaching some results not made out clearly by him, and not at all without the aid of the Pure Mathematical theory, which I, as a logician treating a subject properly logical, do not do.[96]

Peirce proceeds in giving the reason why Cantor could never come up with a definitive treatment of continuum. Instead of treating it as a subject of logic he offers a mathematical study. He then considers his study as that of metaphysics. Indeed, in the words of Hilbert and Wittgenstein, a created paradise:

> This question has not received from Cantor any answer generally deemed conclusive; and it is impossible that it should, since Cantor looks at the subject from a purely mathematical point of view, while this question is a logical one. Perhaps I ought not to omit to say that Cantor has also multitude and number from a metaphysical point of view. But that is, if possible, still less to meet this question properly.[97]

Five years earlier, his 1898 chapters on continuum, from *Scientific Metaphysics*, come as a part of his doctrine of synechism (continuity of nature), following the explanations of tychism (his doctrine on chance). His Cambridge Conference Lecture 3 "The Logic of Relatives" follows his discussions on reasoning, and Lecture 8, "The Logic of

---

[95] Peirce (1976b: xv). For similar view on the priority of logic over mathematics in Peirce's treatment ment of continuity, see Dauben (1979: 321 n.4); Dauben (1982); Hintikka (1980).

[96] Peirce (1976b: 333).

[97] Peirce (1976b: 335. 351, 353, 363, 367).

Continuity". the concluding lecture of the series, follows his treatments of causation and habit. In these two lectures, 3 and 8, where most of his treatments of continuity are, Peirce's main focus is on how inquiry (philosophical, logical, scientific) must proceed: what rules of the argumentation are and how they are to be imposed. In this context, continuum serves as a demonstration of how our mind operates while it is engaged in the process of inquiry. And the central point is that this process of inquiry about the world must reflect the world itself if this process is to arrive at any truth (a veridical interpretation) at all.

If metaphysics enquires about the nature of the world, and ontology on its existence (for Peirce, ontology was part of metaphysics), semantics is the theory of signs as present to our cognition. For Peirce, mind is entirely codependent on the world and sees the world as the totality of *signs* (recall that for Wittgenstein in the TLP the world is (already) a totality of *facts*). This is the central point of his philosophy, semeiotics, logic and scientific method of investigation: that the world and cognition are reflections of each other and that reflection can be read through this theory of signs. This is why semeiotics is part of his metaphysics, and the theory of continuum (and infinitesimals) is part of his semeiotics.

In Lecture 8 Peirce brings together the notions of potentiality, of general, and of the fact that continuum is not a composition of the individual points. He says "That which is possible is in so far *general*, and as general, it ceases to be individual". He then goes on to define what the notion of potentiality means: "the word "potential" means *indeterminate yet capable of determination in any special case*". This can be exemplified in the notion of 'potential aggregates' as "greater in multitude than any possible multitude of individuals. But being a potential aggregate only, it does not contain any individuals at all. It only contains general conditions which *permit* the determination of individuals".[98]

Since continuum contains only possible or potential elements (not points), it is infinite. At the same time Peirce makes a clear distinction between collection and continuum. Unlike in continuum, a collection is made up of individual points and cannot be continuous: "A collection is a whole whose being consists in the independent being of its members; a line, on the contrary, has a being from which the being of its points is derived and in which they, as possibilities, are involved".[99]

If collection is the term for Cantor's *Menge*, then in this period Peirce makes a very clear distinction that continuum should not be treated the way sets are treated in mathematics. His logical treatment of continuum allows him to consider human community as that continuum without individual points, a continuum in which individuality is blended into communal generality.

2.5 Semeiotic Continuum (1908 – 1913)

From 1908 Peirce makes a clear distinction between collection as linearly arranged *actualized* points and continuum as that of *potential* points. Continuum begins to be treated

[98] Peirce (1992b: 247).

[99] From 1904 in Peirce (1976a: 53).

by his topology that studies the way "the parts of places are continuously connected". as he wrote in already in 1891.[100]

In his later semeiotics from 1907, Peirce begins to treat continuum as a goal–oriented denumerable multitude of potential points. It becomes more like time and has a very strong connection with consciousness:

> It means nothing at all to say that time is unbroken. For if we all fall into a sleeping–beauty sleep, and *time itself stops during the interruption*, the instant of going to sleep is absolutely unseparated from the instant of waking; and the interruption is merely in our way of thinking, not in time itself. There are many other curious points in my new analysis. Thus, I show that my true continuum might have room only for a denumeral multitude of points, or it might have room for just any abnumeral multitude of which the units are in themselves capable of being put in a linear relationship, or there might be room for all multitudes, supposing no multitude is contrary to a linear arrangement.[101]

In the same text Peirce distinguished two kinds of continuum: perfect and imperfect. Perfect continuum is: "the absolute generality with which two rules hold good, first, that every part has parts; and second, that every sufficiently small part has the same mode of immediate connection with others as every other has".[102] The imperfect continuum is the one that has topical singularities.

Cognition in general and consciousness particularly is associated with the perfect continuum: "My notion is that we directly perceive the continuity of consciousness; and if anybody objects, that which is not really continuous may *seem* so, I reply, "Aye, but it could not *seem* so, if there were not *some* consciousness that is so". I should like to see a good criticism of that reply".[103]

What continuum really becomes in the last years of Peirce is *modus operandi* of immediate, dynamic and final interpretants. To understand this perfect continuum we must look into his later semeiotics, for in this latest period, collections are treated as dynamic and immediate objects. In 1908 Peirce defines collection as:

> An object whose being is recognized as logically depending upon nothing whatever except upon the Being of each object that is termed a member of the collection, [in] which the Being of no member logically depends either immediately or mediately upon the being of another. The Being of the whole depends not merely on the Being of each Part, but also upon some relation between the parts.[104]

Here again, the treatment of a collection is logical rather than mathematical. But notice the double use of the word 'being' as a collection and 'Being' as an object. In the last sentence Peirce stresses the importance of dependency upon the relation between the parts of the collection. In the following paragraph the connection between being of a

---

[100] Peirce in Whitney (1891: 1360).

[101] From 26 May 1908 in Peirce (1933: 642).

[102] Peirce (1933: 642).

[103] From 1911 "Continuity Redefined" in Peirce (1935: 182).

[104] "Notes for a Paper for the American Mathematical Society: Logical Remarks on Some Mathematical Definitions". 19 December, 1908 in Peirce (1985: 1009).

collection and Being of continuum becomes evident: "A *collection* is an *ens rationis* whose Being consists in the Real Possibility of creating an Object of thought the Being of which object should imply, and should only imply, the Being of whatever Object of thought you'd have come to on any one occasion, in any one way".[105]

Notions of possibility (potentiality), time ("on any one occasion") and interpretation ("in any one way") indicate the necessary link between collection and continuum as between the object and interpretant. Peirce further continues by underlining the essential difference between continuum and collection in their internal relations between their parts and the whole: "The members of a collection may be called *parts* of the collection as their *whole*. But this involves a certain violence to the term *whole* and *part*; since in general a whole is that whole Being [which] consists in the Being of the Parts combined with those reductions between the parts which constitute the *ratio essendi* of the whole".[106]

Here too the use of 'being' and 'Being', 'parts' and 'Parts' is by no means accidental. When we say that collections are made up of parts and that collection is the whole of parts, the language is not exact due to our linguistic "violence to the term *whole* and *part*". Indeed, only of a continuum can we say as a Whole, a Part and a Being. Peirce concludes his comparison by stressing the dependence of a collection (as an object) upon a continuum (as an interpretant): "The Being of the essential whole consists in the essential parts with their essential relations; the state of being of an accidental whole consists in the being of all its parts together with all their relations. But the Being of a collection is entirely independent of the relations between its parts".[107]

Remember the frequent confusion that the notion of the object brings to the readers of Peirce, when he uses object as that which is interpreted and sometimes object as an interpretant. Here, as in no other period of his philosophy, logic and semeiotics are intimately interlinked. When the interpretant (mind or community) is interpreting object (collection), it becomes object itself, viz. the final object of interpretation in a continuum. This entire process of interpretation and metamorphosis of the object of interpretation can be called by the name of cognition.

One final point needs to be made in relation to Peirce's treatment of the continuum. All of his five notions of continuum use the concept of time. What he charged Cantor with on his circular definition of quantity as multitude can be said of Peirce's own definition of continuum:

1. Infinite: as an infinite divisibility.
2. Mathematical: as an "uninterrupted connection of parts in space and time".
3. Categorical: continuum of signs of future interpretations of the object.
4. Logical: as an infinite of possible elements (points).
5. Semeiotic: as a goal–oriented denumerable multitude of potential points in a time–like progression.

---

[105] Peirce (1985: 1009).

[106] Peirce (1985: 1010).

[107] Peirce (1985: 1010).

Peirce was conscious of that problem. In the same note from 26 May 1908 he acknowledged what Cantor stated in his *Grundlagen* by criticizing all previous definitions of continuum as time, from the ancient Greeks to Kant. Peirce writes:

> If my definition of continuity involves the notion of immediate connection, and my definition of immediate connection involves the notion of time; and the notion of time involves that of continuity, I am falling into a *circulus in definiendo*. But on analyzing carefully the idea of Time, I find that to say it is continuous is just like saying that the atomic weight of oxygen is 16, meaning that that shall be the standard for all other atomic weights. The one asserts no more of Time than the other asserts concerning the atomic weight of oxygen; that is, just nothing at all.[108]

According to Havenel, Peirce never succeeded in providing what he considered a satisfactory definition of continuum without the notion of time. As late as 1913, four months before his death, he was still trying to find new ways of treating continuum.[109]

3. Cognitive Semeiosis

3.1 Peirce's Semeiotic Philosophy of Mind

It would not be an overgeneralization to say that Peirce's semeiotics was the alternative to Cartesian philosophy of mind, whether in its original form of Descartes' writings or what permeated into 19th century philosophy of mind and philosophical psychology (See CP 5.81 quoted at the beginning of Chapter II). From 1868 to 1907 Peirce considered thoughts as signs interpreting signs. As T. L. Short indicates, the difference between early and later semeiotics consists in that the former restricted semeiosis to the interpretative activity of the thinking subject (still an overly–Cartesian *ego*), while the later extended the notion of interpreter to that of an interpretant, viz. beyond thinking subject to interpreting community.[110]

There is, however, another important difference already indicated earlier in this chapter. His earlier notion of considering signs as objects to be interpreted ad infinitum, in his later period was reconsidered as semeiotic continuum. Peirce's logical, rather than metaphysical or mathematical treatment of continuum allowed him to improve his semeiotics by exorcising from it the Cartesian ghost of dualism and made ontological, metaphysical and epistemic connections.

As one must recall, a stumbling block for Cartesian philosophy of mind was the problem of connection and interaction between *res cogitans* and *res extensa*, due to the essentially dualistic metaphysics. Peirce avoids that by insisting on human thinking as thinking by interpretation of signs. But the nature of signs is not intrinsically subjective. If that would be so, Peirce would not be saying much more than what John Locke had already

---

[108] Peirce (1933: 642). See Peirce (2010: 160–164).

[109] Havenel (2008: 123). For criticisms of Peirce's notion of continuum as time, see Thompson (1984) and Hartshorne (1983: 80–93, esp. 89–90).

[110] Short (2007: 289).

said, viz. that thoughts are signs. In Locke's division of sciences, 'the Doctrine of Signs' comes after natural philosophy and ethics:

> The Third Branch may be called *σημειωτική*, or *the Doctrine of Signs*, the most usual whereof being Words, it is aptly enough termed also *λογική*, Logick; the business whereof, is to consider the Nature of Signs, the Mind makes use of for the understanding of Things, or conveying its Knowledge to others. For since the Things, the Mind contemplates, are none of them, besides it self, present to the Understanding, 'tis necessary that something else, as a Sign or Representation of the thing it considers, should be present to it: And these are *Ideas*.[111]

However, the subjective treatment of signs brings Locke immediately to the problem of 'other minds':

> And because the Scene of *Ideas* that makes one Man's Thoughts, cannot be laid open to the immediate view of another, nor laid up any where but in the Memory, a no very sure Repository: Therefore to communicate our Thoughts to one another, as well as record them for our own use, Signs of our *Ideas* are also necessary. Those which Men have found most convenient, and therefore generally make use of, are articulate Sounds. The Consideration then of *Ideas* and *Words*, as the great Instruments of Knowledge, makes no despicable part of their Contemplation, who would take a view of human Knowledge in the whole Extent of it.[112]

The problem of signs that are necessary to communicate other signs was the resulting problem of Peirce's earlier semeiotics and the reason for its revisions in 1903. But unlike Locke, Peirce goes much further by implying an epistemological principle according to which all knowledge (including self–knowledge) is an inference from external experience, viz. external to the experiencing subject: "Any emotion is a predication concerning some object, and the chief difference between this and an objective intellectual judgment is that while the latter is relative to human nature or to mind in general, the former is relative to the particular circumstances and disposition of a particular man at a particular time".[113]

An inference from external experience *is* that predication about an object. It is not only semeiotic predication but also logical inference. The consequence (against Descartes and Locke) is that there is no need in any introspection: "It appears, therefore, that there is no reason for supposing a power of introspection; and, consequently, the only way of investigating a psychological question is by inference from external facts".[114]

The obvious conclusion for Peirce is that thinking is only possible through interpretation of the external objects and not as an introspection of the subjective ego. This implies that there is no need in postulating any ontologically separate *res cogitans* from *res extensa*:

---

[111] Locke (1975: IV, 21 §4).

[112] Locke (1975: IV, 21 §4).

[113] From 1868 in Peirce (1934: 247).

[114] Peirce (1934: 249).

> If we seek the light of external facts, the only cases of thought which we can find are of thought in signs. Plainly, no other thought can be evidenced by external facts. But we have seen that only by external facts can thought be known at all. The only thought, then, which can possibly be cognized is thought in signs. But thought which cannot be cognized does not exist. All thought, therefore, must necessarily be in signs.[115]

On the other hand, instead of adopting a reductionist, physicalist position concerning the mind/body problem, Peirce begins by underlining mind's dependence on the external objects, considering it as interpreting force of the physical objects in the world, equating mind to semeiosis, and finally concluding about a special case of a continuum identity between mind and the world. So, neither dualism nor monism (neither mind nor body exclusively) can be considered as Peirce's ontological subject of his philosophy of mind.

His theory of cognition as continuum is based on his consideration of cognition as *dialogic semeiosis*:

> All thinking is dialogic in form. Your self of one instant appeals to your deeper self for his assent. Consequently, all thinking is conducted in signs that are mainly of the same general structure as words; those which are not so, being of the nature of those signs of which we have need now and then in our converse with one another to eke out the defects of words, or symbols. These non–symbolic thought–signs are of two classes: first, pictures or diagrams or other images (I call them Icons) such as have to be used to explain the significations of words; and secondly, signs more or less analogous to symptoms (I call them Indices) of which the collateral observations, by which we know what a man is talking about, are examples. The Icons chiefly illustrate the significations of predicate–thoughts, the Indices the denotations of subject–thoughts. The substance of thoughts consists of these three species of ingredients.[116]

One must be warned against considering that all thinking is similar to verbal communication. By saying that "all thinking is conducted in signs that are mainly of the same general structure as words". Peirce does not imply here that signs are words, although the converse can be true. All thinking is done in signs, some of which cannot be expressed in words. The rest of the sentence after the semicolon makes it clear.

Peirce proceeds by giving two examples of "non–symbolic thought–signs": icons and indices. In the case of icons these are picture and diagrams. However, the list is much longer than the preceding text indicates. Just like Frege, Peirce too considered ordinary language to be inadequate for exact reasoning, praising mathematical and logical notation for their capacity to express the concise meaning. I.e. in his 1887 "Logical Machines". he considers algebra to be "the best of all instruments of thought".[117] His favorite example is the use of parenthesis in algebraic (and logical) notations.

The second example concerns indices that make clear the meaning of someone's state of mind. The same idea, as was shown in the previous chapter, will reemerge in Wittgenstein's refutation of any possibility of a private language.

---

[115] Peirce (1934: 251).

[116] From 1909 in Peirce (1935: 338).

[117] Peirce (1887: 169).

This last point is important in proving that the interpretant of the sign is neither (exclusively) the mind, nor that all sign–interpretation is expressible in ordinary language. By extending semeiosis to the world of objects, Peirce also extends the notion of signs' interpreter to signs' interpretants, viz. from individual subject to community:

> When we come to study the great principle of continuity and see how all is fluid and every point directly partakes the being of every other, it will appear that individualism and falsity are one and the same. Meantime, we know that man is not whole as long as he is single, that he is essentially a possible member of society. Especially, one man's experience is nothing, if it stands alone. If he sees what others cannot, we call it hallucination. It is not "my" experience, but "our" experience that has to be thought of; and this "us" has indefinite possibilities.[118]

"The great principle of continuity" tell us that individualism is plainly false: there are no individual points on the continuous line just as there are no individualistic minds. The remaining part of the above quoted paragraph indicates the communal characteristic of our cognitive processes, which depends on and is posterior to our dealings within our communities. Thus, it is community that comes to be seen by Peirce to be interpreter of signs, not an individual *ego*. How the semeiotic process actually unfolds was already demonstrated at the beginning of this chapter.

Knowledge and consciousness are not some general states of mind, and 'mind' cannot serve as a synonym for either 'knowledge' or 'consciousness'. Instead, both are special instances of cognitive activity.

As far as knowledge is concerned, objects in the world influence the mind to make an inference about the world. Knowledge cannot be defined, therefore, as a series of mental states:

> Every thought, however artificial and complex, is, so far as it is immediately present, a mere sensation without parts, and therefore, in itself, without similarity to any other, but incomparable with any other and absolutely *sui generis.* (…) Finally, no present actual thought (which is a mere feeling) has any meaning, any intellectual value; for this lies not in what is actually thought, but in what this thought may be connected with in representation by subsequent thoughts; so that the meaning of a thought is altogether something virtual.[119]

Knowledge cannot be pointed out, cannot be said to consist in or be located in this or that part. Instead, "the meaning of a thought is altogether something virtual".

The concept of virtuality was borrowed from Duns Scotus[120] and was defined by Peirce as: "A virtual *X* (where *X* is a common noun) is something, not an *X*, which has

---

[118] This is an additional note added in 1903 for the text written in 1893 in Peirce (1934: 402, n. 2).

[119] Peirce (1934: 289). "Accordingly, just as we say that a body is in motion, and not that motion is in a body we ought to say that we are in thought and not that thoughts are in us". From 1869 in Peirce (1934: 289, n.1).

[120] Duns Scotus (1954: 108–201). On Peirce's uses of Scotus' philosophy see Boler (1963). For Scotus, our predication about God as 'infinite being' includes such divine properties as 'infinite goodness', 'infinite love', etc. Thus, in 'infinite being' is 'virtually included' divine goodness, love,

the efficiency (virtus) of an *X*".[121] Virtual is defined in terms of efficiency and contrasted with potentiality. Potential *X* has a nature of *X* but without an actual efficiency. Peirce gives the example of the sun being "virtualiter" on earth, meaning that the sun is present on earth in its efficiency but without actually being *on* the earth. Knowledge is virtuality of objects in our cognition without being present either actually or at any particular moment of time.

Just as knowledge is not treated by Peirce as an essential feature of the mind, neither is consciousness. There is no localization of the consciousness just as there is no localization of the mind. In the following text Peirce not only resolutely denies dualism but distances himself (through ridicule) from physicalism and materialism.

The passage begins by Peirce's exposition of the argumentation by psychologists that consciousness is, in fact, the essential attribute of the mind. If one could also add that mind is the essential attribute of the human being Peirce's psychologist could go by the name of René Descartes:

> The psychologists say that consciousness is the essential attribute of mind; and that purpose is only a special modification. I hold that purpose, or rather, final causation, of which purpose is the conscious modification, is the essential subject of psychologists' own studies; and that consciousness is a special, and not a universal, accompaniment of mind.[122]

Hence, not being conscious does not cancel one's having mind. By attributing final causation to consciousness, Peirce credits mind with its purpose to be or to become conscious, which is not the same as saying that "being conscious" is the essential hallmark of mind as such. So much for dualism. Peirce then directs his attacks to the modern–day physicalist for whom cognitive capacities have their localization in the brain and are explained as brain states:

> A psychologist cuts out a lobe of my brain (*nihil animale me alienum puto*) and then, when I find I cannot express myself, he says, "You see your faculty of language was localized in that lobe". No doubt it was; and so, if he had filched my inkstand, I should not have been able to continue my discussion until I had got another. Yea, the very thoughts would not come to me. So my faculty of discussion is equally localized in my inkstand. It is localization in a sense in which a thing may be in two places at once.

In light of Peirce's previous definition of 'virtual' and his interpretation of the mind as something virtual, the last sentence about localization in two places at once must be also interpreted as this very same virtuality. Thus, mind can neither be localized in the brain lobe nor in the inkstand but virtually in the brain lobe as in the inkstand.

In a different passage Peirce is strongly arguing against physical localization ("in a space and time") of mind: "I do not say that we are ignorant of our states of mind. What

---

etc. For Scotus we can predicate 'infinite being' of God *univocally*. The same use of 'virtual' can be seen in Peirce: sun is virtually present on the earth; brain–lobe (or the inkstand) is virtually present to my faculty of expressing a thought. See Duns Scotus (1960: 258–277).

[121] Peirce's definition of 'virtual' in Peirce (1935: 372).

[122] From 1902 in Peirce (1958a: 366).

I say is that the mind is virtual, not in a series of moments, not capable of existing except in a space of time – nothing so far as it is at any one moment".[123]

One can visualize the continuous line of a thinking process originating from the brain and concluding on the piece of paper, with the inkstand being one of the intermediate objects. In this case, mind cannot be reduced to anything, be it any brain state or visual expression of one's thoughts and language.

Peter Skagestad commented on the passage above extensively and had drawn attention to the importance of the external manifestations of our cognition: inkstands, writing instruments, paper,[124] algebraic and logical notations, graphs and diagrams. It was also noted that similar ideas were expressed by Karl Popper in his 'World 3' where objects are considered as means of interpretation and misinterpretation of our ideas. Popper himself, though, dismissed such comparisons with Peirce.[125]

There is some importance to the external representations of cognition in Peirce. Here one can be reminded of Wittgenstein's insistence on incarnating one's language into one's actions. Likewise, we can recall Chekhov's and Dostoyevsky's plays with objects in expressing one's state of mind.

It seems, however, that Peirce's attention to the external instruments of thought was given no more attention than to the brain lobe in explaining cognition. Neither the 'brain lobe' nor the 'inkstand' are the key terms for his philosophy of mind but the idea of virtuality as efficiency. The remaining part of the paragraph gives only the metaphorical importance to the visual evidence of the working mind, be it physical or psychic:

> On the theory that the distinction between psychical and physical phenomena is the distinction between final and efficient causation, it is plain enough that the inkstand and the brain–lobe have the same general relation to the functions of the mind. I suppose that if I were to ask a modern psychologist whether he holds that the mind "resides" in the brain, he would pronounce that to be a crude expression; and yet he holds that the protoplasmal content of a brain–cell feels, I suppose: there is every evidence that it does so. This feeling, however, is consciousness. Consciousness, per se, is nothing else: and consciousness, he maintains, is Mind. So that he really does hold that Mind resides in, or is a property of, the brain–matter. The early students of electricity, who assumed that an electrical current resides in the metallic circuit, had infinitely more reason for their mistaken opinion. Yes, without exaggeration, infinitely more; for the ratio of something to nothing is infinite.[126]

The words "the inkstand and the brain–lobe have the same general relation to the functions of the mind" imply the general relation of efficiency or virtuality. But that is also a relation of a continuum in which the brain–lobe, the inkstand, pen and paper, etc. are the constitutive elements in that continuum. Physical presence is limited while

---

[123] From 1868–1869 in Peirce (1958b: 248).

[124] See Peirce's praise of Charles Babbage for his use of different colors of paper and ink for publishing tables of logarithms. This was meant, according to Peirce, to "maximize the cognitive value of tables". From 1871 in Peirce (1984: 459). See Skagestad (1999).

[125] Popper (1983).

[126] Peirce (1958a: 366).

virtual is not. This is why virtuality must be thought in a continuum and not as a mode of being for the physical objects, be it brain–lobe or an inkstand.

What the above passage evidently states is Peirce's opposition to any mathematical identification between the mind and any brain state. He reports that if we "were to ask a modern psychologist whether he holds that the mind "resides" in the brain" he would dismiss such a statement as an exaggeration. But Peirce proceeds by demonstrating that a modern psychologist's statements on identifying mental states to the particular brain states does commit him to acknowledge some identification. What sort of identity this might be will be shown by many theories of mind in the 20th century, examined in Chapter II together with presenting their failure in formulating any identity of mind and body. Here Peirce does not deny that there is some identity, just that this sort of identity escapes mathematics, just as his latest notion of continuum escapes mathematical precision.

When Peirce explained his thought experiment with the line being divided in two and the one point in the line becoming two points,[127] the division is not physical but virtual. This same idea is expressed by Peirce when in the previous lecture he explained syllogisms' use of the notion of identity:

> If the inference is drawn from more than one premise, let all the premises be colligated into one copulative proposition. Then this single premise must relate to S; and in that sense, it may be represented thus: S is M. I do not, of course, mean that S need appear formally in this premise as a subject, far less as the sole subject. I only mean that "S is M" may be in a general sense stand for any proposition which virtually relates to S.[128]

If a copulative proposition makes use of identity in a sense of virtual relation then why cannot the same point in a divided line be virtually the same? And if not, how can one avoid an obvious contradiction without applying the concept of virtual identity? Ketner suggests considering the identity of a single point as a monad, but in Peirce's lectures the term 'monad' appears only in his distinction between different verbs (such as medads, monads, dyads, triads, etc.) in his use of syllogisms and not in treating the identity.[129]

In one of the unidentified fragments in his manuscripts Peirce spoke clearly against localization of the mind and for an identity of the dynamic continuity of cognition:

> I have several times argued, at some length, that the unity of personality is in some measure illusory, that our ideas are not so entirely in the grasp of an *ego* as we fancy that they are, that personal identity differs rather in degree than in kind from the unity of "public opinion" and gregarious intelligence, and that there is a sort of identity of dynamic continuity in all intelligence.[130]

Here too Peirce goes against dualism and physicalism, against likewise the separation of the thinking *ego* from the rest of the world as well as against identifying that *ego* in

---

[127] Peirce (1992b: 159).

[128] Peirce (1992b: 131).

[129] Peirce (1992b: 154).

[130] Peirce (1976c: ix).

one form or another of the physical world. What brings unity to personality is "identity of dynamic continuity". That unity allows neither for dualism nor for materialism.

The ontological question posed by Peirce is: "In what does the reality of the mind consist?"[131] The answer is in claiming that the reality of the mind consists in its being a system of linguistic signs. In commenting on that passage (CP 5. 313–314), Torjus Midtgarden individuated two important ontological arguments in Peirce's theory of cognition in relation to his treatment of personal identity.[132] The two ways in which "the reality of the mind" consists coincides with the complex system of semeiotics that we are already familiar with.

The first is "translation of mental signs". viz. those signs that ultimately constitute our "psycho–physical biography". The second is transformation of the mental signs into an abstract linguistic sign system, in other words, our ordinary language. The conclusion is that our empirical self is dependent upon language and that makes the self dependent on the publicly accessible system of signs.[133]

However, one must keep in mind that signs and language were not always equated in Peirce's thought, just as interpretant was not always considered as the subject's mind (an interpreter). This, in fact, puts Peirce at odds with W. James, for whom "the cortex is the sole organ of consciousness in man". and for whom language often coincides with natural speech.[134] Midtgarden further makes some comparisons with John Dewey's functionalism and naturalism that is in contradiction with Peirce's disavowal from any localization of cognition, except the virtual one.[135]

3.2. Word, Belief, and Cognition

I would like to conclude this section by examining three texts, written twenty–five to thirty years apart, belonging to the early and middle periods of Peirce's semeiotics. This comparison will indicate the difference in Peirce's understanding of what language is. In the first text (from 1868) the human subject is equated with the word, with the spoken ordinary language. In the second text (from 1893) the notion of language is extended from word to behavior, human way of life, history and attitude, and it brings Peirce very close to what Wittgenstein will think of as language–games. The third text (from 1898) underlines Peirce's belief in the unity of person and his/her mind without denying the physical reality of our personhood.

3.2.1 *Metaphysics of Mind*

In the first text Peirce answers the ontological question:

---

[131] From 1868 in Peirce (1934: 313).

[132] Midtgarden (2002).

[133] Midtgarden (2002: 110–112. 114).

[134] James (1950: 66. 137).

[135] Midtgarden (2002: 118–121).

> In what does the reality of the mind consist? We have seen that the content of consciousness, the entire phenomenal manifestation of mind, is a sign resulting from inference. Upon our principle, therefore, that the absolutely incognizable does not exist, so that the phenomenal manifestation of a substance is the substance, we must conclude that the mind is a sign developing according to the laws of inference. What distinguishes a man from a word? There is a distinction doubtless. The material qualities, the forces which constitute the pure denotative application, and the meaning of the human sign, are all exceedingly complicated in comparison with those of the word. But these differences are only relative.[136]

The mind "is a sign resulting from inference". and inference changes depending upon what stage of inquiry one is at. This is why "the mind is a sign developing according to the laws of inference". If the mind is a sign, the man is the word and the external sign, for he thinks in terms of signs:

> As the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thought, proves that man is a sign; so, that every thought is an *external* sign, proves that man is an external sign. That is to say, the man and the external sign are identical, in the same sense in which the words *homo* and *man* are identical. Thus my language is the sum total of myself; for the man is the thought.[137]

The spoken word is an interpreted word, according to the rules of inference. Here, semeiotics and philosophy of mind coincide, resulting in the conclusion that the human mind cannot be interpreted in any other way from how signs are interpreted. After 1906 Peirce will also add that whatever localization of human cognition (consciousness included) at any time one imposes it, it must be dynamic and virtual. It is never at any particular time or space but across time and space in a continuum.

3.2.2 *Semantics of Belief*

Compare, however, Peirce's notion of language in 1893:

> Our daily life is full of involuntary determinations of belief. It is the egotism of the ego, or field of attention, which imposes upon [us] with its High German modest conviction that whatever is known is known through it. It is not so. I converse with a man and learn how he is thinking: I fancy he has *told me,* that is, has "stated" the fact in accurate forms of speech. But he has not, and how I have found out[1] his thought is too subtle a process for this psychologist writing to find out[2]. You hear[1] a new slang word: you never ask for a definition of it; and you never get[1] one. You do not get[2] even any simple example of its use; you only hear[2] it in ironical, twisted, humorous sentences whose meaning is turned inside out and tied in a hard knot; yet you know what that word means much better than any abstract definition could have informed you. In

[136] Peirce (1934: 313).
[137] Peirce (1934: 314).

> riding a horse; rider and ridden understand one another in [a] way of which the former can no more give an account than the latter.[138]

The use of the verbs following pronouns 'I' and 'you' give the clue to the epistemological process Peirce invites us to examine in this text. The verbs are arranged in the progressive manner from 'converse' to 'know/inform': converse, learn, fancy, find out (twice), hear (first), ask (in negative), get (twice in negative), hear (second), know and inform.

How does one come to know the definition of a term ("slang word")? If man is the word and mind and the man is likewise indistinguishable from the (external) sign, then understanding language as "the sum total" of the self is equivalent to understanding the person. Hence, the notion of language, just as the notion of the sign, is essential here: we read each other as reading and interpreting signs. Unlike in the previous text, language here is not equated with the spoken or written word. It is not about the inkstand anymore: one must converse, learn, hear and ask to have knowledge about the meaning of the word.

The easy temptation would be the same as with Wittgenstein to read the above passage in behavioristic terms. But every variety of behaviorism is reductive and the richness of Peirce's exercise does not allow for any reduction. It is not just some sum of behavior that informs one but the continuum of learning from different instances of interpretation until nothing is left to be interpreted. Indeed, that last *telos* is the objective certainty which Wittgenstein did not consider to call 'knowledge' and Peirce speaks here of being 'informed' instead. These ideas were foundational for Hintikka's epistemic logic where the goal is not justification of belief but information acquisition (scientific discovery) and manipulation (games of seeking and finding).[139]

The last sentence of the above text uses allegory of the "rider and ridden". which might remind one of the ancient Chariot Allegory in Plato's Phaedrus.[140] The image of the tripartite *psuchê* is the continuous conflict (until enlightenment of the contemplation of the Ideas) between the rider (rational *psuchê*) and the two horses: the passionate and the appetitive ones. In Peirce, just as in Wittgenstein, there is a harmony of understanding without conflict, and the enlightenment of certainty from following the rules of a game without the epistemic process of discovering the true knowledge of Ideas.

The previous passage is also very telling of Peirce's teleological notion of intentionality. I believe that on this notion Wittgenstein and Peirce are on the same page. Intentionality of knowledge *shows* in our purposeful actions, while ordinary language *alone* cannot make distinctions between existing and non–existing objects. Intentionality in both authors is understood as the interpretation of signs, and in Peirce's later semeiotic it is an interpretation of a dynamic object as a sign. Liszka suggests that only a dynamic object can be considered as existing and only of the dynamic object can there be a final interpretant.[141] Short rejects the possibility of extending Peirce's semeiotic to biology for the

---

[138] Peirce (1958a: 447). Italics are the authors; all underlining and numeration of the verbs are mine.

[139] Hintikka (1962); Hintikka (1968: 37–51); Hintikka (1973); Hintikka (2007: 11–37).

[140] Plato (1953: 246a–54e).

[141] Liszka (1996: 111n.2–116n.2).

reason that on a biological level there can be no intentionality and without intentionality there cannot be semeiotic interpretation.[142]

However, making intentionality as an essential process of all semeiosis is an exaggeration. Peirce, at various periods, had argued that some information interpretation does not require intentionality.[143]

3.2.3 *Unity of Cognition*

In the text from 1898 Peirce spoke clearly of the unity and harmony of the mind that allows for mutual understanding and knowledge/information acquisition:

> The consciousness of a general idea has a certain "unity of the ego". in it, which is identical when it passes from one mind to another. It is, therefore, quite analogous to a person; and, indeed, a person is only a particular kind of general idea. Long ago, in the *Journal of Speculative Philosophy* (Vol. II, p. 156[144]), I pointed out that a person is nothing but a symbol involving a general idea; but my views were, then, too nominalistic to enable me to see that every general idea has the unified living feeling of a person.[145]

Here as well we hear Peirce's notion of the unity between person and his/her consciousness. But that unity neither equates consciousness with the mind, nor makes the mind the only true 'substance' ignoring the physical body. The unity in continuum, instead, allows for mutual knowledge and understanding.

Conclusion

The five questions of the philosophy of mind have been answered by exploring Peirce's semeiotic, theory of continuum and his philosophy of mind. In this short conclusion I will only sum up his ideas in relation to contemporary discussions.

Peirce was neither dualist, behaviorist, nor monist. His phenomenology and idealism never permeated into saying that physical reality is unreal or that 'real' objects are only phenomenological in nature. Thus, his notions of mind and body never were reductionistic in any way.

The ontological question, inseparable from his metaphysics and semeiotics, is answered by Peirce with striking clarity: the mind is a system of linguistic sign. The mind is about the sign and the sign is about the world. The correspondence between the mind and the world is never isomorphic and never static. The dynamism of this correspondence depends on our process of inquiry, on the metamorphosis of object in interpretation, from immediate to dynamic.

---

[142] Short (2007: 174–177).

[143] From 1866 in Peirce (1982: 471–475); From 1903 in Peirce (1998: 193).

[144] From 1869 in Peirce (1934: 264–317).

[145] Peirce (1935: 270).

The mind "is a sign resulting form inference". It is here that metaphysics, ontology, epistemology, and logic come together. With Peirce it is difficult to treat them separately. Furthermore, cognition is a virtual identity between its physical and mental elements. This semeiotic continuum is goal–oriented and purposeful in nature. Cognition is a semeiotic continuum resulting from inference. We come to know the world and ourselves through a dialogical semeiosis; epistemology and semantics are inseparable in Peirce's treatment of knowledge.

The notion of virtual identity is logical and semeiotic, not mathematical. It avoids the nonsense and inconsistency pointed out by Wittgenstein. Although Scotus' virtual identity is univocal, it cannot be transcribed in the mathematical symbol of identity. This is not material or formal but efficient and final identity: an identity in progress toward realization. In this context, Peirce's example with the inkstand is very important in determining that brain states are virtually present in every element of the cognitive continuum. What is real is present and what is present on every step of semeiosis is virtual. Finally, Peirce's notion of modality is perfectly compatible with the reversed multiple realizability thesis stating that a single brain state can be multiply realized in many mental states.

Despite Peirce's rich theory of signs and that of continuum, his philosophy of mind is very much incomplete. We find no explanation of how brain states fit as elements of continuum, continuum being semeiotic cognition. We have no clarification about the semeiotic process that unites us with the world. His language is descriptive and general. We have many descriptions of 'what' but few if any of 'how'. Finally, his concepts of continuity were circular in containing in them the notion of time.

In these regards, Peirce is the opposite of Wittgenstein: the later was resistant to theorizing; the former was too general in his explanations. The opposites can complete each other and we must turn to the language–games again for the explanations and answers to our questions.

CHAPTER V

Philosophy of Cognition: Theoretical Foundations and Methodological Proposals

Introduction

In this concluding chapter I will formulate and defend the main thesis of this work that comes as a response to the problem of choosing between forms of dualism and reductionism (Part One, Chapters I and II) and as an application of the alternative methods of Peirce and Wittgenstein in the philosophy of mind (Chapters III and IV).

Part One sets the problem of treating human mind and body in dualistic or reductionist terms, which imposed serious methodological constraints on itself resulting in the impossibility of answering the fundamental questions of existence, nature, knowledge, meaning and identity. These were proposed questions as methodological tools in examining the major theories in philosophical psychology and philosophy of mind at the beginning of this work. In Chapters III and IV we have seen how Wittgenstein and Peirce approached these questions in their own ways and, through their theories of signs and notions of language–games, what their answers and their reformulations of the questions were.

My current and conclusive suggestion that I will elaborate in the remaining pages is to apply Peirce's cognitive semeiotics and Wittgenstein's language–games in, first, reformulating the dualistically defined mind/body problem as the problem of cognition without reductionism, and second, to provide the study of cognition as a study of both, cognitive semeiosis and language–games.

This proposed study, however, should not be considered as a comparative study between Wittgenstein and Peirce. As I explained in the Introduction to Chapter III, there are some fundamental philosophical differences between the two authors that significantly prevent one from conducting a true comparison between them. They belong not to just different but opposite philosophies of language with significant consequences in opposite views in epistemology, logic, treatment of mathematics and semantics. Instead, this present study will use Wittgenstein's language–games and Peirce's semeiotics as instruments for constructing, different from their own philosophies of mind, a philosophy of cognition. Thus, by the end of this chapter I will not ascribe my own

conclusions to the views of either of the authors. Their importance for the topic was demonstrated in the previous two chapters. However, the study that emerges from their theories is different in its conclusions, albeit similar in scope.

Since neither Peirce's theory of signs (semeiotics), the cognitive observations of signs (cognitive semeiosis), nor Wittgenstein's notion of language–games are sufficient for a comprehensive philosophy of mind, I begin the study in this chapter by considering them as completing each other in their explanation of cognition. Section 2 will provide an alternative to Peirce's time–dependent definition of continuum. The suggested notion of continuum is the modality of the rule–guided activities also defined as the language–games. Section 3 will define cognition as a four–dimensional language–game and will suggest to study cognition through the study of these dimensions. Section 4 will further elaborate on Peirce's concept of virtual identity and will suggest to view cognition as a virtual unity between lower and higher cognitive states. Section 5 will spell out the consequences of these studies for the methodology of the philosophy of mind, while the conclusive suggestion would be to change the dualistic concept of mind for the number–neutral notion of cognition.

1. Semeiosis of Language–Games and Language–Games of Semeiosis

Charles S. Peirce's semeiotics, the theory of signs, is an auxiliary theory for his predicate logic and for his theory of knowledge. The first was elaborated as the method of scientific inquiry and the second as a semeiotic process of knowledge acquisition to be the method of scientific discovery. Furthermore, it is an intermediate doctrine between logic and epistemology on the one hand, and his philosophy of mind on the other. It is not enough (at least not for Peirce) to indicate *that* scientific discovery is the process of knowledge acquisition being a triadic relation between the object, the sign of that object and the interpretant. It is also necessary to say *how* we come to regard sign related to the object and in *what* capacity the interpretant stands between the sign and the object. In other words, it is not sufficient to remain on the epistemological level of logical inquiry or on the ontological level of the world–signs relations. The unity of science and the unity of our inquiry will be complete only if we bring together the metaphysical question of the nature of the mind that performs the logical inference and epistemological inquiry of scientific discovery.

This unity between ontology, metaphysics and epistemology through the semantics of semeiotics was essential to Peirce's notion of continuity between what is known (the world) and what knows (subject, mind). This distinction between the known and the knower, however, is only an abstract distinction in philosophical language, while in fact both, the known and the knower are non–individualized elements of the same continuum. In fact, this continuity named later as his theory of synechism ("of prime importance in philosophy"[1]) was a guarantee that our knowledge of the world is true, since the known and the knower are part of that same continuum. This is why Peirce

[1] From 1902 in Baldwin (1905: II, 657); Peirce (1935: 169).

rejected dualism and reductionism: speaking of separate realities is false, while indicating only its physical part is deficient.

In the previous chapter I said that Peirce's philosophy of mind was incomplete precisely on the question of how mind is related to the world of objects, signs and interpretants. Studies have been done to complete this lack by a comparative study with Wittgenstein's language–games.[2] But language–games were incomplete themselves in explaining how they "are [the] ways of using signs"[3] and how signs acquire their living force through the uses within rule–guided activities which are language–games.[4]

Despite these obvious insufficiencies of Peirce's semeiotics and Wittgenstein's language–games, when it comes to the explanation of the mind they complement each other. Wittgenstein did not have semeiotics (nor semiology) as a theory of signs, their use and interpretations. This should not be surprising due to his generally anti–theoretical approach to philosophy. In what way can Peirce's semeiotics be applied to Wittgenstein's language–games?

It is the fourth and the latest account of semeiotics in Peirce that is most compatible in my judgment to Wittgenstein's language–games especially when these are applied to the philosophy of mind. We can recall that this latest revision of semeiotics was brought about by Peirce's revisions in logic and philosophy of science. The two major problems with his earlier accounts of semeiotics were an insufficient theory of meaning and an undeterminable account of scientific inquiry. Both problems can be reduced to one: Peirce's notion of infinite progression in signs interpreting signs. If the chain of signs is infinite then neither can the meaning be determined nor is scientific knowledge complete. While it is true that further interpretations and further enrichments in meaning are possible, it is also true that some meaning is finite and some scientific inquiries are completed. Very often scientific inquiry consists in simple questions and its success is determined by simple answers, no matter how difficult the scientific inquiry itself is.[5] Likewise, communication based on the shared meaning of the words and sentences is made possible due to the determinate character of the newly or previously shared meaning of these words and sentences.

Between 1906 and 1910 Peirce makes a very important correction to his theory of signs when instead of infinity being accountable in explaining the succession of the signs, he suggested the notion of continuum. Unlike infinity, continuum is a goal–oriented and purpose–directed inquiry. The end of the inquiry is complete knowledge of the object. Now, this end can be real (simple questions answered) or can be ideal (open–ended questions without answers). The fact that some questions cannot be answered now does not diminish their nature of being genuine scientific questions. Furthermore, any inquiry is teleological by the fact that it is always oriented toward its end.

---

[2] Pietarinen (2006: ch. 8 and ch. 12); Legg (2008).

[3] Wittgenstein (1958: 17).

[4] See Wittgenstein (2001: §432).

[5] I.e. J. Hintikka's 'yes–or–no' questions and 'wh–questions' approached by the language–games for quantifiers of seeking and finding in Hintikka (1985) and Hintikka (1999: 183–204).

In the last chapter it was said that the end of the inquiry is the dynamic object, while the immediate object of the sign is the object at some point (starting or intermediate) of the inquiry. The later is fallible and indeterminate in its meaning. What determines it, is the use within the rules of the inquiry. These rules can be specified by the way language–games operate.

If my fragmented and incomplete comparative study between Peirce's theory of signs and Wittgenstein's non–theoretical language–games is correct, the immediate object is the set of rules within the primitive language–game. Since primitive language–games are intentionality–, knowledge–, and meaning–neutral, the immediate object is undetermined. The primitive language–games, with their own rules and strategies, are what Peirce considered to be the immediate interpretants. An immediate interpretant is our initial familiarization with the object of semeiotic inquiry. At this stage meaning and knowledge are too incomplete to form judgments. However, the immediate interpretants are essential for non–epistemic perceptions when all we have is the data that must be still accommodated within the larger context of other perceptions, the background of our knowledge and experience. Chapter III spoke about non–epistemic perceptions in terms of the primitive language–games. The immediate interpretants being at the beginning of the continuum of semeiotic inquiry are the elements of these primitive language–games that often stand in continuum of many interrelated non–epistemic perceptions. These are still to be interpreted pieces of information within the larger backgrounds of our already acquired knowledge and meaning.

The dynamic interpretant is where our complete and final knowledge and meaning are formed within the process of semeiotic inquiry. Here is where, i.e. non–epistemic perceptions acquire their epistemic and semantic values and complete language–games are formed. However, at this point nothing tells us that such a process could not go ad infinitum, for complete language–games can be perfectly interpreted as primitive ones for the higher level complete language–games, (see 'Psychoanalysis through the Language–Games') and so on. Peirce further introduced a final (or 'normal') interpretant as the "effect that would be produced on the mind by the sign after sufficient development of thought".[6] When the final interpretant is reached, the immediate and the dynamic objects now coincide in one making possible final meaning and final knowledge. Again, such a goal might be actual or ideal, but it is still the goal and the end of any semeiotic, scientific, semantic and epistemic inquiry.

What does it mean in the language–game nomenclature? It cannot mean no more strategies, for when such great complexity of rules is reached, the amount of strategies is truly infinite. Just think of the infinite amount of strategies in a relatively simple (in terms of defining rules) game of chess. Instead, the final interpretant means that no more rules can be instantiated over the complete language–game whose outcome is the final knowledge or understanding of the object of inquiry. Strategies can implement further uses of that knowledge but they cannot instantiate new rules and new language–games. To give the simplest example, one can throw a ball in all possible directions but it will follow the force of gravity with every throw. In relation to the inquiry about gravity (i.e. its center) no further language–games can be suggested no matter

[6] From 1908 in Peirce (1958b: 343).

how the game of throwing the ball could be played. In terms of the meaning of a word or an expression, we can think of the infinite ways the same word can be used but with the limited number of senses we assign to the word (i.e. its dictionary entries). While the senses of a single word or expression can increase with time, the new senses will not be the continuation of the previous language–games but results of instantiations of new ones from the primitive language–games, or from the newly instantiated immediate objects. For example, the relatively new senses of the words 'window', 'mouse' and 'desktop' which we use in our computer language are not extensions of the language–games of architecture, the study of rodents or new kinds of furniture. In sum, these are not just new senses of the same words, these are completely new language–games, or new tokens of the completely different types.

Some scientific inquiries, even if formulated in simple and straightforward syllogisms, will always remain open for further investigations and potentially new information acquisition. However, simplicity in science translates into precision of formulating premises in any inquiry. While every inquiry will always be open to an infinitely many objects and questions about them, one of the most important tasks for a scientist is to select the limited amount of premises that can be admitted to the questioning. In other words, a scientist cannot admit every fact as an evidence, for this would render inquiry impossible to conduct.

Here we must make a clear distinction between the infinite number of the possible language–games and infinity within a single language–game. Admitting the first is necessary, but admitting the second is to make a language–game impossible to play, for one single language–game cannot be played forever. Similarly, Peirce abandoned his earlier notion of infinity in semeiotics for the same reason: the meaning of the sign cannot be infinite. It can be *continuous*, viz. it can develop from one interpretant to another, but at least potentially it must be fixed.

As will be explained later, continuity as modality of the language–games allows for an infinite amount of language–games to be played and infinite ways in which one and the same language–game can be played. Still, once the rules are given and strategies are being played, once we transform primitive language–games into complete and more complex ones, each language–game, just as each inquiry, must come to some kind of completion or, using the game–theory jargon, must have the end–game. The end–game, or game's outcome, is encoded in the rules of each game. The single game of chess, i.e. cannot go forever for it will violate the rule of the game of chess (triple repetition of the same move in chess automatically means a draw). Inquiry, just as the game of chess, can also end up in a draw, viz. without providing any definite answer, but that does not mean that that particular inquiry is not ended: given the premises and their application rules, the result is inconclusive. Introduction of the new information being an introduction of new premises means the beginning of a new game. Here we are dealing with two infinities: one of the games in general and one of the single inquiry. The first kind of infinity was accepted by both Wittgenstein (see PI §23a) and Peirce, while the second was refuted by both of them.

In the section six of Chapter III I contrasted Wittgenstein's notion of the language–games with mathematical game theory proposed by the formalists. I also indicated that neuroeconomics, a new discipline, uses the notion of games in their description of hu-

man behavior. Furthermore, game theory was widely applied to the evolutionary theory by J. Maynard Smith and G. R. Price.[7] The authors introduced the notion of 'limited war' strategy being a conflict between animals of the same species. Since 'limited war' strategy causes little harm to the same species of animals, the strategy serves as species selection for the benefit of the species rather than that of individuals.

Again, the difference between game theory and language–games consists in the fact that even among animals a winning strategy comes through selection, whether one accepts egoistic behavior within the species or altruistic behavior within a particular group. Language–games, on the other hand, operate in terms of logical inference even on the most basic, primitive level. This is why I insist on the strong connection between semeiotic theory and the language–games notion. Both are essentially logical inferences to which natural selection is rather irrelevant. On the other hand the reverse is not true, for natural selection as the major force of evolution continues to be relevant for social, cultural and scientific development.

However, here I do not refer to the logical inference we find in sentential or predicate logic. Both semeiotics and language–games are more basic kinds of logical inferences than simple conclusions following premises (treated by Wittgenstein as tautologies anyway). The kind of logical inference that both semeiotics and language–games allow for strategies come from the rules, one game from another, one form of inference (or game) being instantiated by another, etc. These are inferences of discovery of new information when unlike in predicate or sentential inferences nothing seems to follow from the information already given. When we observe somebody's facial expression (the object for Peirce and the primitive language–game for Wittgenstein), we cannot conclude anything from that information alone. We need to add another object or set of rules from a different domain or a different game (viz. from a different family of language–games) to make sense of the information to consider it as a sign, index or a symbol. What makes these connections possible is the continuum of games and signs from the continuum of objects and rules.

Here the dynamics of the semeiotics and language–game inference are the same: from the possible inference of the immediate objects and primitive language–games, through the best inference of the dynamic objects and complete language–games, to the only possible inference of the final objects and the family–resembled language–games forming into entire forms of life.

However, cognition explained through language–games is not limited to human cognition alone. The kind of inference present in semeiotics and in the language–games is perfectly compatible with animal cognition as well. The difference between animal and human cognition is that the former is incomplete due to the lack of symbolic language, subjective self–consciousness and free will. Without entering into a detailed discussion on the difficult problems surrounding the origin of language, self–consciousness and free will, nor attempting to explain them in this work, it will suffice to say that these are manifestations of the complexity of the family–resembling language–games constituting entire forms of life. Therefore, any further explanations of self–consciousness and free will should be done in terms of the developments *between*

[7] Maynard Smith and Price (1973).

rather than *within* particular language–games. Concerning the origin of language, it could be described as the evolution of signs into symbols, thus implementing Peirce's notion of semeiotic continuum.

The conclusion of section six in Chapter III was that game theory cannot be applied to the study of human cognition in the univocal way, that is, the way that language–games and semeiotics can. Their use of the notions of the game's rules and strategies apply equivocally. The major method of any game theory is the choice of the best available strategy which is spelled out in decision making. The major problem of any game theory is to explain how rules generate strategy. Formalist mathematics cannot solve this problem because solving it would go contrary to the essence of mathematical formalism, which consists in the assumption that mathematics is nothing more than the manipulation of axioms, theorems and other rules. Rearranging rules *somehow* brings out the strategies of a game, but an explanation of *how* this 'somehow' happens would have to include a non–circular reference from the rules toward something that the rules generate. In language–games it is human activity but in formalist game theory these are rules themselves. For someone like Frege, Wittgenstein, and Peirce after 1903, this would not do, for one must find a better explanation for the foundations of mathematics than pure symbolism and symbol manipulations.

An animal, or human, makes a decision between attacking and withdrawing. This is a decision between two strategies, and game theory studies how this decision comes about and which strategy should be chosen given a specific situation (game's rules and available strategies). Language–game study, on the other hand, explains how particular rules generate strategies but it will not inform about which strategy is the best nor say how we come to choose one strategy over the other.

Game theory had great success and continues to have it in economics and in general decision theories, even if it deals with the already available sets of strategies without asking how and where the strategies came about. As I indicated in Chapter III, the fact that strategies can be rewritten solely by their constitutive rules cannot be an indication that strategies come about by mere reshuffling of their rules. Similarly, language–games, unlike games in the game theory, cannot be simply regarded as some unity of rules and strategies. Language–games come in four dimensions and can only serve as representative links between reality and language as four–dimensional activities.

If we explain semeiotics in terms of language–games and the operations of language–games in terms of semeiotic inquiry, we use all four dimensions of the language–games. The first dimension of the rules and strategies is essential in explaining how the immediate object is used in the immediate interpretant. The second dimension of the primitive and complete language–games explains further the development of the semeiotic inquiry and the ways in which the immediate object is further understood by the dynamic interpretant. The third dimension of the family–resembled language–games and these families extended into entire forms of life elaborates on the final interpretant of the dynamic object. Because the semeiotic process of inquiry is not infinite but continuous, the fourth and final dimension of the language–games as continuum and continuum as the language–games' modality cuts through all other three dimensions. This last point must be explained in detail.

2. Continuum as the Language–Game's Modality

In the previous two chapters we have seen that both Wittgenstein and Peirce, mostly in opposition to the that–day empiricist (W. James) and reductionist (B. Russell) approaches to the study of cognition, have spoken against its localization the same way physical objects are. Thoughts and feelings cannot be time–and space–dependent.[8] This was particularly clear in Peirce's later 'logical' and 'semeiotic' continua according to which our thoughts, feelings and consciousness in general are not part of some continuum, like individual points on a line, but are a continuum themselves.

The fact that cognition cannot be localized not only implied that it cannot be treated mathematically, quantitatively, and logically, but also brought Peirce to consider it as a semiotic continuum. This comes as a consequence of his long struggle against Cantor's treatment of continuum in terms of mathematical set theory. This struggle brought him to his own alternative to Cantor's philosophical, more than mathematical notions, of continuum applied in logic and semeiotics. We have seen, however, that the time–dependency of Peirce's later notions of continuum was an important deficiency of which he was well aware.

In 1868 Peirce wrote, "The mind is a sign developing according to the laws of inference".[9] In the previous chapter I suggested to view this semeiotic development also in terms of Wittgenstein's language–games, keeping in mind that "every sign *by itself* is dead", just as a lifeless point on the line or a subset of the power set, "but *what* gives it life? – It *lives* by use".[10] I specified that this use is the rule–guided activity, viz. language–games. Regardless of the fundamental philosophical differences between Peirce and Wittgenstein, it is in this situation that both complement each other. Wittgenstein lays down the explanation of how we use the signs through the strategies instantiated from the rules and through the web of language–games that implement the signs. What are the signs, how they can be classified is not explained in Wittgenstein but in Peirce, in his most complex and rich theory of signs.

If cognition is a dynamic language–game, then continuum is its fourth and final dimension. In the last chapter I indicated Peirce's own dissatisfaction with his five notions of continuum being defined in temporal terms. Although from 1892 he considers his notion of continuum in more philosophical than mathematical terms, he could not avoid the temporal treatment of it. His categorical, logic and semeiotic continua became more incorporated into his metaphysics, logic and semeiotics, thus being implemented into his philosophy of mind. Cognition itself, as we have seen in the previous chapter's section three, is semeiotic continuum, but it cannot be seen only in temporary terms.

In what ways can language–games help to break away from the time–dependency of Peirce's continuum? If continuum is one of the four dimensions of the language–

[8] See the examined texts in Wittgenstein, i.e. Wittgenstein (2001: §§243–315, §339; II, viii); Wittgenstein (1981: §§608–611), and in Peirce, i.e. Peirce (1934: 289); Peirce (1958a: 366); Peirce (1984: 240ff); Peirce (1976c: ix).

[9] Peirce (1934: 289).

[10] Wittgenstein (2001: §432).

games, I suggest defining it as the mode of the language–game's development. Bringing together the other three dimensions, semiotic continuum is the mode in which language–games within these three dimensions develop. This means the mode in which rules constitute strategies, strategies are composed from the rules; the mode in which primitive games constitute complete games and complete games become primitive for further complete games; the mode in which family–resemblances are organized and constitute the entire forms of life.

Here I understand the word 'mode' in its most basic etymological definition. *The Concise Oxford Dictionary* defines 'mode' as "1 a way or manner in which a thing is done; a method of procedure. 2 a prevailing fashion or custom". Thus, mode and modality are the ways in which we use the signs according to the rules of inference. In other words, it is the way or a custom in which a particular language–game or its families are developing.

However, as the same dictionary instructs, the word comes from the Latin *mŏdus* and in Latin the meaning is much more precise. *Cassell's Latin Dictionary* defines *mŏdus* as "a measure, standard of measurement; size, quantity, length". It has a special use as a musical term for "rhythm, melody, time". We can also speak of *mŏdus* as poetry's *tempo*. Finally, its similar use permeated into English as "manner, mode, fashion, way, method", such as "in that manner, of that kind".

The word 'mode' in terms of measurement and quantity is neutral and unsaturated. It tells us nothing about which measurement should be used. In this sense continuum does not need to be temporal. Continuum as a mode defines the way language–game is played. It could be defined as a way of performing some task, a style of action, strength or weakness of performance. Continuum is an adjective and an adverb applied to the each and every step within every language–game: "Proceed cautiously", "Be quiet", "Be persistent", etc. In PI §23a Wittgenstein speaks of the countless language–games as countless kinds of sentences ("*Arten der Sätze*"). Although he himself admitted only three dimensions of the language–games, what I suggest by continuum as a mode being their fourth dimension is the adjectival and adverbial mode of these countless kinds of sentences.

Mode as measurement also includes degree in performing particular language–games. These modes, together with strategies, family–resemblances and forms of life, are other modalities of the language–games. In other words, one and the same language–game can be performed in many different modes.

In how many ways can a language–game develop? In the same PI §23a, just before giving a variety of language–games, Wittgenstein says that there are countless ("*unzählige*") kinds of sentences that are used in always increasing new language–games. One single language–game can be developed in infinite ways as well, as long as it is played, as long as new strategies can be introduced and tested. Language–game is the process of inquiry in search of meaning, understanding, discovering new information. But each language–game is fixed in view of its endgame, which is also the goal of any inquiry. An example, again, from chess can be instructive here.

Every (good) chess player comes to the game hoping to implement his or her elaborated strategies. But at the game, before the first move, the possibilities of the game developments are, indeed, literally infinite. When the first moves at the game's opening

stage are made these possibilities are significantly diminished, yet the middle stage puts to the test the opponents strategies and determines the endgame more than the opening stage determines the middle. Toward the endgame there are limited statistical possibilities of the moves, so much so that a decent chess program can easily calculate *all* of the possible remaining moves in that particular game. Now let us implement this analysis to any game that is already completed. It makes no sense to say that the 'Immortal Game' between Adolf Anderssen and Lionel Kieseritzky on 21 June 1851 has infinite ways to develop. There is only one way being only one description of that game now studied by every student of chess. If I say that before Anderssen and Kieseritzky made their first move the game has infinitely many possible ways of development, I say absolutely nothing informative about the game.

Continuum as the modality of games is a basic and obvious fact about the game. If the game is somehow decided between the players (a sad but also illegal act in any sport), we have no right to call it a game. If I know on 20 June 1951 that Anderssen will win due to some agreement between him and Kieseritzky, not only there will be no 'Immortal Game' but there will be no game at all. I might predict that Goliath will necessarily win against David if the two are to fight against each other. My prediction can be based on my knowledge of the physical fitness of the opponents, psychological facts, etc. But what makes any game, language–games as much as military and sport games, so fascinating is the modality of strategies that determines the modality of each game's outcome. A game decided ahead might be even interesting to observe, but it is no game at all, or, in the language–game terminology, it is a completely different language–game altogether (i.e., the language–game of lying[11]).

Fixed scientific experiments are no more scientific experiments than fixed games are games. If I throw the ball standing on the surface of the earth, I cannot consider it to be the test to prove gravity as if the ball might not be falling down. That the ball will fall down is part of the physics and it is part of the grammar, rules, of the language–game. This is why Wittgenstein considered in OC G. E. Moore's sentences as expressing knowledge to be completely misguided.

Thus, continuum is the modality of the language–games. It defines and describes the ways in which language–games are played. Modalities are possibilities of different realizations of virtual identities between rules and strategies within language–games.

In Chapter III strategy was defined in terms of planning, thinking, establishing goals and in this chapter it was associated with free action and conscious choice. It was also thought of as synonymous to the notion of game itself. Now we can associate the notions of strategy and game to that of a mode. Mode is the way strategy is played in a particular game.[12] It is in this way that continuum as a mode of play comes as a fourth dimension of the language–game.

Recalling Peirce's thought experiment of the inkstand and the brain lobe in CP 7.366, we can place these two elements in the continuum of the language–game of thinking and expressing one's thoughts. The brain states (rules) initiate thinking (strategies) are

[11] Wittgenstein (2001: §249b).

[12] However, the way strategy is played should not be confused with practical arrangements to implement a particular strategy, being a tactic.

virtually identical in the entire process of thinking (language–game). These brain states are not just present at the beginning of the continuum (as a mode of the game's development) but present at every stage of the game, hence at every element of the continuum there is the element of the rule, and at every element of the strategy there are necessarily present elements of the rule. Because of the virtual identity between rules and strategies, brain states and thinking process, these two are indistinguishable except in the abstract and somewhat artificial distinctions of rules vs. strategies, brain states vs. thinking.

The way of the language–game development, contrary to Wittgenstein's one–world language–game understanding, is essentially modal.[13] Just as there is an infinite number of strategies from the rules, there are infinite possibilities of the brain states realizations in the process of thinking. This infinite realization of the limited brain states in the infinite strategies of thinking constitute the cognitive continuum of family resemblances and forms of life of thinking.

Above I made a distinction between the infinite possibilities of one game development against language–games being played infinitely, saying that the later is nonsensical. Continuum within a single language–game, however, does not mean that there cannot be continuum of the many language–games nor that virtual identity cannot be implemented between them. Again, here the notion of identity and continuity are not used merely as analogies. For instance, when someone says, "You have your father's nose" it is a statement of virtual identity. It is neither a mathematical nor a logical identity (unless somebody is holding his/her father's nose), nor it is an analogical predication. There is a genetic continuity between that person and his/her father that is actualized in the shape of that person's nose (among other physical features). On the other hand, when someone says, "You have a Roman nose", this is purely an analogical predication. It does not mean that this person is somehow related to the ancient Romans or is of Mediterranean descent. Even if that is the case, the above predication is no more than a metaphor for a particular type of nose.

The concept of Wittgenstein's '*Familienähnlichkeit*' was already introduced in Chapter II in the context of connectionism. There it was suggested that the family–resemblance of the language–game can serve as an alternative to the type–token understanding of the parallel distributed processes in cognition. Different but related language–games share a complex network of related similarities that cannot be explained in terms of mathematical identity of types and tokens or set theory. In PG 66–71 Wittgenstein con-

[13] "As Stenius has pointed out himself in his book on Wittgenstein's *Tractatus* (1964: 157), there is both in the *Tractatus* and among logical empiricists a tendency to consider language as fundamentally descriptive, without any modal element. See also Føllesdal (1967). However, in the PI Wittgenstein did not change his mind about modality. Although the notion does not appear in the text, it is clearly implied that the meaning does not depend on the states of the affairs in one world or another, but in the particular language–game being played. If not, Wittgenstein would have had to embrace the model of '*Gegenstand und Bezeichnung*' meaning. Thus, if the meaning is changed, the language–game is changed as well. See Wittgenstein (2001: §§41–42). When I define continuum as language–game modality I imply that one language–game can instantiate another by implementing a different meaning which is an implementation of a different set of rules that in turn creates a different language–game.

siders one family sharing similar physical features, i.e. one type of a Roman nose, to continue with my analogy. Concepts and people are like language–games (and vice–versa): being distinct they can share similar features that are virtually identical and continuous (the sameness of nose), while at the same time preserve their own distinctiveness.

Whenever we use the notion of language–game as semiotic continuum of cognition we do not use it analogically or metaphorically. Virtual identity cannot be analogical. Similarly, the notion of a language–game cannot be analogical either. For Scotus as for Peirce, virtual identity was univocal predication. The discourse on cognition as semiotic continuum and language–game in this work is also univocal. In this regard I do not depart from Wittgenstein either: language–games are an univocal description of our rule–guided activities. We cannot step outside of our language by means of language. For these reasons, Peirce's virtual identity in his semiotics and Wittgenstein's notion of language–games come together in a perfect fit.

The question asked at the beginning of this work, 'How can we speak univocally of our cognition?' is now answered with the concept of language–game developing according to the four dimensions of cognitive semeiosis. This suggestion, however, imposes restrictions upon neuroscience's ability of explanation. If physicalism as reductionism is refuted then neuroscience has a limited role to play in explaining human cognition. Limited to its discourse on the rules of the language–game, however, its role is absolutely essential and indispensible. Without rules we are not capable to define anything in this suggested philosophy of cognition.

Another restriction is imposed on the strong AI thesis. First, by declaring that all analogical discourse on cognition is senseless. Analogy outside of mathematics has no accuracy. Second, by indicating that computationalism is limited only to the description of rule–following. This is, however, true only of the computationalism and the AI thesis of the present generation. The notion of computation as rule–following does not have to be the definitive understanding of what computation is. Wittgenstein and Peirce were in complete agreement that what divides us from machines is our ability to instantiate strategies out of the rules (for Peirce it simply meant the ability of creativity in thinking). If the second generation of computationalism focuses on strategy–development rather than the rule–following (from actions and states to goals and purposes, from rule–following to decision making), then there is good reason to hope of having a univocal AI thesis in cognitive science.

The way the second generation of computationalism can begin is by taking these two steps. First, by abandoning analogical language. Second, by dropping the exact distinction between syntax and semantics, as well as by eliminating the analogy between brain states being syntax and mental states being semantics of our cognition. Thinking machines rather than computing machinery would be no longer analogous but an univocal statement of the AI thesis.

3. Cognition as a Four–Dimensional Language–Game

There are two kinds of definitions of cognition in contemporary literature. First, the most widespread, is the dictionary use adopted by cognitive psychology. *The Oxford English Dictionary* defines cognition through the use of synonymous terms:

> 1. a. The action of faculty of knowing; knowledge, consciousness; acquaintance with a subject. b. Apprehension, perception. 2. *Philos*. a. The action or faculty of knowing in its widest sense, including sensation, perception, conception, etc., as distinguished from feeling and volition; also, more specifically, the action of cognizing an object in perception proper. b. A product of such an action: a sensation, perception, notion, or higher intuition.

Cognitive psychology is essentially defined by this same distinction between perception and knowledge on the one hand, and emotion and volition on the other.[14] According to this definition, cognition is only part of what 'mind' is supposed to be. But why cannot cognition be extended to volitions and emotions? Well, what is the hallmark of cognition according to the above definition? The definition presumes consciousness to be the essential part of cognition. Why is consciousness excluded from emotion and volition? Because consciousness is still widely understood in terms of awareness, as it is clearly stated in the very same dictionary (omitting the first two entries of the legal senses): "3. The state or fact of being mentally conscious or aware *of* anything. 4. a. *Philos*. The state or faculty of being conscious, as a condition and concomitant of all thought, feeling, and volition". *The Concise Oxford Dictionary* adds: "The state of being conscious (*lost consciousness during the fight*); awareness, perception (*had no consciousness of being ridiculed*); awareness of (*class–consciousness*); the totality of a person's thoughts, feelings, and sensations, or of a class of these (*moral consciousness*)".

The association of consciousness with awareness was already criticized in discussions of the identity theories in Chapter II. The theory would presume us to be unconscious, i.e. every time we fall asleep. The state of awareness is clearly associated with behavior of "the totality of a person's thoughts, feelings, and sensations". When comatose patients clearly manifest no such behavior they are assumed to be unconscious and unaware. These are dangerous assumptions for these patients are conscious and some of them even aware of their surroundings. Their inability to manifest it should not count as lack of consciousness.

The second definition of cognition comes from the works of Maturana and Varela: "Cognition is a biological phenomenon and can only be understood as such; any epistemological insight into the domain of knowledge requires this understanding".[15] For them, all living systems come as cognitive systems: "Living as a process is a process of cognition".[16]

---

[14] Alan Garnham, defining cognitive psychology, begins with the same definition from the very same dictionary. See Garnham (1995).

[15] Maturana and Varela (1972: 7).

[16] Maturana and Varela (1972: 13).

The authors consider (any) living system as a 'self–referring domain of interactions' with a circular nature of its organization.[17] Evolution, reproduction, and self–reproduction of living systems are considered starting from this working definition. Upon this idea is based their basic distinction between *autopoiesis* and *allopoiesis*. *Autopoiesis* is the process of self–production which is essential to all living systems. *Allopoiesis* is the mechanistic process in which the system does not produce itself but it is produced by something else.

This distinction can be translated into what Aristotle called *psuchê*: what distinguishes animated matter from non–animated. *Psuchê* explains this distinction and even today some scientists recur to Aristotelian psychology for general explanation of living systems. It seems that Maturana's and Varela's *autopoiesis* can in some way be compared with Aristotle's *psuchê*'s animating matter.

In the previous chapter it was already mentioned that according to Peirce information interpretation is essential to semeiosis. What Varela's and Maturana's notion of self–organized living (autopoietic) systems is connected to is a dyadic relation. Their interpretation of the 'self' and intrinsically linked to it, their definition of 'cognition' is precisely in this dimension of dyadic relation. On the other hand, language and intentionality cannot be thought of unless the relation is triadic.

Semeiotics, albeit incompletely, can be regarded in dyadic relations of signs as indices. However, such semeiotics cannot be held accountable for language and intentionality. On the other hand, although language–games are not essentially linked to either language or intentionality (unlike Austin's 'performative utterances'[18]), they cannot be thought of otherwise than triadic relations.

If we consider Varela's and Maturana's definition of cognition as applicable to any living organism, we cannot extend the same definition to Peirce's full account of semeiotics and Wittgenstein's notion of the language–games. Also, it cannot account for (nor was it meant to be by the authors) mental events.

Every living organism by itself displays teleonomic functions of information interpretation, control, and adaptation, while biological organisms as wholes also manifest teleologic functions of being goal–oriented. Semeiotics connects teleology and teleonomy and treats external signals (object in general sense) as signs.

This is what Wittgenstein implied when he spoke against mental causation of the brains causing minds in Z §§608–611. In this context he spoke against efficient, not final or teleological causation. However, the difference between biological organisms (brains included) and cognition is that the latter also includes the symbolic interpretation of information.

Symbols and language–games share this important unity of syntax and semantics. Likewise, as was mentioned already several times, within language–games rules and strategies cannot be separated, unless by abstraction: all rules are teleologically formed in view of strategies and no strategy can be even imagined without its constitutive rules. But when it comes to the view of the language–games in their complete understanding as a four–dimensional activity, these distinctions are of no help. Cognition

[17] Maturana and Varela (1972: 10).

[18] Austin (1979: 233–252).

just as language–games, must be regarded according to the suggested notion of continuum which can only be achieved when all other three dimensions are in play. Unlike any biological function, cognition is universal, or "general" as Peirce considered symbols and thirdness.

As we have seen, Peirce defined symbols in his earlier semeiotics as "one which upon being presented to the mind – without any resemblance to its object and without any reference to a previous convention",[19] and in his later accounts he identified symbols with thirdness and interpreted symbols as a triadic relation.[20] Finally, he would refer to icons and indices as "non–symbolic thought–signs".[21]

This is why I insist that semeiotics is conceptually a larger notion (indeed a theory) than language–games, which are unimaginable as monadic or even as dyadic relations. It is no surprise that for Peirce, unlike for Wittgenstein, semeiotics is applied not only to the world of living organisms but even to the physical world.[22]

Language–games are restricted not only to living organisms but to the symbolic interpretation of the information. This, of course, is not limited to humans alone, nor to the symbolic language. Unlike performative utterances in Austin, language–games quite often operate without the use of the ordinary language. This was already explained in relation to the primitive language–games, but the same can be said in terms of the language–games formation in their first dimension, that of the rules and strategies.

I suggest defining cognition within this notion of the language–games as the symbolic activity of information interpretation but, most importantly, information manipulation. Here by information I mean a simple set of rules interpreted either iconically or indexically. Rule following, therefore, is a dyadic relation of information acquisition and interpretation, while information interpretation symbolically always involves triadic relations. This is also the difference between rules and strategies within a particular game. Rules are dyadic relations between information and action according to that information. Strategies are triadic symbolic relations between information, action and the choice of action.

In Chapter III the definition of strategies was that of purposeful actions of implementation of one set of rules over another. It is also important to stress further that this choice of the rules is purposeful itself for the manipulating of the information.

For Peirce, symbols are a triadic relation being the result of generating an interpretant due to the observed connection between sign and its object. Any inference would demand the presence of symbols since any inference is a triadic relation itself. Therefore, cognition is a symbolic, inferential and triadic relation explained through the workings of the language–games in their four–dimensional reality. First, rules generate strategies when a choice is given and made between one rule and another, or between one set of rules and another. Second, simple, primitive language–games give rise to more complex, complete language–games the same way in which we infer from the set

[19] From 1866 in Peirce (1982: 258).

[20] "Notes on the Categories" from 1885 in Peirce (1993: 235–241).

[21] From 1909 in Peirce (1935: 338).

[22] From 1903 in Peirce (1998: 193).

of primitive language–games one complete language–game. This occurs when from meaning–neutral facts we construct meaning–relevant (but also knowledge relevant) events. This was already explained in detail in Chapter III in relation to the non–epistemic vs. epistemic perceptions. Third, out of this continuity between primitive and complete language–games we form, what Wittgenstein called, forms of life and family–resembling language–games. Thus, consciousness, being neither awareness nor restricted to the human self–consciousness, is relational activity. It can be defined as the relations between different language–games families. Fourth, modality of the language–games is something exclusively human and it is related to the subjective self–consciousness not present in other animals.

This distinction between consciousness and awareness, and between self–consciousness and self–awareness is important in defining cognition that is inclusive of both, perception, thinking, and emotion, volition. This was already explained in rejecting the cognitivist definition of cognition.

In conclusion, cognition is the ability of instantiating strategies out of the available rules or, in other words, the ability of creating language–games. This instantiating of the strategies is the choice between two or more alternative rules. When the person (the player) is given the choice between two or more rules to follow, the mere fact or the mere ability to make even the simplest choice is the sign of the presence of cognition.

This can be demonstrated by imagining that somebody is given questions and multiple answers. Presuming that (1) the person answering questions understands them, (2) the person also understands that only one answer is correct, and (3) is capable of providing the answer to the question, we can assume that this person by choosing among the available possibilities is manifesting cognitive ability. In the case of a computer such ability is missing, viz. such ability is present *outside* of the program completely dependent on the programmer's command.

In the case of the present cognition such ability of choosing between different rules to follow is where the strategy comes about and this ability itself is cognition. The choice of answering at all is to follow one rule, the choice of not answering means to follow another rule, but whichever rule one chooses one instantiates a strategy. Strategies are also given in the multiple answers, but before one chooses any strategy, one must first choose to answer (or not to answer) at all. That primary or even primitive ability to chose to begin the game is cognition.

4. Virtual Identity

4.1 Identity of Rules and Strategies

For Wittgenstein the notion of identity was essentially one, but it could be expressed either as a tautology or as a contradiction. Peirce, on the other hand, had two notions of identity. One was mathematical and logical. Unlike Wittgenstein, Peirce never made any clear distinction between the two and his logical identity was carried out from mathematical equality. The other identity was already discussed in the previous chapter and it was virtual identity.

Virtual identity, borrowed from John Duns Scotus, presented for Peirce a very important methodological tool in dismissing dualism and monism alike. The 'brain lobe' was *virtualiter* present in the mental act of thinking and thought expression, just like the sun is *virtualiter* present on the earth, being present only in its efficiency and not actually being *on* the earth.[23] By 'present' here is meant 'identical to' in the act of that to which *a* is identical to *b*. Thus, the brain state that initiates thinking is present to the act of thinking and further, to the act of thought expression by being present to these acts in its every stage and not just serving as an efficient cause only at the beginning.

Recalling my earlier distinction between the analogy of saying that one has a Roman nose and the identity of the same type of nose in one family, let us show now through the distinction of the different senses of identity suggested in the Introduction to Chapter II how virtual identity is different from all others. This time I will elaborate Peirce's example of the sun *virtualiter* present on earth.

1. Tautology, being the strongest identity, tells us nothing about the sun or the earth. It cannot be used meaningfully to express any identity apart of the self–identity in mathematics.

2. Univocal identity is the virtual identity. This identity, according to Peirce, is an identity of efficiency. Thus, the sunlight and sunrays always present on earth are not something separate from the sun but are "actual efficiency" of the sun. When we say that the sun is *virtualiter* on earth we do not mean that the sun is actually present on the earth nor vice versa. There is no sun, being an object, physically placed or attached to the earth, being another object. However, by univocal predication (i.e. "lots of sun today") we can speak of the sun being on the earth.

3. Equivocal identity presumes three elements of predication. When we say that the sun is a star and the sun is the king, in the first instance we speak of the center of the solar system, while in the second we refer to Louis XIV, *'le Roi Soleil'*.

4. Proportion identity makes use of the same three elements of predication except in reverse. Saying 'shining sun' and 'shining example' uses the same predication about completely different nouns.

5. Analogy proper, or identity of proportionality, uses four elements of predication and can be demonstrated through the use of the 'sun' as an analogy, model or paradigm of something. I.e. the Standard Solar Model is a mathematical theory of treating the sun as a sphere composed of gases implemented for the study of other suns in the universe, especially in the study of their evolutions, that cannot be observed either directly or as well as 'our' sun.[24] Although a genuine mathematical theory (model), the identity here cannot be treated univocally due to the significant differences in chemical compositions and other physical features between the suns. Unlike in univocal identity, analogy does not treat or speak of the same object but of two or more.

6. Metaphors also use four–element predication. I.e. when someone says "You are the sunlight of my life", the analogy is between the living force of the sun on earth and the relationship between the lovers. Notice, however, that the difference between analogy and metaphor is very important in predication and should not, as they often are,

[23] Peirce (1935: 372).

[24] Bahcall (2003).

be considered to be the same. We say that there is an analogy *of* something, but we also say that metaphor is *for* something.

Returning to the identity of rules and strategies, the identity is a virtual presence of the rules to strategies and vice–versa. This implies that rules are virtually present to the strategies (by defining them) at every stage of the game, and that strategies are present to the rules (by being their goal of realization) also at every stage of the game. Just as we speak of the same sun located about 150 million kilometers from the earth and the sun shining through my window, we have the very same strategies in that one particular game. I.e. the Sicilian Defense, being one of the most common chess opening strategies, can be easily transcribed as the first move of the chess game: "1. e4 c5". What we see here is a description of a simple rule, but this is also the Sicilian Defense strategy, being a chess opening strategy. It is in this way that we say that rules are virtually identical to the strategies without being the very same thing. Sunlight is virtually identical to the sun without sun being physically present as an object on the earth.

When we take a single chess strategy we can transcribe it in a series of moves of the chess pieces made by the players. When seeing just that list, the reader can have a perfect understanding from that notation not only the rules implemented in that strategy, but the whole strategy as well. In the same way, when a well–trained musician reads the musical score without actually playing it, he or she hears the musical melody as if it was played. But no serious chess player would imply that any chess strategy is reducible to the list of the chess moves. Strategy presupposes long studies, planning (strategy was defined as planning in Chapter III) and many tests in actual games. The chess notation of any strategy or the entire game is the end product of these long processes. In the same way, no musician will ever imply that music can be reduced to its notation. Beethoven's symphonies are very different when they are performed by Daniel Barenboim (2012), Leonard Bernstein (1970s), or Herbert von Karajan (1977). I presume, however, that they were using the same score written by Beethoven. One sees music, one hears music based on the very same notation; one sees the rules, one for–sees the strategies based on the very same set of rules.

This notion of virtual identity is easier to express with the notion of cognition rather than the mind. The ability of making decisions out of the available rules implies the virtual identity of the rules in strategies. Again, rules are as they are in view of the strategies while strategies without rules are unthinkable. Imagine, i.e. any chess strategy: we can deconstruct every single strategy into several rules. Yet, knowledge of all defining rules would give us no knowledge of even one single strategy.

Peirce's notion of efficiency connected with his concept of virtual identity fits perfectly within this rules–strategy dichotomy. Rules make possible all the strategies of the game. "Queen captures knight" is a short description of a simple strategy (always within some larger strategy being played in a single game), that implies the virtual identity (presence) of the underlining rule of one piece capturing another, and a description of how this rule is being played (i.e., a4–c6x). Notice that "queen can capture knight under certain circumstances" is not a strategy but rule description.

The virtual identity is strictly connected with cognition and the notion of a language–game also because whenever there is simple rule–following without strategies there is neither cognition nor virtual identity. If the rules are alone, their own identity

cannot be described by virtual identity that presupposes two elements. This is why even in a computer program we cannot speak of rules without strategies. All rules are made in view of strategies and the fact that strategies are the work and the invention of the programmer outside the program does not change the fact that every rule demands strategy.

The brain states are virtually identical to the cognitive processes they instantiate. No cognitive process can be without them, but brain states alone cannot be held accountable for cognition. The role the brain state plays is that of instantiation on the level of primitive language–games strategies. Whenever the complexity of the brain states reaches a point of choice between being replayed again or taking a different set of rules (different brain state), cognition is instantiated by the instantiation of a strategy in this primitive game. Complete language–games are formed, more games constitute the entire families of these games, and continuum supplies more choices and possibilities for strategies to be played, for rules to be implemented. Rule following is related to action performance. Purposes and goals, on the other hand, come with strategies. The complexity of rules must be oriented toward purpose achievement and goal meeting in order for rules to generate strategy.

Chapter II suggested viewing the functionalist theory of multiple realizability in reverse. Instead of speaking of one mental state potentially realizable in many physical states (for computational functionalism not necessarily in the brain states), I suggested to consider the fact that the same set of the brain states can give rise to many mental states. Virtual identity confirms this suggestion and the notion of the same sets of rules in a particular language–game being able to generate many different strategies implies univocal interpretation of the brain states as rules and mental states as strategies.

In Chapter III the narcissistic character of the sensory receptors transmitting relevant but unrelated information was mentioned. I indicated that we cannot speak of the isomorphic correspondence between the world and the transmitted information and that further interpretations are needed and are provided by the different brain areas.

There is no isomorphic representation of the world in the transmitted information, being the set of often unrelated rules, but it has the need of further interpretation, being strategies in charge of integration and interpretation. Isomorphism, nonetheless, comes only on the more general level of the interrelated families of the language–games. Here, however, one cannot speak of a mathematical equality of the rules and strategies, nor of a discontinuity between the two. This is why there is a need of redefining both the notions of identity and the concept of continuity in which neither logic nor mathematics are of significant help *anymore*.

The same goes for the notion of efficient causation. One advantage of applying language–games to cognition instead of applying it to mind is that the problem of efficient causation can be stated in terms of non–temporal causation. It makes sense to say that there are brain states at time t_1 causing mental states at time t_2, but it makes no sense to say the same about one strategy being caused *after* such and such a rule was given. Although "minds are caused by brains" awaits a detailed explanation, it is grammatically and semantically a sound statement. On the contrary, "strategies caused by rules" is an unintelligible statement if by that we mean that there is some time at which there were rules without strategies.

Why do I insist on claiming that when virtual identity is applied to the mind/body relation this is not an analogy or equivocation? Notice that in the above list of the six senses of identity only tautology and univocal identity use the same physical object to identify. Leaving tautology as irrelevant, the only truly informative theory of identity that predicates only of one and the same object is the univocal one. This is why the choice in the philosophy of mind was always between eliminativism and property dualism, while the desired identity should have been univocal. It makes no sense to speak of the mind being an emergent property of the brain, just as it makes no sense to speak of the strategies being emergent properties of the rules, or sunlight being an emergent property of the sun. This is why there must be some unified term to indicate not the two properties of the same object and certainly not the two realities (even if they operate by the same set of laws, unlike in substance dualism), but the same object, the brain, being in its complexity of functions present at every stage of the cognitive processes of perceiving, knowing, feeling, speaking, etc. in the one and the same subject.

In conclusion, there is no analogy between mind/body and strategies/rules. First, because instead of the term mind I suggest the concept of cognition. Second, because cognition is not treated as separate reality or property (property of what, of itself?) but as a continuum of the brain states. Third, because cognition is not only limited to the rules composing strategies (this is only, as I said, the beginning of cognition) but to all four dimensions of the language–games. Fourth, because any separation of the 'physical' and 'mental' is only artificial and very abstract.

4.2 Identity of Syntax and Semantics

One of the side effects of substance and property dualism is in the sharp distinction between syntax and semantics. Hintikka and Pietarinen indicate that such a distinction did not come about until the 1930s in the works of R. Carnap and Charles W. Morris. However, even in the earlier works of Carnap such distinction was not clear.[25]

Sharp distinctions between rules and strategies are abstractive. In reality rules are made for the sake of strategies and strategies are composed out of the rules. This relationship is not circular but spiral in the context of modality brought by continuum. Rules are less prone to revisions than strategies if we consider one single game being played constantly. It is highly unlikely that chess rules can be amended even though the game presupposes that an infinite number of strategies are to be played. In language–games, which are constantly created, rules and strategies are made up all the time.

Who is creating the rules and strategies in the language–games of our cognition? Are we not back to the Cartesian *ego*, this time in charge of the game? In the previous sections, defining cognition as an ability of creating strategies out of the defining rules, it was mentioned that what makes the threshold between computer and human is the complexity of rules that makes one think and decide. This threshold must be explained

[25] Hintikka (2003: 13); Pietarinen (2006: 377).

in terms of goals and the purposes of the rules–becoming–strategies process in view of the purpose of problem solution.

The notion of the rule or syntax used in some context is not always related to the strategies in which the rule is used. By rule one can mean a rock falling down from the hill, and the strategy of a primitive game can be the action of avoiding to be hit by it. The same rock can be used in a house construction; in this case the rule is employed for the sake of the strategy of house building. The entire physical world can be seen in terms of the out–there present rules that can be part of somebody's language–game.

What makes an object a rule is the interpretation of it as a sign in the process of inquiry: "Every sign *by itself* is dead, but *what* gives it life? – It *lives* in use. Was the living breath in it? Or is the *use* its breath?"[26] To explain the use Wittgenstein recures to his notion of the language–games, but the notion itself, as we had seen, is not explained theoretically. In the text of the *Investigations* we are constantly asked to be part of the authors language–games instead. The explanation of the use of signs is given much thoroughtfully by Peirce especially in his mature semiotics. It is an essentially a non–reductive triadic semeiosis.

The identity of rules and strategies cannot be considered otherwise but as a virtual one. For Wittgenstein we should not speak of identity at all but learn that rules and strategies come naturally together when the game is played. For Peirce we cannot speak of localization of one or another. In applying rules we do not just follow them but construct new strategies.

This is why virtual identity cannot be treated as mathematical equality of logical identity. It is neither tautology nor contradiction. For being either of these, each element would have to be localized within the game itself. One day neuroscience will most likely achieve such excellency as to be able not only to point at the brain state from observing the action, but from the observation of the brain state be able with precision determine to what action it corresponds.[27]

In the language of rules and strategies, the fact that one single strategy can be perfectly transcribed by its constitutive rules does not indicate that it can be reduced to these rules. Complexity is not the only problem here. The ability to use the rules does not depend on the presence of these rules alone. Their use depends also on the goals for which they are chosen and on the intermediate link between rules and goals, viz. the strategies, made of rules for the sake of the goals and purpose achievements.

4.3 Identity of Facts and Events

Therefore, no mathematical or logical identity can be applied in explaining human cognition in general. They all demand localization and proportion/proportionality of analogy between one and another. Chapter II had demonstrated why this remains no more than analogy. On the contrary, virtual identity is univocal identity.

---

[26] Wittgenstein (2001: §432).

[27] Schauer (2010).

Virtual identity defies the numerical identities of Frege's sense and reference, and Peirce's type and tokens. Statements of equality 'rules are (=) strategies' and 'strategies are (=) rules' are false because even the transcription of strategies through the list of its constitutive rules used in the game does not explain anything except the procedure of the played strategy and strategy cannot be reduced to its constitutive rules as their simple arrangement.

The material equivalence can be applied to 'there are strategies if and only if (≡, or ↔) there are rules', but the reverse is false, since rules can exist on their own, i.e. in a program. It was explained before that presence of *any* rule implies some strategy, even in a computer program. However, that strategy is not in the program itself but outside. It is possible to implement this strategy within the program but only in a form of another rule, i.e. as the power set of the sets of rules for which it is a strategy. The same is true in relation to material implication (⊃, or →): while all strategies imply rules, only in non–pseudo language–games do rules imply strategies.

Finally, rules can be members of strategies but the reverse would only be true if we could reduce strategies to their rules. The material implication by the set's theory relation of membership within a set (∈), however, implies that membership of the members is localized. While we can perfectly well localize every rule in one or many strategies, only a reductive approach to strategies would allow for treating all the elements in a strategy as localizable. The reductive approach would have to get rid of the goals and decision making which are essential features of every strategy ("planning and implementation"), and these two are not expressible in the list of the rules.

One might object that at least the goal of the game can be sometimes described purely in terms of the rules. For instance, 'checkmating the king' is a description of the rule as well as statement of the game's purpose. This is true, however, of the single game's goal. If we look on the many dimensions of the language–games we will see that one game can be part of the larger games, a primitive, part of the complete ones, and that sets of games can fit within even larger family–resembled games. In this context, any hope of describing these complex goals in terms of the strategies' constitutive rules alone is unreasonable.

Decision making, on the other, hand, is a choice between different rules that by itself constitutes a strategy. On some primitive level, when the decision is between one set of rules and another, it can be described by the rules alone and in this case we can speak of localized memberships of the elements in one or another set. However, just as in the case of the goals, it is impossible to describe decisions in terms of choices simply between two sets of rules when we decide between strategies with goals. In other words, the greater the complexity of the game, the harder it is to refer to the rules of the game alone.

Virtual identity is univocal identity but it cannot be transcribed by logical or mathematical equivalence. This would imply the reduction of goals and decision–making within a single and especially within a larger game. For this reason, rules are localizable only within single strategies but rules cannot be localized within the games the same way we localize elements as members within sets and collections. In this regard Peirce's criticisms of Cantor's set theory in which sets are collections of their own ele-

ments is crucial. Rules define strategies but rules do not constitute language–games the way strategies do.

In the second proposition in the TLP (1.1) Wittgenstein affirms that the world is a totality of facts, not of things. For him 'things' are not only individual objects but also predicates. Both were considered as atomic states of affairs and as logically atomic things in Wittgenstein's logical atomism.[28] These things and predicates constitute his phenomenological objects which were given up in 1929, but the idea of the world being the totality of facts persisted.

Most importantly, as we have seen in Chapter III, our knowledge of the world is our knowledge of the facts. These facts in the *Investigations* are primitive language–games as constitutive parts of the complete language–games. Unlike things and predicates they can be individualized as events within a particular language–game and further be transformed from primitive to a complete one. Thus, the fact of one's facial expression is part of the event of pain. We identify one's facial expression with the pain while the observation of the objects of perception brought to our consciousness and named by our predication is a much later act of language and logic. Likewise, we count objects and assign them numbers only after and never prior to our experience with the fact that the objects are perceived and individualized.

Our knowledge of the world is based on the identity of facts, not of things. We come to know facts before we know things just as our knowledge of the world is prior to language and logic, viz. prior to predication about things. Virtual identity is implemented to that of facts, not of things. For this reason, since language and logic apply mathematical identity, virtual identity is not applicable in them, nor can the logical or mathematical operators transcribe it.

Since cognition is the ability of transforming rules into strategies at every dimension of the language–game, virtual identity has a very important but limited role to play in cognition. The limit consists in the constraint of the relation between rules and strategies, or in the non–reductive nature of strategies into their rules. The importance of virtual identity is that unlike mathematical equality it is not limited to single–word semantics but it is realized in the modality of continuum.

5. The Mind/Body Problem Dis–Solution

The purpose of Chapter I was to show that substance dualism is unable to answer the four major questions of the philosophical psychology. The life–giving force of *psuchê* in *phusis* was the common explanation also for all the cognitive states. However, already in Aristotle the unsolved problem in explaining how the *nôus* as the highest faculty of the *psuchê* is united with the body, in the Middle Ages again took shape of dualism. In Descartes, once the *anima* was dismissed as the life–giving force in favor of a mechanic explanation and efficient causation of living matter, dualism took its strongest shape and the mind/body problem was defined as the problem of causation and interaction.

---

[28] Stenius (1964: 65).

Since substance dualism was not able to provide a satisfactory explanation, the next obvious solution would be to dismiss it in favor of physicalistic monism. The purpose of Chapter II was to show that physicalism is not able to meet the challenge either. It furthermore resulted in that the physicalistic methodology in the philosophy of mind rendered the five questions meaningless.

The formulation of the mind/body problem presupposes an essentially dualistic ontology and metaphysics. To say that we have mind implies that we have body as well, and vice versa. Whether we treat them as two substances or properties does not change the fact that we are using generally dualistic imagery and language. Property dualism was meant in many ways to be, if not the final, at least the temporary substitution of substance dualism until some better alternative would appear. No doubt, the mind/body problem can be as easily formulated in property dualism as it was in substance dualism.

However, property dualism is not nearer to the solution of the problem than substance dualism was in the past. Saying that mind is an emergent property and/or a function of the brain only reiterates the question of causation. In terms of the treatment of the interaction problem between mind and body, property dualism by speaking of 'emergent properties' remains on the level of a very general description. At most it answers the question 'what' but not 'how'.

More radical versions of physicalism render the mind/body problem completely meaningless. The folk psychology use of propositional dualism was never seen as any tangible solution to the problem. It serves more as a necessary terminological tool in face of the linguistic destitution of the philosophy of mind.

Paradoxically, if one really wants to solve the mind/body problem, one must embrace some form of dualism in which the question was formulated. For substance dualism the solution of the mind/body problem presupposes solving the interaction problem first. Any solution of the interaction problem from the point of view of substance dualism would necessarily suspend our neuroscientific judgments. In other words, either neuroscience or substance dualism are inconsistent. All mutual consistency is excluded by two neuroscientific facts: by distribution of the brain functions in different parts of the brain, and consequently by the implication that there is no one and only center of consciousness.

Some might argue for a different definition of substance altogether. Whitehead and Hartshorne insisted on defining substance in terms of process, and being in terms of becoming. This would bring dynamism to the notions of being and substance. Process metaphysics brings more clarity to the substance of the mind than to the substance of the brain however. If process is defined in terms of evolution, process metaphysics would have to explain somehow the identity of different stages in the evolutionary process.[29]

If being is becoming then process metaphysics would need to offer an alternative to materialist explanations (such as survival and adaptation) as to why there is anything becoming or being at all. Finally, the social characteristic of the mind insists on the communities of subjects, but any subject (according to Peirce) is a semeiotic interpreta-

[29] For in–depth criticisms of process metaphysics see Strawson (1964).

tion of the object. Thus, any talk of subject presumes a great deal of the study of objects first. Does not this discourse bring us back to substance dualism? If we continue to resist it, are we not offered some strong version of idealism instead?

The two authors discussed in the previous two chapters agree on their dismissal of dualism and monism alike. In terms of their answers to the five questions, however, there are some significant differences.

Wittgenstein dismisses the problem of identity either as the question of mathematics, or logic, or philosophy. Whenever any of these three attempts to state the question, the answers turn out to be either nonsense (contradiction) or tautology (saying the same by different terms). The ontological question of existence and the epistemic question of certainty for Wittgenstein must be answered through the activity of the player in the language–games. Outside the language–games these questions are meaningless. Likewise, semantics is reduced not to some theory of meaning but to the very same meaning as use in the language–games.

For Peirce, the metaphysics of the mind is thought interpreting thought. The ontological question is answered by stating that the reality of the mind consists in its being a system of linguistic signs. Just as his ontology is part of metaphysics, his epistemology is indistinguishable from logic and results in the methods of inquiry as the mind's process of acquisition of information and its interpretation. Peirce's epistemology is that of the cognitive semeiotic continuum. The question of identity between mind and body is answered in terms of virtual identity, where no priority is given to either, no separation but at the same time no reduction of one into another.

As was already stated, neither Wittgenstein's nor Peirce's solutions are satisfactory nor final. Language–games resist the theoretical approach to the questions of the philosophy of mind or any other philosophical discipline. They show more than they tell. Wittgenstein's resistance to dualism and monism did not prevent him from the overall use of dualistic language in his own philosophy of mind.

Peirce's accounts of semeiotics and continuum were certainly philosophy of mind oriented. Although his use of the term 'cognition' was often synonymous with the term mind, his own resistance to dualism and monism moved him to prefer 'cognition' over 'mind/body'. However, semeiotics alone is not enough to reform the philosophy of mind, and his theory of continuum was not final without a non–circular definition of thereof. Finally, the notion of virtual identity and continuum must come together for greater power in the explanation of cognition.

The mind/body problem can only be solved in the theoretical and linguistic context of dualism, in which it was defined in the first place. Any monistic attempt to solve the problem would have to compromise on monism in its most general definition. Property dualism represents such a compromise. Outside of dualism, the problem is a conceptual mistake. Inside dualism, it is a theoretical conundrum. The true solution to the mind/body problem is a resolute refutation of it as a genuine problem altogether. We need a different term without exclusivity for either mental or physical, but implicitly for both. The suggested term is 'cognition'.

CONCLUSION

I began this work by examining the evolution of the terms *psuchê* and *nôus,* into soul and mind. I indicated that changes in philosophical conceptions were indicators of more significant changes in the ontology, metaphysics, epistemology, semantics and logic of the major figures in the history of the philosophical psychology and philosophy of mind. Chapters III and IV presented views of Wittgenstein and Peirce that often stood aside of the mainstream philosophical evolutionary process with their criticisms of dualism and reductionism. Finally, in the last chapter, based on Wittgenstein's and Peirce's philosophies, I suggested a different terminology to avoid the usual choices between dualism and reductionism.

But 'cognition,' being such a term that is not new but newly defined, is itself a product of evolution in philosophy. Defined as multidimensional rule–guided activity (language–games), it brings together, separated by the cognitivists, thought and volition. This term 'cognition' rejects both mental (efficient) causation and multiple–realization of the mental in the physical in favor of a teleonomic and teleological inferential process coupled with the idea that limited physical states are multiply–realized in many mental states. It postulates unity between syntax and semantics on the basis of the inseparability of rules and strategies within one and many language–games.

I suggested the term 'cognition' as an alternative to the dualistic terminology of mind and body. The mind/body dichotomy throughout history employed not only many senses but many references as well. The dual term itself was the indicator of many different senses of identity between the mind and the body, as Chapter II has shown. However, suggested changes in philosophical terminology are no more efficient for the ordinary usage of the mind/body language than heliocentrism for the use 'sunrise/sunset'. This work has no ambition of challenging the use of phrases such as 'keep in mind', 'to set one's mind on something', and 'dedicating one's soul and heart'. We have seen that analogies and metaphors can be unhelpful in defining the mind/body relationship but this failure has no effect on the analogical and metaphorical uses in science (analogy originated from ancient mathematics long before Aristotle[1]), rhetoric, literature and generally in the ordinary language. This conclusion will review the work that has been done to improve the philosophical language in the philosophy of mind and reevaluate the proposed solutions.

---

[1] "Gr. *ἀναλογία* equality of ratios, proportion (orig. a term of mathematics, but already with transf. sense in Plato)". (OED).

Chapter I demonstrated the failure in the discourse of the mind/body relationship due to the overall dualistic conception of human nature. Aristotelian hylomorphism implies such a dualistic view; naturalizing Aristotelianism does injustice to his ontology and metaphysics by defining hylomorphism as a thesis of supervenience or functionalism. The Kantian problem of language was solved by Wittgenstein's transition from phenomenological to physicalistic language and by making language–games as the main representative link between language and reality.

One terminological clarification is in order in view of my definition of cognition. Kant uses the term *Erkenntnis* throughout his first *Critique*, and it is usually translated as 'cognition' or 'practical cognition'. The Kantian notion of *Erkenntnis* is much broader than what is meant here by cognition. For Kant, practical cognition includes our knowledge of God and morality.[2] From the perspective of cognition as an ability to initiate strategies from the rules of the language–games, there must be a lot of theoretical elaboration done in order to reach to the intelligible language of God and morality. My intent in this work was to elaborate only the foundations for many further philosophical discourses.[3]

Chapter II covered a variety of different theories in the philosophy of mind. All these theories fall under physicalism and can be divided, in general, between reductive and non–reductive physicalism. It was demonstrated that any explanation on the purely physical level fails, if by physical we mean what we know so far about the brain states. There is always hope that greater advances in science will grant us a deeper understanding of what the physical is (*ignoramus sed non ignorabimus*).

The notion of cognition places our current understanding of the physical in a different perspective than the one given to us by physicalism. 'Physical' cannot be reduced to our image in observing the operations of the neurons, networks, brain areas, and brain operations alone. While what is observed is essential to our understanding of the cognitive activities, it is only the basis of what cognition really is.

The goal of this work was to find a non–metaphorical and non–analogical language for our discourse in the philosophy of mind. It was found in the philosophy of Wittgenstein, especially in his later period from the 1930s on. Unlike the mathematical (formalist) game theory, the language of the language–games is univocal and physicalistic in its description.

My interpretation of Wittgenstein's texts on AI, the alleged private language for private experiences, on language and mind in general, is very radical. I argued that by the mid–1940s Wittgenstein completely got rid of the phenomenological obectry, from

---

[2] See W. S. Pluhar's commentary in Kant (1996: 5 n. 6).

[3] Without attempting to start new discussion on how the notion of cognition would substitute for Christian concept of the soul, we can, however, recall Locke's attempt to explain the resurrection of the body through his theory of personal identity. Cognition could be thought to survive death if the reality of the resurrection of the new body is admitted. The old rules of the *sōma phusikón* (physical body) are substituted by the new rules of the resurrected *sōma pneumatikón* (spiritual body), with the old and the new strategies in play. This argument, of course, needs a philosophico–theological elaboration that significantly extends beyond the purposes of the current work. Cf. 1 Corinthians 15: 44. On the problem of the resurrection of the body for Platonism, see Nussbaum and Putnam (1995).

his philosophy such as sense–data and his *Tractarian* phenomenological objects. There is plenty of textual evidence for this argument with very little support for it in the secondary literature. This refutation of phenomenology, however, was in no way the impulse for Wittgenstein to embrace some form of eliminative physicalism. Wittgenstein, just like Peirce, consistently avoided any form of reductionism.

I made some extensions beyond what Wittgenstein wrote. I implemented the notion of infinity and modality to the language–games by stressing that all strategies are infinite in number and open to different possible developments. It must be stressed that infinity and modality were not part of his philosophy.

Chapter IV presented two sets of evolutionary development in Peirce's semiotic and continuity. This chapter was meant as an addition to the notion of the language–games, as an elaboration of how language–games are about interpretation of signs. Rejection of any notion of identity by Wittgenstein was answered by bringing in the notion of virtual identity by Peirce and by elaborating it further.

The last chapter answers the question of 'How ought we do philosophy of mind?' by saying that instead we ought to do philosophy of cognition. The mind/body problem was defined by dualistic philosophy and can be answered only by dualistic philosophy or eliminated as 'the body' problem by reductive physicalism. I suggested a different kind of elimination: the way it is formulated, the question is incorrect. There is no distinction of the mind and body, there is no mind being some supervenient property of the body or some of its functions. What was eliminated is *the choice* between dualistic and reductive languages.

In the context of a discussion on the importance of the grammatical investigation of our philosophical language, Wittgenstein remarked that philosophical terms can become idolized abstractions that would mean more than they say. The only solution is the Socratic one of giving rid of the idols without creating any new ones. Perhaps alluding to F. Nietzsche's *Götzen–Dämmerung* Wittgenstein concludes: «All that philosophy can do is to destroy idols. And that means not creating new ones – such as in the "absence of idols"».[4]

If I postulated the terminological absence of mind and body, does it not mean that the new term of 'cognition' is proposed instead as the new idol? If the evolution of terminology from *psuchê* and *nôus* into soul and mind brings us gradually to the newly suggested redefined term of cognition, does this evolution not mean rather a genealogy of idols, viz. of creating new idols in the absence of the old ones?

Two reasons could be given to justify myself from the possible charge of philosophical–linguistic idolatry. First, idols are meant to be transcendental beings representing the unseen, over– and beyond–the–physical reality. While idols exist, what they represent does not. In the context of the philosophy of mind (but philosophy in general as well), it means that a three–partite *psuchê* explaining life, intellect and perception in fact serves as a general term whose purpose is to explain more than it states. When what it is supposed to explain is explained not in that term (i.e. the explanation of living systems does not require the notion of the soul), we realize at the

[4] „Alles, was die Philosophie tun kann ist, Götzen zerstören. Und das heißt, keinen neuen – etwa in der „Abwesenheit eines Götzen" – zu schaffen". In Wittgenstein (1993: 170–171).

same time that the term is no more accurate, true, or even needed. At least here eliminativists are correct: soul is no more a scientific term in biology than phlogiston in thermodynamics.

The term of cognition is a unifying term for the multidimensional language–games. It does not state more than that. It uses univocal identity of terms and, therefore, a direct description by rules and strategies. It lacks transcendentality of terminological description of meaning more than it says. It suggests simple description by univocal language.

If the above justification from the possible charge of saying the same only in different terms (viz. the terms of cognition and language–games being just another analogy) seems insufficient, consider the following. When the claim is made in Chapter V that the brain states in relation to the somehow corresponding to them mental state are univocal to the rules of the language–game, what is it in the brain state that is qualitatively more or less in relation to the rules, or what in any mental state is qualitatively greater or lesser than in a given strategy in language–game? The answer is 'nothing', because here we can have only difference of degree, not of kind. On the other hand, analogical and metaphorical descriptions are not only stating differences of degree but also of kind. Computational functionalism, for instance, compares the brain states to the functions of computing machine implying an analogy which, obviously, is playing not only on the difference of degree but also of kind. Machines, after all, are quantitatively and qualitatively different from brains. Univocal predication as difference in description of rules allows only quantitative difference.

Description of rules related to strategies does not specify what that qualitative difference is. It will be further specified by the particular language–game where the focus will be on the way strategies develop further.[5] Brain states are univocally rules of language–games instantiating multiple strategies which, in the old philosophy of mind, we called them as mental states. But unlike 'mind/body', 'rules/strategies' avoid dualism in language as well as in ontology: rules without strategies are meaningless (pseudo language–games) and strategies without rules are impossible (unimaginable).

Thus, on its most basic level of rules related to strategies, brain states to mental states, the difference can only be quantitative. What accounts for the qualitative characteristic of cognition is its fourth and final dimension, continuum understood in adjectival and adverbial terms, being modality of the language–games. It is here that the qualitative characteristic of our experiences are accounted for. It is also for this reason that the non–temporal notion of cognition, independent from time as much as from mathematical calculation was suggested. As was explained in Chapter IV, my aim was to fulfill Peirce's unrealized goal of defining continuum in a non–mathematical and non–temporal manner.

The second justification, following from the first, is that idols can be counted. In philosophy of mind it translates into simple correspondence between term and its designation. Wittgenstein combated such an approach to meaning from the *Tractarian* isomorphism to the language–games of the *Investigations*. Following Peirce's later

[5] Thus, our knowledge of bat's neurology would not give us knowledge into what is it like to be a bat.

notion of continuum, the redefined notion of cognition resists any mathematical treatment. The notion of continuum *is* mathematical through and through, but the notion of semeiotic continuum is not. It avoids individualization of its elements thus preventing it from accepting efficient causation. In the functionalist accounts of the multiple realizations of the physical states for one single mental state (as if mental state can be individualized), a functionalist would be expected to give some sort of mathematical account for these physical states. Instead, what is individualized are the rules of the given games and some primitive strategies. The complexity of the brain states, however, allows only for very general descriptions which even by themselves do not explain which single mental state it is supposed to explain or generally represent.

Cognition cannot be localized in any specific physical state; it resists such descriptions, it resists counting. It is not limited to the set of rules or strategies but to the unity of the language–games. The language of syntax or rules is mathematical, but such language cannot be extended to the language of cognition. Rules require organization, decision, and finality given by the strategies. As demonstrated in the last chapter in the context of virtual identity, mathematical description does not extend to strategies.

I defined cognition as a language–game. What is, then, language–game to cognition? As a definition it itself implies some sort of identity. That it is not a tautology (repetition of the same through the use of different words) I think is clear by now. It is not an analogy or a metaphor either. To be an analogy or metaphor it would have to presuppose that both, cognition and language–games, enjoy some separate existence, like day and night to youth and old age, etc. The same would go for paradigms and models.

Cognition is understood here as a set of four dimensions, the four dimensions of the language–games. Here the word 'dimension' must be understood as a geometrical term, neutral in itself, not saying anything specific, not settling any qualitative notions of the length, width or depth (nor time as past, present or future). But as in geometry any initial discourse about dimensions must soon be applied to the actual geometrical shapes and figures, so the dimensions of cognitions should have actual applications in the cases of cognition: perception, thinking, feeling, emotions, language, and so on.

Even in my interpretation of the language–game, it continues to remain a relatively non–theoretical term or notion. Something self–contradictory seems to emerge from this: language–games are non–conceptual concepts, a non–theoretical theory. This is not just some paradox that can be solved (as any logical paradoxes) by moving to the next level of discourse. This is nonsense, a *Tractarian Unsinn*, that same ladder we are to throw away once we have climbed by means of it to the desired goal of understanding. Once we have understood what it tells us, it becomes useless, nonsensical in itself.

What are we to understand from the use of the language–game notion in relation to cognition? How cognition works. It works within these four dimensions; it operates by and through them. Language–game as notion is rather neutral even if we specify it in details through its four dimensions. What are these four dimensions we can only *name*, but each dimension will have to *show* itself in the use of the continuous and virtually

identical signs within the cognitive process. What is essential here is not even how we call each dimension but how each dimension develops: it develops through the many uses within the infinite richness of our actions.

LIST OF ABBREVIATIONS

1. L. Wittgenstein's works

AWL	*Wittgenstein's Lectures, Cambridge 1932 – 1935.*
BB	*The Blue and Brown Books.*
CL	*Ludwig Wittgenstein: Cambridge Letters.*
CV	*Culture and Value. Vermischte Bemerkungen.*
LC	*Lectures and Conversations on Aesthetics, Psychology and Religious Belief.*
LFM	*Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge 1939.*
LPP	*Wittgenstein's Lectures on Philosophical Psychology 1946 – 1947.*
LW I, II	*Last Writings on the Philosophy of Psychology. Letze Schriften zur Philosophie der Psychologie.*
LWL	*Wittgenstein's Lectures, Cambridge 1930 – 1932.*
MS	*Wittgenstein's Nachlass.*
NB	*Notebooks 1914 – 1916. Tagebücher 1914 – 1916.*
OC	*On Certainty. Über Gewißheit.*
PG	*Philosophical Grammar. Philosophische Grammatik.*
PI I, II	*Philosophical Investigations. Philosophische Untersuchungen.*
PO	*Philosophical Occasions.*
PR	*Philosophical Remarks. Philosophische Bemerkungen.*
RFM	*Remarks on the Foundations of Mathematics. Bemerkungen zu den Grundlagen der Mathematik.*
RPP I, II	*Remarks on the Philosophy of Psychology. Bemerkungen zu Philosophie der Psychologie.*
TLP	*Tractatus Logico – Philosophicus. Logisch – Philosophische Abhandlung.*
WVC	*Ludwig Wittgenstein and the Vienna Circle. Ludwig Wittgenstein und der Wiener Kreis.*
Z	*Zettel.*

2. C. S. Peirce's works

CD	*The Century Dictionary and Cyclopedia.*
CP	*The Collected Papers of Charles Sanders Peirce.*
DPP	*Dictionary of Philosophy and Psychology*.
EP	*The Essential Peirce.*
HP	*Historical Perspectives on Peirce's Logic of Science: A History of Science.*
LM	«Logical Machines».
NE	*The New Elements of Mathematics.*
RLT	*Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898.*
SW	*Philosophy of Mathematics: Selected Writings.*
W	*Writings of Charles S. Peirce: A Chronological Edition.*

BIBLIOGRAPHY

Ahn, C. H. (1995). Aristotle on Thinking. In M. C. Nussbaum and A. O. Rorty (Eds.), *Essays on Aristotle's De Anima* (pp. 346–366). Oxford: Oxford University Press. [11]

Aizawa, K. (1997). Expaining Systematicity. *Mind and Language, 12*, 115–136. [68]

Akins, K. (1996). Of Sensory Systems and the 'Aboutness' of Mental States. *The Journal of Philosophy, 93*, 337–372. [150] [151]

Akmajian, A., and Demers, R. A. (1990). *Linguistics: An Introduction to Language and Communication*. Cambridge, MA: MIT Press. [100]

Alexander, S. (1920). *Space, Time, and Deity*. London: Macmillan. [85]

Ambrose, A. (1993). Transfinite Numbers. In S. Canfield and J. Shanker (Eds.), *Wittgenstein's Intentions* (pp. 136–146). London & New York: Routledge. [94]

Ameriks, K. (2000). *Kant's Theory of Mind. An Analysis of the Paralogisms of Pure Reason*. Oxford & New York: Oxford University Press. [30]

Anscombe, G. E. M. (1959). *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson University Library. [97]

Anselm of Canterbury. (1969). *Proslogion with a Reply on Behalf of the Fool by Gaunilo and the Author's Reply to Gaunilo* (M. J. Charlesworth, Trans.). Notre Dame: University of Notre Dame Press. [139] [141]

Aquinas, T. (1889). Pars prima Summae theologiae. In *Sancti Thomae de Aquino opera omnia* (Vol. V). Roma: Ex typographia polyglotta. [16] [17] [18]

Aquinas, T. (1984). *Sententia libri De anima* (Vol. XLV). Roma: Ex typographia polyglotta. [11] [17]

Aquinas, T. (2000). Quaestio disputata de spiritualibus creaturis. In *Sancti Thomae de Aquino opera omnia* (Vol. XXIV). Roma: Ex typographia polyglotta. [17]

Aristotle. (1907). *De Anima with Translation, Introduction and Notes* (R. D. Hicks, Trans.). Cambridge, MA: Harvard University Press. [10] [12] [13]

Aristotle. (1926). *Rhetoric* (J. H. Freese, Trans.). Cambridge, MA: Harvard University Press. [41]

Aristotle. (1933). *Metaphysics, Books I–IX* (H. Tredennick, Trans.). Cambridge, MA: Harvard University Press. [15] [40]

Aristotle. (1934). *Nicomachean Ethics* (H. Rackham, Trans.). Cambridge, MA: Harvard University Press. [15] [40]

Aristotle. (1937). *Parts of Animals. Movement of Animals* (A. L. Peck and E. S. Forster, Trans.). Cambridge, MA: Harvard University Press. [12]

Aristotle. (1955). *On Sophistical Reputations. On Coming–to–be and Passing Away. On the Cosmos* (E. S. Foster and D. J. Furley, Trans.). Cambridge, MA: Harvard University Press. [12] [40]

Aristotle. (1970). *Metaphysics. A Revised Text with Introduction and Commentary* (W. D. Ross, Trans.). Oxford: Oxford University Press. [13]

Aristotle. (1975). *Categories and De Interpretatione* (J. L. Ackrill, Trans.). Cambridge, MA: Harvard University Press. [14]

Aristotle. (1995). *Poetics* (S. Halliwell, Trans.). Chicago: University of Chicago Press. [41]

Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. London & New York: Routledge. [48] [66]

Armstrong, D. M. (1997). Mind–Body Problem: Philosophical Theories. In R. L. Gregory (Ed.), *The Oxford Companion to the Mind* (pp. 490–491). Oxford: Oxford University Press. [66]

Atherton, M. (Ed.) (1999). *The Empiricist: Critical Essays*. Lanham: Rowan & Littlefield Publishers. [26]

Atkin, A. (2008). Peirce's Final Account of Signs and the Philosophy of Language. *Transactions of the Charles S. Peirce Society, 44*, 63–85. [181]

Audi, R. (Ed.) (1999). *The Cambridge Dictionary of Philosophy*. Cambridge: Cambridge University Press. [6] [40]

Augustine, A. (1955). *De Civitate Dei contra Paganos*. Turnhout: Brepols. [22]

Augustine, A. (1962). *De doctrina christiana. De vera religione*. Turnhout: Brepols. [22]

Augustine, A. (1968). *De Trinitate libri XV*. Turnhout: Brepols. [22]

Austin, J. L. (1979). *Philosophical Papers*. Oxford: Oxford University Press. [224]

Ayer, J. L. (1966). Can There Be a Private Language? In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 251–266). New York: Anchor Books. [113]

Backer, G. P. (2004). *Wittgenstein's Method: Neglected Aspects*. Oxford: Blackwell. [113]

Backer, G. P., and Hacker, P. M. S. (1980). *Wittgenstein: Understanding and Meaning. Analytical Commentary on the Philosophical Investigations* (Vol. 1). Oxford: Blackwell. [113]

Bahcall, J. N. (2003). Solar Models: An Historical Overview. *Nuclear Physics B, 118*, 77–86. [227]

Baldwin, J. M. (1905). *Dictionary of Philosophy and Psychology*. New York: Macmillan. [179] [212]

Barber, A., and Stainton, R. J. (Eds.) (2010). *Concise Encyclopedia of Philosophy of Language and Linguistics*. Oxford & Boston: Elsevier. [155]

Bayne, S. M. (2004). *Kant on Causation: On the Fivefold Routes on the Principle of Causation*. Albany, NY: SUNY Press. [28]

Beaney, M. (1997). *The Frege Reader*. Oxford: Blackwell. [157]

Beardsley, M. C. (1962). The Metaphorical Twist. *Philosophy and Phenomenological Research, 22*, 293–307. [42]

Bechtel, W., and Hamilton, A. (2007). Reduction, Integration, and the Unity of Science. In D. M. Gabbay, T. Kuipers, P. Thagard, and J. Woods (Eds.), *General Philosophy of Science. Focal Issues* (pp. 377–430). Amsterdam: North Holland. [74]

Behmann, H. (1922). Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem. *Mathematische Annalen, 86*, 163–229. [194]

Bell, J. L. (2006). *The Continuous and the Infinitesimal in Mathematics and Philosophy*. Monza: Polimetrica. [193]

Bennett, M. R., and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Hoboken, NJ: Wiley–Blackwell. [129]

Berkeley, G. (1837). *The Works of George Berkeley, Bishop of Cloyne, To Which Are Added, An Account of His Life, and Several of His Letters to Thomas Prior, Esq, Dean Gevais, Mr. Pope, &c. In One Volume*. London: Thomas Tegg. [23]

Bickle, J. (2006). Reducing Mind to Molecular Pathways: Explicating the Reductionism Implicit in Current Cellular and Molecular Neuroscience. *Synthese, 151*, 411–434. [74]

Biro, J. (2006). Hume's New Science of the Mind. In D. F. Newton (Ed.), *The Cambridge Companion to Hume* (pp. 33–63). Cambridge: Cambridge University Press. [27]

Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science, 9*, 261–325. [72]

Block, N. (1995). Functionalism (2). In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 323–332). Oxford: Blackwell. [70] [72]

Block, N. (Ed.) (1980). *Readings in the Philosophy of Psychology* (Vol. I). Cambridge, MA: Harvard University Press. [47] [50]

Boethius, A. M. S. (1918). *Theological Tractates and the Consolation of Philosophy* (H. F. Stewart, Trans.). London & Cambridge, MA: Harvard University Press. [17]

Boler, J. F. (1963). *Charles Peirce and Scholastic Realism. A Study of Peirce's Relation to John Duns Scotus*. Seattle: University of Washington Press. [194]

Boole, G. (1952). *Collected Logical Works*. La Salle: Open Court. [84]

Bouveresse, J. (1991). *Philosophie, mythologie et pseudo–science. Wittgenstein lecteur de Freud*. Combas: Eclat ed de L'. [145]

Brentano, F. (1867). *Die Psychologie des Aristoteles, insbesondere seine Lehre vom nôus poietikos*. Mainz: Franz Kirchheim. [11]

Broad, C. D. (1925). *The Mind and Its Place in Nature*. London: Routledge. [84] [85]

Broadie, S. (2001). Soul and Body in Plato and Descartes. *Proceedings of the Aristotelian Society, 101*, 295–308. [19]

Broughton, J. (2005). Impression and Ideas. In S. Traiger (Ed.), *The Blackwell Guide to Hume's A Treatise of Human Nature* (pp. 43–58). Oxford: Blackwell. [27]

Brown, R. (1962). Meaning and Rules of Use. *Mind, 71*, 494–511. [100]

Budd, M. (1989). *Wittgenstein's Philosophy of Psychology*. London: Routledge. [129]

Burnyeat, M. F. (1995). How Much Happens when Aristotle Sees Red and Hears Middle C? In M. C. Nussbaum and A. O. Rorty (Eds.) *Essays on Aristotle's De Anima* (pp. 421–434). Oxford: Oxford University Press. [11] [12]

Buroker, J. V. (2006). *Kant's Critique of Pure Reason: An Introduction*. Cambridge: Cambridge University Press. [31]

Butler, J. (1852). *The Analogy of Religion, Natural and Revealed*. London: George Bell. [26]

Candlish, S. (1996). Wittgenstein and the Doctrine of Kinaesthesis. *Australasian Journal of Philosophy, 74*, 581–597. [156]

Candlish, S., and Wrisley, G. (2011). Private Language. *The Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/private–language/ [123]

Canfield, S., and Shanker, J. (Eds.). (1993). *Wittgenstein's Intentions*. London & New York: Routledge. [94]

Cantor, G. (1878). Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die Reine und Angewandte Mathematik, 84*, 242–258. [190]

Cantor, G. (1966). *Gesammelte Abhandlungen. Mathematischen und philosophischen Inhalts. Mit erläuternden Anmerkungen sowie mit Ergänzungen aus dem Briefwechsel Cantor–Dedekind*. Berlin: Springer–Verlag. [189]

Carnap, R. (1931a). Die logizistische Grundlegung der Mathematik. *Erkenntnis, 2*, 91–105. [100]

Carnap, R. (1931b). Die Physikalische Sprache als Universalsprache der Wissenschaft. *Erkenntnis, 2*, 432–465. [194]

Carnap, R. (1936). Testability and Meaning. *Philosophy of Science, 3*, 419–471. [44]

Carnap, R. (1937). Testability and Meaning – Continued. *Philosophy of Science, 4*, 1–40. [44]

Carnap, R. (1968). *Logische Syntax der Sprache*. Wien & New York: Springer. [122] [159]

Carpenter, B. E., and Doran, R. W. (Eds.) (1986). *A. M. Turing's Ace Report of 1946 and Other Papers*. Cambridge, MA: MIT Press. [58] [63]

Caruana, L. (2003). Realism and Rule–Following. In R. Egidi, M. dell'Utri, and M. De Caro (Eds.), *Normatività fatti, valori* (pp. 143–152). Marcerata: Quodlibet. [55] [56]

Caruana, L. (2004). Wittgenstein and the Status of Contradictions. In A. Coliva, and E. Picardi (Eds.), *Wittgenstein Today* (pp. 223–232). Padova: Il poligrafo. [114]

Caston, V. (2006). Aristotle's Psychology. In M. L. Gill, and P. Pellegrin (Eds.), *A Companion to Ancient Philosophy* (pp. 316–346). Oxford: Blackwell. [15]

Cavell, S. (1979). *The Claim of Reason*. Oxford: Oxford University Press. [113]

Chakrabarti, A. (1997). *Denying Existence: The Logic, Epistemology and Pragmatics of Negative Existentials and Fictional Discourse*. Dordrecht: Springer. [139]

Chalmers, D. J. (1993). Why Fodor and Pylyshyn Were Wrong: The Simplest Refutation. *Philosophical Psychology, 6*, 305–319. [68]

Chalmers, D. J. (1996). *The Conscious Mind. In Search of a Fundamental Theory*. Oxford: Oxford University Press. [48]

Chalupa, L. M., and Werner, J. S. (Eds.) (2004). *The Visual Neurosciences* (Vol. II). Cambridge, MA: MIT Press. [52]

Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language, 35*, 26–58. [45] [51]

Church, A. (1936). An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics, 58*. [57]

Churchland, P. M. (1988). *Matter and Consciousness*. Cambridge, MA: MIT Press. [81]

Churchland, P. S. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press. [81]

Cioffi, F. (1998). *Wittgenstein on Freud and Frazer*. Cambridge: Cambridge University Press. [145]

Cobb, J. B., Hartshorne, C., and Gamwell, F. I. (Eds.) (1984). *Existence and Actuality: Conversations with Charles Hartshorne*. Chicago & London: University of Chicago Press. [199]

Cohen, M. (2005). *Wittgenstein's Beetle and Other Classic Thought Experiments*. Oxford: Wiley–Blackwell. [146]

Cohen, P. J. (1963). The Independence of the Continuum Hypothesis, I. *Proceedings of the National Academy of Sciences, USA, 50*, 1143–1148. [190]

Cohen, P. J. (1964). The Independence of the Continuum Hypothesis, II. *Proceedings of the National Academy of Sciences, USA, 51*, 105–110. [190]

Coliva, A., and Picardi, E. (Eds.) (2004). *Wittgenstein Today*. Padova: Il poligrafo. [114]

Conant, J. (2004). Why worry abouth the Tractatus? In B. Stocker (Ed.), *Post–Analytic Tractatus* (pp. 167–192). Aldershot: Routledge. [112] [113]

Copeland, B. J., and Proudfoot, D. (1996). On Alan Turing's Anticipation of Connectionism. *Synthese, 108*, 361–377. [66]

Copi, I. M. (1958). Objects, Properties, and Relations in the Tractatus. *Mind, 67*, 145–165. [97]

Cortazar, J. (1847). Démonatration des analogies de Néper. *Nouvelles annales de mathématiques 1re série, 6*, 218–220. [42]

Couturat, L. (1914). *The Algebra of Logic* (L. G. Robinson, Trans.). Chicago: The Open Court. [96]

Crary, A., and Read, R. (Ed.) (2000). *The New Wittgenstein*. London & New York: Routledge. [112]

Crelle, A. L. (1834). Comment, dans la trigonométrie sphérique, les formules de Gauß et les analogies de Neper, qui en découlent, peuvent être tirées immédiatement et facilement des formules fondamentales. *Journal für die Reine und Angewandte Mathematik, 12*, 348–349. [42]

Crick, F. (1995). *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribner. [38]

Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, MA: MIT Press. [149]

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Vintage Arrow. [150]

Damasio, A. R. (2000). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Orlando: Harcourt Brace. [12]

Dancy, J., and Sosa, E. (Eds.) (1993). *A Companion to Epistemology*. Oxford: Blackwell. [149]

Dauben, J. W. (1979). *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Princeton: Princeton University Press. [193] [195]

Dauben, J. W. (1982). Peirce's Place in Mathematics. *Historia Mathematica, 9*, 311–325. [195]

Davidson, D. (1987a). Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association, 60*, 441–458. [90]

Davidson, D. (1987b). Problems in the Explanation of Action. In P. Pettit, Sylvan, R., and Norman, J. (Eds.), *Metaphysics and Morality. Essays in Honour of J. J. C. Smart* (pp. 101–116). Oxford: Blackwell. [32]

Davidson, D. (1993). Thinking Causes. In J. Heil, and A. Mele (Eds.), *Mental Causation* (pp. 3–17). Oxford: Clarendon Press. [90]

Davidson, D. (1995a). Davidson, Donald. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 231–236). Oxford: Blackwell. [95]

Davidson, D. (1995b). Laws and Cause. *Dialectica, 49,* 263–279. [90]

Davidson, D. (1999a). Reply to Bruce Vermazen. In L. E. Hahn (Ed.), *The Philosophy of Donald Davidson* (pp. 653–656). Chicago: Open Court. [89]

Davidson, D. (2001). *Essays on Actions and Events*. Oxford: Clarendon Press. [85] [88] [89]

Davidson, D. (Ed.) (1999b). *Spinoza's Causal Theory of the Affects*. New York: Fordham University Press. [23]

Davis, M. (Ed.) (1965). *The Undecidable. Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. New York: Dover Publications. [54]

Day, W., and Krebs, V. J. (Eds.) (2010). *Seeing Wittgenstein Anew*. Cambridge: Cambridge University Press. [100]

Debrock, G., and Hulswit, M. (Eds.) (1994). *Living Doubt. Essays Concerning the Epistemology of Charles Sanders Peirce*. Dordrecht: Springer. [180]

Dennett, D. (1971). Intentional Systems. *Journal of Philosophy, 8*, 87–106. [73]

Dennett, D. (1988). Quining Qualia. In A. Marcel, and E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 42–77). New York: Oxford University Press. [81]

Dennett, D. (1991). Two Contrasts: Folk Craft Versus Folk Science, and Belief Versus Opinion. In J. Greenwood (Ed.), *The Future of Folk Psychology* (pp. 135–148). New York: Cambridge University Press. [82]

Dennett, D. (Ed.) (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery: Bradford Books. [73] [81]

Des Chene, D. (2001). *Spirits and Clocks. Machine and Organism in Descartes*. New York: Cornell University Press. [21]

Descartes, R. (1996). *Oeuvres de Descartes. I–XI* (C. Adam, and P. Tannery, Trans.). Paris: Vrin/C.N.R.S. [20]

Devereux, D. (2003). Plato: Metaphysics. In C. Shields (Ed.), *The Blackwell Guide to Ancient Philosophy* (pp. 75–99). Oxford: Wiley–Blackwell. [20]

Diamond, C. (1991). *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind*. Cambridge, MA: MIT Press. [165]

Diamond, C. (2009). Ludwig Josef Johann Wittgenstein. In J. Kim, G. S. Rosenkrantz, and E. Sosa (Eds.), *A Companion to Metaphysics* (pp. 628–632). Oxford: Wiley–Blackwell. [165]

Diamond, C., and Gerrard, S. (1999). How Old Are These Bones? Putnam, Wittgenstein and Verification. *Proceedings of the Aristotelian Society, 73*, 99–150. [138]

Dicker, G. (2001). *Hume's Epistemology*. London: Routledge. [29]

Dreben, B., and Floyd, J. (1991). Tautology: How Not to Use a Word. *Synthese, 87*, 23–49. [36] [37]

Dreben, B., and von Heijenoort, J. (1986). Introductory Note (S. Feferman, Jr., S. C. Kleene, G. H. Moore, R. N. Solovay and J. v. Heijenoort, Trans.). In K. Gödel, *Collected Works: Publications 1929–1936* (Vol. I, pp. 44–59). New York & Oxford: Oxford University Press. [194]

Dretske, F. I. (1969). *Seeing and Knowing*. London: Routledge & K. Paul. [147]

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press. [150]

Dretske, F. I. (1988). *Explaining Behavior. Reasons in a World of Causes*. Cambridge, MA: MIT Press. [72]

Dretske, F. I. (1991). Reply to J. Heil. In B. P. McLaughlin (Ed.), *Dretske and His Critics* (pp. 180–185). Oxford: Blackwell. [151]

Dretske, F. I. (1993). Perceptual Knowledge. In J. Dancy and E. Sosa (Eds.), *A Companion to Epistemology* (pp. 333–338). Oxford: Blackwell. [149]

Dretske, F. I. (1995). *Naturalizing the Mind. The Jean Nicod Lectures*. Cambridge, MA: MIT Press. [149]

Dretske, F. I. (2000). *Perception, Knowledge and Belief. Selected Essays*. Cambridge: Cambridge University Press. [147]

Dretske, F. I. (2003). Intentionality of Perception. In B. Smith (Ed.), *John Searle. Contemporary Philosophy in Focus* (pp. 154–168). Cambridge: Cambridge University Press. [77]

Dreyfus, H. L. (1972). *What Computers Can't Do: A Critique of Artificial Reason*. New York: Harper & Row. [65]

Dreyfus, H. L. (1986). *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Oxford: Blackwell. [65]

Dreyfus, H. L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press. [65]

Duns Scotus, J. (1954). *Opera Omnia* (Vol. III). Vatican: Typis polyglottis vaticanis. [202]

Duns Scotus, J. (1960). *Opera Omnia* (Vol. XVI). Vatican: Typis polyglottis vaticanis. [203]

Egidi, R., dell'Utri, M., and De Caro, M., (Eds.) (2003). *Normatività Fatti, Valori*. Marcerata: Quodlibet. [55] [56]

Everson, S. (Ed.) (1991). *Companions to Ancient Thought: Psychology* (Vol. II). Cambridge: Cambridge University Press. [7]

Feferman, S. (2011). Is the Continuum Hypothesis a Definite Mathematical Problem? *Exploring the Frontiers of Independence. Harvard Lecture Series*. Retrieved from https://math.stanford.edu/~feferman/papers/BernaysLecture2.pdf/ [191]

Feigl, H., and Scriven, M. (Eds.) (1956). *The Foundations of Science and the Concepts of Psychology and Psychoanalysis: Minnesota Studies in the Philosophy of Science* (Vol. I). Minneapolis: University of Minnesota Press. [81]

Fisch, M. H. (1978). Peirce's General Theory of Signs. In T. A. Sebeok (Ed.), *Sight, Sound, and Sense* (pp. 31–70). Bloomington: Indiana University Press. [168]

Flew, A. (1951). Locke and the Problem of Personal Identity. *Philosophy, 26*, 53–68. [26]

Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press. [51]

Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press. [81] [133]

Fodor, J. A., and Lepore, E. (1999). Impossible Words? *Linguistic Inquiry, 30*, 445–453. [69]

Fodor, J. A., and McLaughin, B. P. (1991). Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work. In C. Macdonald and G. Macdonald (Eds.), *Connectionism: Debates on Psychological Explanation* (pp. 199–222). Oxford: Wiley–Blackwell. [66]

Fodor, J. A., and Pylyshyn. Z. (1988). Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition, 28*, 3–71. [68]

Føllesdal, D. (1967). Comments on Stenius's 'Mood and Language–Game'. *Synthese, 17*, 275–280. [221]

Forster, P. (2011). *Peirce and the Threat of Nominalism*. Cambridge: Cambridge University Press. [175]

Fourneret, P., and Jeannerod, M. (1998). Limited Conscious Monitoring of Motor Performance in Normal Subjects. *Neuropsychologia, 36*, 1133–1140. [76]

Franzén, T. (2005). *Gödel's Theorem. An Incomplete Guide to Its Use and Abuse*. Wellesley: CRC Press. [63] [64]

Frede, D., and Reis, B. (Eds.) (2009). Body and Soul in Ancient Philosophy. Berlin: de Gruyter. [10]

Freedman, E. (1983). *The Relevance of Charles Peirce*. La Salle: Monist Library of Philosophy. [199]

Frege, G. (1950). *The Foundations of Arithmetic. A Logico–Mathematical Enquiry into the Concept of Number* (J. L. Austin, Trans.). New York: Harper. [38]

Frege, G. (1966). *Gründgesetze der Arithmetik* (Vol. I). (I. Angelelli Ed.). Hildesheim: Georg Olms. [141]

Frege, G. (1967). *Kleine Schriften*. (I. Angelelli Ed.). Hildesheim: Georg Olms. [41] [84] [96] [157]

Frege, G. (1988). *Begriffsschrift und andere Aufsatze* (I. Angelelli Ed.). Hildesheim: Georg Olms. [156]

Gabbay, D. M., and Woods, J. (Eds.) (2004). *Handbook of the History of Logic: The Rise of Modern Logic: From Leibniz to Frege* (Vol. III). Amsterdam: Elsevier. [96]

Gabbay, D. M., Kuipers, T., Thagard, P., and Woods, J. (Eds.) (2007). *General Philosophy of Science. Focal Issues*. Amsterdam: North Holland. [74]

Gardner, E. P., Martin, J. H., and Jessell, T. M. (2000). The Bodily Senses. In Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (Eds.), *Principles of Neural Science* (pp. 430–451). New York: McGraw–Hill. [156]

Garnham, A. (1995). Cognitive Psychology. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 167–176). Oxford: Blackwell. [223]

Garson, J. (1997). Syntax in a Dynamic Brain. *Synthese, 110*, 343–355. [67]

Garson, J. (2010). Connectionism. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/win2010/entries/connectionism/ [67]

Gilbert, P. (1991). *La simplicité du principe. Prolégomènes à la métaphysique*. Namur: Lessius. [36]

Gilbert, P. (1995). *La patience d'être. Métaphysique*. Namur: Lessius. [36]

Gill, M. L., and P. Pellegrin, P. (Eds.), (2006). A Companion to Ancient Philosophy. Oxford: Blackwell. [15] [42]

Glimcher, P. W. (2011). *Foundations of Neuroeconomic Analysis*. New York: Oxford University Press. [161]

Gödel, K. (1986). *Collected Works: Publications 1929–1936* (S. Feferman, Jr., S. C. Kleene, G. H. Moore, R. N. Solovay and J. v. Heijenoort, Trans. Vol. I). New York & Oxford: Oxford University Press. [63] [64] [158] [194]

Gödel, K. (1990). *Collected Works: Publications 1938–1974* (J S. Feferman, Jr., S. C. Kleene, G. H. Moore, R. N. Solovay and J. v. Heijenoort, Trans. Vol. II). New York & Oxford: Oxford University Press. [190]

Gödel, K. (1995). *Collected Works: Unpublished Essays and Lectures* (J S. Feferman, Jr., S. C. Kleene, G. H. Moore, R. N. Solovay and J. v. Heijenoort, Trans. Vol. III). Oxford: Oxford University Press. [59] [64]

Graham, G. (2011). Behaviorism. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2010/entries/behaviorism/ [43]

Greenwood, J. (Ed.) (1991). *The Future of Folk Psychology*. New York: Cambridge University Press. [82]

Gregory, R. L. (1987). 'Emergence' and 'Reduction' in Explanations. In R. L. Gregory (Ed.), *The Oxford Companion to the Mind*. Oxford: Oxford University Press. [83]

Gregory, R. L. (Ed.) (1987). *The Oxford Companion to the Mind*. Oxford: Oxford University Press. [66]

Guttenplan, S. (Ed.) (1995). *A Companion to the Philosophy of Mind.* Oxford: Blackwell. [20] [32] [50] [70] [71] [72] [73] [75] [86] [95] [223] [239]

Guzzo, A., Mathieu, V., and Lia, P. (2006). Analogia. In V. Melchiorre (Ed.), *Enciclopedia filosofica* (pp. 402–410). Milan: Bompiani. [42]

Hacker, P. M. S. (1990). *Wittgenstein: Meaning and Mind, Analytical Commentary on the Philosophical Investigations* (Vol. III). Oxford: Wiley–Blackwell. [112] [120] [123] [129]

Hacker, P. M. S. (2000). Was He Trying to Whistle It? In A. Crary and Read, R. (Eds.), *The New Wittgenstein* (pp. 353–388). London & New York: Routledge. [113]

Hacker, P. M. S. (2001). *Connections and Controversies*. Oxford: Clarendon Press. [113]

Hacker, P. M. S. (2010). Robinson Crusoe Sails Again: The Interpretative Relevance of Wittgenstein's Nachlass. In N. Venturinha (Ed.), *Witttgenstein After His Nachlass* (pp. 91–109). Chicago: Open Court. [112]

Hahn, L. E. (Ed.) (1999). *The Philosophy of Donald Davidson*. Chicago: Open Court. [89]

Haldane, J. (1995). History: Medieval and Renaissance Philosophy of Mind. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 333–338). Oxford: Blackwell. [20]

Haller, R., and Puhl, K. (Eds.) (2002). *Wittgenstein and the Future of Philosophy: A Reassessment after 50 Years. Proceedings of the 24th International Wittgenstein Symposium*. Wien: obv & hpt. [163]

Harel, D. (2000). Computers Ltd.: What They Really Can't Do. Oxford & New York: Oxford University Press. [65]

Hark, M. T. (1990). *Beyond the Inner and the Outer: Wittgenstein's Philosophy of Psychology*. Dordrecht: Springer. [110]

Hark, M. T. (1995). Electric Brain Fields and Memory Traces: Wittgenstein and Gestalt Psychology. *Philosophical Investigations*, *18*, 113–138. [115]

Harré, R. (2001). Wittgenstein: Science and Religion. *Philosophy*, *76*, 211–237. [97]

Hartshorne, C. (1929). Continuity, the Form of Forms in Charles Peirce. *The Monist*, *39*, 521–534. [188]

Hartshorne, C. (1983). A Revision of Peirce's Categories. In E. Freedman (Ed.) *The Relevance of Charles Peirce* (pp. 80–93). Chicago: La Salle. [199]

Haufeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge, MA: MIT Press. [59] [61]

Havenel, J. (2008). Peirce's Clarifications of Continuity. *Transactions of the Charles S. Peirce Society*, *44*, 86–133. [186] [187] [199]

Hegel, G. W. F. (1978). *Wissenschaft der Logik. Erste Band: Die objektive Logik*. Hamburg: Felix Meiner Verlag. [37]

Hegel, G. W. F. (1989). *The Science of Logic* (A. V. Miller, Trans.). Atlantic Highlands: Humanity Books. [37]

Heil, J. (1991). Perceptual Experience. In B. P. McLaughlin (Ed.), *Dretske and His Critics* (pp. 1–16. 180–185). Oxford: Blackwell. [151]

Heil, J., and Mele, A. (Eds.) (1993). *Mental Causation*. Oxford: Clarendon Press. [90]

Hellett, G. (1977). *A Companion to Wittgenstein's Philosophical Investigations*. Ithaca & London: Cornell University Press. [123]

Hess, R. F. (2004). Spatial Scale in Visual Processing. In L. M. Chalupa and J. S. Werner (Eds.), *The Visual Neurosciences* (Vol. II, pp. 1043–1059). Cambridge, MA: MIT Press. [52]

Hills, D. (2012). Metaphor. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/win2012/entries/metaphor/ [42]

Hilpinen, R. (2004). Peirce's Logic. In D. M. Gabbay and J. Woods (Eds.), *Handbook of the History of Logic: The Rise of Modern Logic: From Leibniz to Frege* (Vol. III, pp. 611–658). Amsterdam: Elsevier. [96]

Hintikka, J. (1962). *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca: Cornell University Press. [208]

Hintikka, J. (1967). Time, Truth, and Knowledge in Ancient Greek Philosophy. *American Philosophical Quarterly, 4*, 1–14. [15]

Hintikka, J. (1968). Epistemic Logic and the Methods of Philosophical Analysis. *Australasian Journal of Philosophy, 46*, 37–51. [208]

Hintikka, J. (1969a). *Models for Modalities: Selected Essays*. Dordrecht: D. Reidel. [137] [140]

Hintikka, J. (1969b). Wittgenstein on Private Language: Some Sources of Misunderstanding. *Mind, 78*, 423–425. [126]

Hintikka, J. (1973). *Logic, Language–Games and Information: Kantian Themes in the Philosophy of Logic*. Oxford: Oxford University Press. [137] [208]

Hintikka, J. (1980). C. S. Peirce's 'First Real Discovery' and Its Contemporary Relevance. *The Monist, 63*(3), 304–315. [96] [195]

Hintikka, J. (1985). A Spectrum of Logics of Questioning. *Philosophica, 35*, 135–150. [213]

Hintikka, J. (1988). On the Development of the Model–Theoretic Viewpoint in Logical Theory. *Synthese, 77*, 1–36. [96]

Hintikka, J. (1997). *Lingua Universalis versus Calculus Ratiocinator. Selected Papers* (Vol. II). Dordrecht: D. Reidel. [96]

Hintikka, J. (1999). *Inquiry as Inquiry: A Logic of Scientific Discovery. Selected Papers* (Vol. III). Dordrecht: D. Reidel. [213]

Hintikka, J. (2000). *On Gödel*. Belmont: Wadsworth. [63] [64]

Hintikka, J. (2003). What Does the Wittgensteintian Inexpressible Express? *The Harvard Review of Philosophy, 11*, 9–17. [230]

Hintikka, J. (2007). *Socratic Epistemology: Explorations of Knowledge–Seeking by Questioning*. Cambridge: Cambridge University Press. [208]

Hintikka, J. (Ed.) (1995). *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*. Dordrecht & Boston: Springer. [158]

Hintikka, J., and Hintikka, A.–M. (2002). Wittgenstein the Bewitched Writer. In Haller, R., and Puhl, K. (Eds.), *Wittgenstein and the Future of Philosophy: A Reassessment after 50 Years. Proceedings of the 24th International Wittgenstein Symposium* (pp. 131–150). Wien: obv & hpt. [63] [163]

Hintikka, J., and Merrill, M. B. (1986). *Investigating Wittgenstein*. Oxford: Blackwell. [97] [106] [110] [111] [116] [138] [145]

Hobbes, T. (1839). *The English Works of Thomas Hobbes of Malmesbury. Now First Collected and Edited by Sir William Molesworth, Bart*. (Vols. I–III). London: Longman. [23]

Hodges, A. (1983). *Alan Turing: The Enigma*. New York: Vintage Books. [53]

Hoffman, P. (1990). St. Thomas Aquinas on the Halfway State of Sensible Being. *Philosophical Review, 99*, 73–92. [17]

Hofstadter, D. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books. [64]

Holborow, L. (1973). The 'Prejudice in Favour of Psychophysical Parallelism'. *Royal Institute of Philosophy Lectures, 7*, 193–207. [129]

Honderich, T. (1982). The Argument for Anomalous Monism. *Analysis, 42*, 59–64. [90]

Hook, S. (Ed.) (1960). *Dimensions of Mind. A Symposium*. New York: New York University Press. [65]

Hookway, C. J. (1985). *Peirce*. London: Routledge and K. Paul. [185]

Horgan, T. (1990). Soft Laws. *Midwest Studies in Philosophy, 15*, 256–279. [68]

Horgan, T., and Tienson, J. (1989). Representations without Rules. *Philosophical Topics, 17*, 147–174. [68]

Huffman, C. A. (2009). The Pythagorean Conception of the Soul from Pythagoras to Philolaus. In D. Frede and B. Reis (Eds.), *Body and Soul in Ancient Philosophy* (pp. 21–43). Berlin: de Gruyter. [10]

Hull, D. L. (1972). Reduction in Genetics – Biology or Philosophy? *Philosophy of Science, 39*, 491–499. [74]

Hull, D. L. (1974). *The Philosophy of Biological Science*. Eaglewood Cliffs: Prentice Hall College Div. [74]

Hulswit, M. (Ed.) (1994). *Living Doubt. Essays Concerning the Epistemology of Charles Sanders Peirce*. Dordrecht: Springer. [180]

Hume, D. (1975). *An Enquiry Concerning Human Understanding, in Id., Enquiries Concerning Human Understanding and Concerning the Principles of Morals* (L. A. Selby–Bigge and P. H. Nidditch Eds.). Oxford: Oxford University Press. [28] [29]

Hume, D. (1978). *A Treatise of Human Nature* (L. A. Selby–Bigge and P. H. Nidditch Eds.). Oxford: Clarendon Press. [29]

Iordache, O. (2011). *Modeling Multi–Level Systems*. Berlin: Springer. [193]

Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly, 32*, 127–136. [72]

Jackson, F. (1986). What Mary Didn't Know. *The Journal of Philosophy, 83*, 291–295. [72]

Jackson, F., Pargetter, R., and Prior, E. W. (1981). Functionalism and Type–Type Identity Theories. *Philosophical Studies, 42*, 209–225. [92]

James, W. (1950). *The Principles of Psychology* (Vols. I–II). New York: Dover Publications. [153] [206]

James, W. (1996). *Essays in Radical Empiricism*. Lincoln: Nebraska Press. [162]

Johnson, W. E. (1922). *Logic*. Cambridge: Cambridge University Press. [158]

Jones, O. R. (Ed.) (1971). *The Private Language Argument*. London: St. Martin's Press. [113]

Kandel, E. R., and Wurtz, R. H. (2000). Constructing the Visual Image. In Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (Eds.), *Principles of Neural Science* (pp. 492–506). New York: McGraw–Hall. [151]

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (Eds.) (2000). *Principles of Neural Science*. New York: McGraw–Hall. [151]

Kant, I. (1922–1923). *Immanuel Kants Werke* (O. Buek Ed. Vols. I–XI). Berlin: Bruno Cassirer. [37]

Kant, I. (1956). *Kritik der reinen Vernunft*. Hamburg: Felix Meiner Verlag. [31]

Kant, I. (1968). *Gesammelte Schriften* (Vol. XXVIII/1). Berlin & Leipzig: De Gruyter. [30]

Kant, I. (1988). *Logic* (R. S. Hartman and W. Schwarz, Trans.). New York: Dover Publications. [37]

Kant, I. (1996). *Critique of Pure Reason* (W. S. Pluhar, Trans.). Unified Edition. Indianapolis: Hackett Publishing. [238]

Kenny, A. (1966). Cartesian Privacy. In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 352–370). Melbourne: Macmillan. [113]

Kenny, A. (1971). Verification Principle and the Private Language Argument. In O. R. Jones (Ed.), *The Private Language Argument* (pp. 204–228). New York: St. Martin's Press. [113]

Kenny, A. (2004). *Aquinas on Mind*. London: Routledge. [17]

Ketner, K. L. (Ed.) (1955). *Peirce and Contemporary Thought: Philosophical Inquiries*. New York: Fordham University Press. [186]

Kim, J. (1987). 'Strong' and 'Global' Supervenience Revisited. *Philosophy and Phenomenological Research, 68*, 315–326. [86]

Kim, J. (1993). Can Supervenience and 'Non–Strict Laws' Save Anomalous Monism? In J. Heil and A. Mele (Eds.), *Mental Causation* (pp. 19–26). Oxford: Clarendon Press. [90]

Kim, J. (1995). Supervenience. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 575–583). Oxford: Blackwell. [86]

Kim, J. (1999). Making Sense of Emergence. *Philosophical Studies, 95*, 3–36. [85]

Kim, J., Sosa, E., and Rosenkrantz, G. S. (Eds.) (2009). *A Companion to Metaphysics*. Oxford: Wiley–Blackwell. [165]

Klage, J. C. (1989). Wittgenstein and Neuroscience. *Synthese, 78*, 319–343. [129]

Knuuttila, S., and Hintikka, J. (Eds.) (1986). *The Logic of Being. Historical Studies*. Dordrecht: Springer. [41]

Köhler, W. (1930). *Gestalt Psychology*. London: Bell & Sons. [153]

Kretzmann, N. (1992). Aquinas' Philosophy of Mind. *Philosophical Topics, 20*, 77–101. [17]

Kripke, S. (1971). Identity and Necessity. In M. K. Munitz (Ed.), *Identity and Individuation* (pp. 93–121). New York: New York University Press. [68]

Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press. [49] [84] [93]

Kuhn, S. T. (1999). Sortal Predicate. In R. Audi (Ed.), *The Cambridge Dictionary of Philosophy* (p. 865). Cambridge: Cambridge University Press. [40]

Kuusela, O., and McGinn, M. (Eds.) (2011). *The Oxford Handbook of Wittgenstein*. Oxford: Oxford University Press. [100] [112] [138] [159] [164]

Lardner, D. (1828). *An Analytical Treatise on Plane and Spherical Trigonometry and the Analysis of Angular Sections*. London: John Taylor. [42]

Lazerowitz, M. (1977). *The Language of Philosophy: Freud and Wittgenstein*. Dordrecht: Springer. [145]

Leder, D. (1985). Troubles with Token Identity. *Philosophical Studies, 47*, 79–94. [89]

Legg, C. (2008). The Problem of the Essential Icon. *American Philosophical Quarterly, 45*, 207–232. [213]

Leiber, J. (1991). *An Invitation to Cognitive Science*. Oxford: Blackwell. [66]

Leibniz, G. W. (1890). *Die philosophischen Schriften* (C. I. Gerhardt Ed. Vol. VII). Berlin: Weidmannsche Buchhandlung. [38]

Leibniz, G. W. (1904). *Hauptschriften zur Grundlegung der Philosophie* (A. Buchenau Ed. Vol. I). Leipzig: Dürr. [96]

Leibniz, G. W. (1976). *Philosophical Papers and Letters* (L. E. Loemker, Trans.). Chicago: Springer. [38]

Leibniz, G. W. (1991). *La Monadologie*. Paris: MAPbook. [23]

Leibniz, G. W. (2010). *Essais De Théodicée Sur La Bonté De Dieu, La Liberté De L'homme et L'origine Du Mal*. Paris: Nabu Press. [23]

Lesher, J. H. (1971). Aristotle on Form, Substance, and Universals: A Dilemma. *Phronesis, 16*, 169–178. [13]

Leslie, S. (2004). *The English Utilitarians* (Vol. III: John Stuart Mill). Cambridge: Cambridge University Press. [83]

Lewis, C. I., and Langford, C. H. (1932). *Symbolic Logic*. New York: The Century Co. [38]

Lewis, D. (1980). Mad Pain and Martian Pain. In N. Block (Ed.), *Readings in the Philosophy of Psychology* (Vol. I, pp. 216–222). Cambridge, MA: Harvard University Press. [47] [50]

Libet, B. (1985). Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *Behavioral and Brain Sciences, 8*, 529–539. [77]

Liszka, J. J. (1996). *A General Introduction to the Semeiotic of Charles Sanders Peirce*. Bloomington & Indianapolis: Indiana University Press. [208]

Locke, J. (1975). *An Essay Concerning Human Understanding* (P. H. Nidditch Ed.). Oxford: Clarendon Press. [24] [25] [26] [27] [28] [36] [168] [200]

Lovibond, S. (1991). Plato's Theory of Mind. In S. Everson (Ed.), *Companions to Ancient Thought: Psychology* (Vol. II, pp. 35–55). Cambridge: Cambridge University Press. [7]

Löwenheim, L. (1915). Über Möglichkeiten im Relativkalkül. *Mathematische Annalen, 76*, 447–470. [96] [194]

Lucas, J. (1961). Minds, Machines and Gödel. *Philosophy, 36*, 112–127. [64]

Lycan, W. G. (1981). Form, Function, and Feel. *Journal of Philosophy, 78*, 24–49. [73]

Lycan, W. G. (1995). Functionalism (1). In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 317–323). Oxford: Blackwell. [71]

Lycan, W. G. (Ed.) (2008). *Mind and Cognition. A Reader*. Oxford: Blackwell. [73]

Macdonald, C., and Macdonald, G. (1986). Mental Causes and Explanation of Action. *Philosophical Quarterly, 36*, 145–158. [90]

Macdonald, C., and Macdonald, G. (Eds.) (1991). *Connectionism: Debates on Psychological Explanation*. Oxford: Wiley–Blackwell. [66]

Mach, E. (1918). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Jena: Verlag der Wissenschaften. [162]

Malcolm, N. (1966). Wittgenstein's Philosophical Investigations. In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 65–103). Melbourne: Macmillan. [112] [113]

Malcolm, N. (1986). *Nothing Is Hidden: Wittgenstein's Criticism of His Early Thought*. Oxford: Basil Blackwell. [129]

Malebranche, N. (1711). *Entretiens sur la métaphysique, sur la religion et sur la mort* (Vol. I). Paris: Michel David. [23]

Manson, N. (2003). Consciousness. In B. Smith (Ed.), *John Searle. Contemporary Philosophy in Focus* (pp. 128–154). Cambridge: Cambridge University Press. [79]

Marcel, A., and Bisiach, E. (Eds.) (1988). *Consciousness in Contemporary Science*. New York: Oxford University Press. [81]

Marion, M. (1995). Wittgenstein and Ramsey on Identity. In J. Hintikka (Ed.), *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics* (pp. 343–371). Dordrecht & Boston: Springer. [158]

Marr, D. (1983). *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*. London & Cambridge, MA: MIT Press. [51]

Maturana, H. R., and Varela, F. J. (1972). *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht: Springer. [223] [224]

Maynard Smith, J., and Price, G. R. (1973). The Logic of Animal Conflict. *Nature, 246*, 15–18. [216]

McCann, E. (1995). History: Philosophy of Mind in the Seventeenth and Eighteenth Centuries. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 338–347). Oxford: Blackwell. [20] [27] [32]

McCann, E. (1999). Locke on Identity: matter, life, consciousness. In M. Atherton (Ed.), *The Empiricist: Critical Essays* (pp. 68–83). Lanham: Rowan & Littlefield Publishers. [26]

McClelland, J. L., and Rumelhart, D. E., and the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Psychological and biological models* (Vol. II). Cambridge, MA: MIT Press. [66] [67]

McGinn, M. (1997). *Wittgenstein and the Philosophical Investigations*. London: Routledge. [144]

McLaughlin, B. P. (1999). Philosophy of Mind. In R. Audi (Ed.), *The Cambridge Dictionary of Philosophy* (pp. 684–694). Cambridge: Cambridge University Press. [6]

McLaughlin, B. P. (Ed.) (1991). *Dretske and His Critics*. Oxford: Blackwell. [151]

McLaughlin, B. P., and Bennett, K. (2011). Supervenience. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/win2011/entries/supervenience/ [85]

Melchiorre, V. (Ed.) (2006). *Enciclopedia filosofica* (Vol. I). Milan: Bompiani. [42]

Meltzer, B., and Michie, D. (Eds.) (1969). *Machine Intelligence* (Vol. 5). Edinburgh: Edinburgh University Press. [61] [66]

Midtgarden, T. (2002). Peirce on the Notion of Self and Personal Identity. *History of Philosophy Quarterly, 19*, 109–124. [206]

Mill, J. S. (1974). *Collected Works of John Stuart Mill* (Vols. VIII–IX). Toronto: University of Toronto Press. [83] [84]

Millican, P., and Beebee, H. (2007). Humes Old and New. *Proceedings of the Aristotelian Society 81*, 163–199. [29]

Minar, E. (2010). The Philosophical Significance of Meaning–Blindness. In W. Day and V. J. Krebs (Eds.), *Seeing Wittgenstein Anew* (pp. 183–203). Cambridge: Cambridge University Press. [144]

Minar, E. (2011). The Life of the Sign: Rule–Following, Practice, and Agreement. In O. Kuusela and M. McGinn (Eds.), *The Oxford Handbook of Wittgenstein* (pp. 276–293). Oxford: Oxford University Press. [100]

Misak, C. (Ed.) (2006). *The Cambridge Companion to Peirce*. Cambridge: Cambridge University Press. [177] [182]

Monk, R. (1990). *Ludwig Wittgenstein: The Duty of Genius*. London: Penguin Books. [44] [52] [125] [140] [145]

Moore, A. W. (2011). Wittgenstein and Infinity. In O. Kuusela and M. McGinn (Eds.), *The Oxford Handbook of Wittgenstein* (pp. 105–121). Oxford: Oxford University Press. [159]

Moore, S., Gielgud, J., and Logan, J. (1984). *The Stanislavski System: The Professional Training of an Actor. Second Revised Edition*. New York: Penguin Books. [147]

Moravcsik, J. (1995). Explaining Various Forms of Living. In M. C. Nussbaum and A. O. Rorty (Eds.), *Essays on Aristotle's De Anima* (pp. 129–145). Oxford: Oxford University Press. [11]

Moyal–Sharrock, D. (2004). *Understanding Wittgenstein's On Certainty*. Basingstoke: Palgrave Macmillan. [144]

Mueller, I. (2006). Greek Mathematics (Arithmetic, Geometry, Proportion Theory) to the Time of Euclid. In M. L. Gill and P. Pellegrin (Eds.), *A Companion to Ancient Philosophy* (pp. 686–718). Oxford: Blackwell. [42]

Mulhall, S. (2007). *Wittgenstein's Private Language: Grammar, Nonsense, and Imagination in Philosophical Investigations, §§243–315*. Oxford: Oxford University Press. [112]

Munitz, M. K. (Ed.) (1971). *Identity and Individuation*. New York: New York University Press. [68]

Murphey, M. (1961). *The Development of Peirce's Philosophy*. Cambridge, MA: Harvard University Press. [191] [193]

Nadelhoffer, T. (2011). Neural Lie Detection, Criterial Change, and Ordinary Language. *Neuroethics, 4*, 205–213. [129]

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review, 83*. [72]

Neurath, O. (1932). Protokollsätze. *Erkenntnis, 3*, 204–214. [100]

Newton–Smith, W. H. (Ed.) (2000). *A Companion to the Philosophy of Science*. Oxford: Blackwell. [20]

Newton, D. F. (Ed.) (2006). *The Cambridge Companion to Hume*. Cambridge: Cambridge University Press. [27]

Nielsen, K. S. (2008). *The Evolution of the Private Language Argument*. Aldershot: Palgrave Macmillan. [112]

Noonan, H. W. (2003). *Personal Identity*. London: Routledge. [26]

Nussbaum, M. C., and Putnam, H. (1995). Changing Aristotle's Mind. In M. C. Nussbaum and A. O. Rorty (Eds.), *Essays on Aristotle's De Anima* (pp. 27–56). Oxford: Oxford University Press. [238]

Nussbaum, M. C., and Rorty, A. O. (Eds.) (1995). *Essays on Aristotle's De Anima*. Oxford: Oxford University Press. [11] [12] [238]

O'Connor, T., and Wong, H. Y. (2009). Emergent Properties. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/spr2009/entries/properties–emergent/ [85]

Olson, E. T. (2010). Personal Identity. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/entries/identity–personal/ [26]

Overgaard, S. (2004). Exposing the Conjuring Trick: Wittgenstein on Subjectivity. *Phenomenology and the Cognitive Sciences, 3*, 263–286. [44]

Pasnau, R. (2003). *Thomas Aquinas on Human Nature: A Philosophical Study of Summa theologiae 1ª 75–89*. Cambridge: Cambridge University Press. [17]

Paull, R. C., and Sider, T. R. (1992). In Defense of Global Supervenience. *Philosophy and Phenomenological Research, 52*, 833–854. [86]

Peirce, C. S. (1887). Logical Machines. *The American Journal of Psychology, 1*, 165–170. [128] [201]

Peirce, C. S. (1931). *Collected Papers of Charles Sanders Peirce. Volume I: Principles of Philosophy* (C. W. Hartshorne and P. Weiss Eds.). Cambridge, MA: Harvard University Press. [84] [172] [182] [183] [194]

Peirce, C. S. (1932). *Collected Papers of Charles Sanders Peirce. Volume II: Elements of Logic* (C. W. Hartshorne and P. Weiss Eds.). Cambridge, MA: Harvard University Press. [92] [179] [183] [187]

Peirce, C. S. (1933). *Collected Papers of Charles Sanders Peirce. Volume III: Exact Logic* (C. W. Hartshorne and P. Weiss Eds.). Cambridge, MA: Harvard University Press. [93] [182] [185] [193] [197] [199]

Peirce, C. S. (1934). *Collected Papers of Charles Sanders Peirce. Volume V: Pragmatism and Pragmaticism* (C. W. Hartshorne and P. Weiss Eds.). Cambridge, MA: Harvard University Press. [35]

Peirce, C. S. (1935). *Collected Papers of Charles Sanders Peirce. Volume VI: Scientific Metaphysics* (C. W. Hartshorne and P. Weiss Eds.). Cambridge, MA: Harvard University Press. [186] [189] [197] [201] [203] [209] [212] [225] [227]

Peirce, C. S. (1958a). *Collected Papers of Charles Sanders Peirce. Volume VII: Science and Philosophy* (A. W. Burks Ed.). Cambridge, MA: Harvard University Press. [203] [204] [208] [218]

Peirce, C. S. (1958b). *Collected Papers of Charles Sanders Peirce. Volume VIII: Review, Correspondence, and Bibliography* (A. W. Burks Ed.). Cambridge, MA: Harvard University Press. [181] [187] [189] [204] [214]

Peirce, C. S. (1976a). *The New Elements of Mathematics. Volume 2: Algebra and Geometry* (C. Eisele Ed.). Highlands, NJ: Humanities Books. [191] [192] [196]

Peirce, C. S. (1976b). *The New Elements of Mathematics. Volume 3: Mathematical Miscellanea* (C. Eisele Ed.). Highlands, NJ: Humanities Books. [188] [190] [191] [195]

Peirce, C. S. (1976c). *The New Elements of Mathematics. Volume 4: Mathematical Philosophy* (C. Eisele Ed.). Highlands, NJ: Humanities Books. [192] [205] [218]

Peirce, C. S. (1982). *Writings of Charles S. Peirce. A Chronological Edition. Volume 1: 1857–1866* (E. C. Moore Ed.). Bloomington, IN: Indiana University Press. [167] [170] [195] [209] [225]

Peirce, C. S. (1984). *Writings of Charles S. Peirce. A Chronological Edition. Volume 2: 1867–1866* (E. C. Moore Ed.). Bloomington, IN: Indiana University Press. [167] [171] [172] [173] [174] [187] [199] [204] [218]

Peirce, C. S. (1985). *Historical Perspectives on Peirce's Logic of Science: A History of Science* (C. Eisele Ed.). Berlin & New York: Mouton Publishers. [198]

Peirce, C. S. (1986). *Writings of Charles S. Peirce. A Chronological Edition. Volume 3: 1872–1878* (C. J. W. Kloesel Ed.). Bloomington, IN: Indiana University Press. [175] [176] [185]

Peirce, C. S. (1989). *Writings of Charles S. Peirce. A Chronological Edition. Volume 4: 1879–1884* (C. J. W. Kloesel Ed.). Bloomington, IN: Indiana University Press. [178]

Peirce, C. S. (1992a). *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898* (K. L. Ketner Ed.). Cambridge, MA: Harvard University Press. [183] [194]

Peirce, C. S. (1992b). *The Essential Peirce. Volume 1: Selected Philosophical Writings (1867–1893)* (N. Houser and C. Kloesel Eds.). Bloomington & Indianapolis: Indiana University Press. [194] [196] [205]

Peirce, C. S. (1993). *Writings of Charles S. Peirce. A Chronological Edition. Volume 5: 1884–1886* (C. J. W. Kloesel Ed.). Bloomington, IN: Indiana University Press. [178] [179] [225]

Peirce, C. S. (1998). *The Essential Peirce. Volume 2: Selected Philosophical Writings (1893–1913)* (The Peirce Editorial Project Ed.). Bloomington & Indianapolis: Indiana University Press. [180] [209] [225]

Peirce, C. S. (2010). *Philosophy of Mathematics. Selected Writings* (M. E. Moore Ed.). Bloomington & Indianapolis: Indiana University Press. [188] [194] [199]

Penrose, R. (1996). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford: Oxford University Press. [64]

Penrose, R. (2002). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press. [64]

Perler, D. (2001). Ancient and Medieval Theories of Intentionality. In P. Pettit, Sylvan, R., and Norman, J. (Eds.), *Metaphysics and Morality. Essays in Honour of J. J. C. Smart*. Oxford: Blackwell. [11]

Pettit, P. (2008). *Made with Words. Hobbes On Language, Mind, and Politics*. Princeton: Princeton University Press. [24]

Pettit, P., Sylvan, R., and Norman, J. (1987) (Eds.), *Metaphysics and Morality. Essays in Honour of J. J. C. Smart*. Oxford: Blackwell. [32]

Pietarinen, A.–V. (2006). *Signs of Logic: Peircean Themes on the Philosophy of Language, Games, and Communication*. Dordrecht: Springer. [160] [191] [194] [213] [230]

Pitcher, G. (Ed.) (1966). *Wittgenstein: The Philosophical Investigations*. Melbourne: Macmillan. [112] [113]

Place, U. T. (1954). The Concept of Heed. *British Journal of Psychology, 45*, 243–255. [47]

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

Plato. (1926a). *Cratylus, Parmenides, Greater Hippias, Lesser Hippias* (H. N. Fowler, Trans.). Cambridge, MA: Harvard University Press. [15]

Plato. (1926b). *Laws, Books 7–12* (R. G. Bury, Trans.). Cambridge, MA: Harvard University Press. [9]

Plato. (1935). *The Republic, Books 6–10* (P. Shorey, Trans.). Cambridge, MA: Harvard University Press. [8]

Plato. (1953). *Euthyphro. Apology. Crito. Phaedo. Phaedrus* (H. N. Fowler, Trans.). Cambridge, MA: Harvard University Press. [8] [9] [208]

Plato. (1977). *Theaetetus* (J. McDowell, Trans.). Oxford: Oxford University Press. [9]

Poincaré, H. (1935). *La science et l'hypothèse*. Paris: Ernest Flammarion. [37]

Polansky, R. (2007). *Aristotle's De anima*. Cambridge: Cambridge University Press. [10]

Popper, K. (1976). *Logik der Forschung*. Tübingen: Mohr. [44]

Popper, K. (1983). Freeman on Peirce's Anticipations of Popper. In E. Freeman (Ed.), *The Relevance of Charles Peirce* (pp. 78–79). La Salle: Monist Library of Philosophy. [204]

Potter, V. G., and Shields, P. B. (1977). Peirce's Definitions of Continuity. *Transactions of the Charles S. Peirce Society, 13*, 20–34. [188]

Proudfoot, D. (1997). On Wittgenstein on Cognitive Science. *Philosophy, 72*, 189–217. [163]

Proudfoot, D. (2004). Robots and Rule–Following. In C. Teuscher (Ed.), *Alan Turing: Life and Legacy of a Great Thinker* (pp. 359–379). New York: Springer. [52]

Proudfoot, D., and Copeland, B. J. (1994). Turing, Wittgenstein and the Science of the Mind. *Australasian Journal of Philosophy, 72*, 497–519. [163]

Putnam, H. (1955). Peirce's Continuum. In K. L. Ketner (Ed.), *Peirce and Contemporary Thought: Philosophical Inquiries* (pp. 1–22). New York: Fordham University Press. [186]

Putnam, H. (1960). Minds and Machines. In S. Hook (Ed.), *Dimensions of Mind. A Symposium* (pp. 148–179). New York: New York University Press. [65]

Putnam, H. (1975). *Mind, Language and Reality. Philosophical Papers* (Vol. II). Cambridge: Cambridge University Press. [73]

Putnam, H. (1992). *Renewing Philosophy*. Cambridge, MA: Cambridge University Press. [110]

Putnam, H. (1995). Putnam, Hilary. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 507–513). Oxford: Blackwell. [73] [238]

Quine, W. V. (1987). *Quiddities: An Intermittently Philosophical Dictionary*. Cambridge, MA: Harvard University Press. [92]

Ramsey, F. P. (1923). Critical Notice of L. Wittgenstein's Tractatus. *Mind, 32*, 465–478. [103]

Ramsey, W. (2011). Eliminative Materialism. *The Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/materialism–eliminative/ [80]

Ramsey, W., Stich, S., and Garon, J. (1990). Connectionism, Eliminativism and the Future of Folk Psychology. *Philosophical Perspectives, 4*, 499–533. [81]

Ransdell, J. (1977). Some Leading Ideas of Peirce's Semiotic. *Semiotica, 19*, 157–178. [186]

Rescher, N., and Brandom, R. (1980). *The Logic of Inconsistency: A Study in Non–Standard Possible–World Semantics and Ontology*. Oxford: Basil Blackwell. [114]

Rhees, R. (1966). Can There Be a Private Language? In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 267–285). Melbourne: Macmillan. [113]

Rhees, R. (2002). Five Topics in Conversation with Wittgenstein (Numbers; Concept–Formation; Time–Reactions; Induction; Causality). *Philosophical Investigations, 25*. [95] [160]

Richards, I. A. (1936). *The Philosophy of Rhetoric*. London & New York: Oxford University Press. [42]

Richardson, A., and Uebel, T. (2008) (Eds.), *The Cambridge Companion to Logical Empiricism*. New York: Cambridge University Press. [100]

Romdenh–Romluc, K. (2010). Immunity to Error through Misidentification. In A. Barber and R. J. Stainton (Eds.), *Concise Encyclopedia of Philosophy of Language and Linguistics* (pp. 305–308). Oxford & Boston: Elsevier. [155]

Rosenblatt, F. (1958). Immunity to Error through Misidentification. *Psychological Review, 65*, 386–408. [66]

Rosental, D. M. (1995). Identity Theories. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 348–355). Oxford: Blackwell. [50]

Rumelhart, D., McClelland, J., and The PDP Research Group (1986). *Parallel Distributed Processing. Explorations in the Microstructure of Cognition: Foundations* (Vol. I). Cambridge, MA: MIT Press. [66] [67]

Russell, B. (1905). On Denoting. *Mind, 14*, 479–493. [84]

Russell, B. (1921). *The Analysis of Mind*. London & New York: G. Allen & Unwin. [153]

Russell, B. (1954). *The Analysis of Matter*. New York: Dover Publications. [162]

Russell, B., and Whitehead, A. N. (1957). *Principia Mathematica* (Vol. I). Cambridge: Cambridge University Press. [157]

Savan, D. (1988). *An Introduction to C. S. Peirce's Full System of Semeiotic*. Toronto: Victoria College University of Toronto. [185]

Sayre, K. (1976). *Cybernetics and the Philosophy of Mind*. London: Routledge and Kegan Paul. [162]

Schauer, F. (2010). Neuroscience, Lie–detection, and the Law. *Trends in Neuroscience, 14*, 101–103. [231]

Schneider, S., and Velmans, M. (Eds.) (2007). *The Blackwell Companion to Consciousness*. Malden, MA: Blackwell. [74] [149]

Seage, W., and Bourget, D. (2007). Representationalism About Consciousness. In S. Schneider and M. Velmans (Eds.), *The Blackwell Companion to Consciousness* (pp. 261–276). Malden, MA: Blackwell. [149]

Searle, J. R. (1979). What Is an Intentional State? *Mind, 88*, 74–92. [72]

Searle, J. R. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences, 3*, 417–457. [76]

Searle, J. R. (1983). *Intentionality. An Essay in the Philosophy of Mind*. Cambridge & New York: Cambridge University Press. [77] [78] [92]

Searle, J. R. (1984). *Minds, Brains and Science*. Cambridge, MA: Harvard University Press. [76]

Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press. [52] [75] [76] [77]

Searle, J. R. (1995). Searle, John. In S. Guttenplan (Ed.), *A Companion to the Philosophy of Mind* (pp. 544–550). Oxford: Blackwell. [75]

Searle, J. R. (2000). Consciousness. *Annual Review of Neuroscience 23*, 557–578. [75]

Searle, J. R. (2002). Why I Am Not a Property Dualist. *Journal of Consciousness Studies, 9*, 57–64. [79]

Searle, J. R. (2007). Biological Naturalism. In S. Schneider and M. Velmans (Eds.), *The Blackwell Companion to Consciousness* (pp. 325–334). Malden, MA: Blackwell. [74]

Sebeok, T. A. (Ed.) (1978). *Sight, Sound, and Sense*. Bloomington: Indiana University Press. [168]

Sellars, W. (1956). Empiricism and the Philosophy of Mind. In H. Feigl and M. Scriven (Eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis: Minnesota Studies in the Philosophy of Science* (Vol. I, pp. 253–329). Minneapolis: University of Minnesota Press. [81]

Shanker, S. G. (1987). Wittgenstein versus Turing On the Nature of Church's Thesis. *Notre Dame Journal of Formal Logic, 28*, 615–649. [54] [55] [57] [60] [61]

Shanker, S. G. (1998). *Wittgenstein's Remarks on the Foundations of AI*. London & New York: Routledge. [61]

Shapiro, L. (2000). Multiple Realizations. *Journal of Philosophy, 97*, 635–654. [74]

Shapiro, S. (1998). Incompleteness, Mechanism, and Optimism. *The Bulletin of Symbolic Logic, 4*, 273–302. [65]

Shields, C. (Ed.) (2003). *The Blackwell Guide to Ancient Philosophy*. Oxford: Wiley–Blackwell. [20]

Shoemaker, D. (2012). Personal Identity and Ethics. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/entries/identity–ethics/ [26]

Shoemaker, S. (1984). Identity, Cause, and Mind. Cambridge: Clarendon Press. [26]

Short, T. L. (2006). The Development of Peirce's Theory of Signs. In C. Misak (Ed.), *The Cambridge Companion to Peirce* (pp. 214–240). Cambridge: Cambridge University Press. [182]

Short, T. L. (2007). *Peirce's Theory of Signs*. Cambridge & New York: Cambridge University Press. [168] [172] [174] [199]

Skagestad, P. (1999). Peirce's Inkstand as an External Embodiment of Mind. *Transactions of the Charles S. Peirce Society, 35*, 551–561. [204]

Skinner, B. F. (1974). *About Behaviorism*. New York: Vintage Books. [45]

Skinner, B. F. (1977). Why I Am Not A Cognitive Psychologist. *Behaviorism, 5*, 1–10. [45]

Smart, J. C. C. (1959). Sensations and Brain Processes. *Philosophical Review, 68*, 141–156. [47] [71]

Smart, J. C. C. (2011). The Mind/Brain Identity Theory. *The Stanford Encyclopedia of Philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2011/entries/mind–identity/ [47] [48] [49] [71]

Smith, B. (Ed.) (2003). *John Searle. Contemporary Philosophy in Focus*. Cambridge: Cambridge University Press. [77] [79]

Smolensky, P. (1987). The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn. *The Southern Journal of Philosophy, 26*, 137–161. [66]

Smolensky, P. (1991). Reply: Constituent Structure and Explanation in an Integrated Connectionist/Symbolic Cognitive Architecture. In C. Macdonald and G. Macdonald (Eds.), *Connectionism: Debates on Psychological Explanation* (pp. 223–290). Oxford: Wiley–Blackwell. [70]

Snowdon, P. (2011). Private Experience and Sense Data. In O. Kuusela and M. McGinn (Eds.), *The Oxford Handbook of Wittgenstein* (pp. 402–428). Oxford: Oxford University Press. [138]

Sober, E. (2008). Teleological Functionalism. Putting the Function Back into Functionalism. In W. G. Lycan (Ed.), *Mind and Cognition. A Reader* (pp. 97–106). Oxford: Blackwell. [71]

Sorabji, R. (1974). Body and Soul in Aristotle. *Philosophy, 49*, 63–89. [11]

Sorabji, R. (2001). Aristotle on Sensory Processes and Intentionality: A Reply to Burnyeat. In D. Perler (Ed.), *Ancient and Medieval Theories of Intentionality* (pp. 49–61). Leiden: Brill. [11]

Sorell, T. (2000). Descartes. In W. H. Newton–Smith (Ed.), *A Companion to the Philosophy of Science* (pp. 79–84). Oxford: Blackwell. [20]

Spinoza, B. (1846). *Opera quae supersunt omnia. Principia philosophiæ, Cogitata metaphysica, Ethica* (Vol. I.). Leipzig: Bernh. [23]

Stenius, E. (1964). *Wittgenstein's Tractatus: A Critical Exposition of Its Main Lines of Thought*. Ithaca: Cornell University Press. [97] [221] [233]

Stenius, E. (1967). Mood and Language–Game. *Synthese, 17*, 254–274. [138]

Stern, D. G. (2004). *Wittgenstein's Philosophical Investigations: An Introduction*. Cambridge: Cambridge University Press. [113]

Stern, D. G. (2008). Wittgenstein, the Vienna Circle, and Physicalism. In A. Richardson and T. Uebel (Eds.), *The Cambridge Companion to Logical Empiricism* (pp. 305–331). New York: Cambridge University Press. [100]

Stern, D. G. (2010). Tracing the Development of Wittgenstein's Writing on Private Language. In Venturhina, N. (Ed.), *Wittgenstein After His Nachlass* (pp. 110–127). New York: Palgrave Macmillan. [112]

Stern, D. G. (2011). Tracing the Development of Wittgenstein's Writing on Private Language. In O. Kuusela and M. McGinn (Eds.), *The Oxford Handbook of Wittgenstein* (pp. 333–350). Oxford: Oxford University Press. [112]

Stich, S. (1984). Is Behaviorism Vacuous? *Behavioral and Brain Sciences, 7*, 615–667. [45]

Stich, S. (1996). *Deconstructing the Mind*. New York: Oxford University Press. [82]

Stocker, B. (Ed.) (2004). *Post–Analytic Tractatus*. Aldershot: Routledge. [112] [113]

Strawson, P. F. (1964). *Individuals: An Essay in Descriptive Metaphysics*. London: Routledge. [234]

Strawson, P. F. (1966). Review of Wittgenstein's Philosophical Investigations. In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 22–64). Melbourne: Macmillan. [113]

Stroud, B. (1977). *Hume*. London: Routledge & Paul Kegan. [28]

Stump, E. (1995). Non–Cartesian Substance Dualism and Materialism without Reductionism. *Faith and Reason, 12*, 505–531. [17]

Swinburne, R. (1997). *The Evolution of the Soul*. Oxford: Oxford University Press. [17]

Teuscher, C. (Ed.) (2004). *Alan Turing: Life and Legacy of a Great Thinker*. New York: Springer. [52]

Thompson, M. (1984). Hartshorne and Peirce: Individuals and Continuity. In J. B. Cobb and F. I. Gamwell (Eds.), *Existence and Actuality: Conversations with Charles Hartshone*. Chicago & London: University of Chicago Press. [199]

Traiger, S. (Ed.) (2005). *The Blackwell Guide to Hume's A Treatise of Human Nature*. Oxford: Blackwell. [27]

Turing, A. M. (1965). On Computable Numbers, with an Application to the Entscheidungsproblem. In M. Davis (Ed.), *The Undecidable. Basic Papers On Undecidable Propositions, Unsolvable Problems And Computable Functions* (pp. 116–153). New York: Dover Publications. [54]

Turing, A. M. (1969). 'Intelligent Machinery'. National Physical Laboratory Report. In B. Meltzer and D. Michie (Eds.), *Machine Intelligence* (Vol. 5, pp. 3–23). Edinburgh: Edinburgh University Press. [61] [66]

Turing, A. M. (1986). Lecture to the London Mathematical Society on 20 February 1947. In M. Woodger (Ed.), *A. M. Turing's Ace Report of 1946 and Other Papers* (pp. 106–124). Cambridge, MA: MIT Press. [58] [63]

van Gulik, R. (2008). Functionalism, Information and Content. In W. G. Lycan (Ed.), *Mind and Cognition. A Reader* (pp. 107–129). Oxford: Blackwell. [71]

van Heijenoort, J. (1967). Logic as Calculus and Logic as Language. *Synthese, 17*, 324–330. [96]

Venturhina, N. (Ed.) (2010). *Wittgenstein After His Nachlass*. New York: Palgrave Macmillan. [112]

Vericat, J. F. (1994). Color as Abstraction. In G. Debrock and M. Hulswit (Eds.), *Living Doubt. Essays Concerning the Epistemology of Charles Sanders Peirce*. Dordrecht: Springer. [180]

von Neumann, J., and Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press. [152] [161]

von Wright, G. H. (1980). *Wittgenstein*. Oxford: Blackwell. [153]

Wang, H. (1974). *From Mathematics to Philosophy*. London: Routledge. [59]

Watson, J. (1930). *Behaviorism*. New York: Norton. [44]

Webb, J. C. (1968). Metamathematics and the Philosophy of Mind. *Philosophy of Science, 35*, 156–178. [64]

Webb, J. C. (1980). *Mechanism, Mentalism, and Metamathematics. An Essay on Finitism*. Boston: Springer. [59] [60]

Wedin, M. (1995). Keeping the Matter in Mind: Aristotle on the Passions and the Soul. *Pacific Philosophical Quarterly, 76*, 183–221. [15]

Weisstein, E. W. (1999). *CRC Concise Encyclopedia of Mathematics*. Boca Raton, London, New York, Washington: CRC Press. [42]

Wetzel, L. (2008). *Types and Tokens: An Essay on Universals*. Cambridge, MA: MIT Press. [92]

Whitney, W. D. (Ed.) (1889–1891). *The Century Dictionary and Cyclopedia* (Vols. I–X). New York: W. G. Blackie and Co. [197]

Wilkes, K. V. (1995). Psuchê vs. Mind. In M. C. Nussbaum and A. O. Rorty (Eds.), *Essays on Aristote's De Anima* (pp. 109–128). Oxford: Oxford University Press. [11]

Williams, M. (1999). *Wittgenstein, Mind and Meaning. Toward a Social Conception of Mind*. London & New York: Routledge. [112] [145]

Witherspoon, E. (2011). Wittgenstein on Criteria and the Problem of Other Minds. In O. Kuusela and M. McGinn (Eds.), *The Oxford Handbook of Wittgenstein* (pp. 472–498). Oxford: Oxford University Press. [164]

Wittgenstein, L. (1958). *The Blue and the Brown Book* (R. Rhees Ed.). Oxford: Blackwell. [53] [69] [109] [110] [134] [213]

Wittgenstein, L. (1969). *On Certainty* (D. Paul, G. E. M. Anscombe, Trans.). Oxford: Basil Blackwell. [58] [109] [130] [135] [136] [137] [140] [141] [142] [144]

Wittgenstein, L. (1974a). *Tractatus Logico–Philosophicus* (C. K. Ogden, Trans.). London: Routledge and Kegan Paul. [97] [102] [103] [131] [157]

Wittgenstein, L. (1974b). *Philosophical Grammar* (A. Kenny, Trans. R. Rhees Ed.). Oxford: Basil Blackwell. [53] [69] [107] [159]

Wittgenstein, L. (1975). *Philosophical Remarks* (R. Hargreaves and R. White, Trans. R. Rhees Ed.). Oxford: Basil Blackwell. [77] [103] [133] [168]

Wittgenstein, L. (1978). *Lectures and Conversations on Aesthetics, Psychology and Religious Belief* (C. Barrett Ed.). Oxford: Blackwell. [105]

Wittgenstein, L. (1979a). *Ludwig Wittgenstein and the Vienna Circle* (B. F. McGuinness Ed.). Oxford: Blackwell. [97] [99] [159]

Wittgenstein, L. (1979b). *Wittgenstein's Lectures: Cambridge, 1932–1935. From the Notes of A. Ambrose and M. MacDonald* (A. Ambrose Ed.). Oxford: Blackwell. [158]

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology, Volume 1* (G. E. M. Anscombe, Trans. G. H. von Wright and G. E. M. Anscombe Eds.). Chicago: Chicago University Press. [52] [99] [130] [134] [141] [154]

Wittgenstein, L. (1981). *Zettel* (G. E. M. Anscombe, Trans. G. E. M. Anscombe and G. H. von Wright Eds.). Oxford: Blackwell. [218]

Wittgenstein, L. (1983). *Remarks on the Foundation of Mathematics* (G. E. M. Anscombe, Trans. G. H. von Wright and G. E. M. Anscombe Eds.). Cambridge, MA: MIT Press. [55] [56] [57] [60] [61] [114]

Wittgenstein, L. (1988). *Wittgenstein's Lectures on Philosophical Psychology 1946–1947* (P. Geach Ed.). Hassocks: Harvester Press. [143] [154]

Wittgenstein, L. (1992). *Last Writings on the Philosophy of Psychology: 1949–1951. Volumes I and II* (C. G. Luckhardt, M. A. E. Aue, Trans. G. H. von Wright, H. Nyman Eds.). Oxford: Blackwell. [144]

Wittgenstein, L. (1993). *Philosophical Occasions, 1912–1951* (J. Klagge, A. Nordman Eds.). Indianapolis and Cambridge: Hackett Publishing. [139]

Wittgenstein, L. (1995). *Ludwig Wittgenstein: Cambridge Letters* (B. F. McGuinness, G. H. von Wright Eds.). Oxford: Blackwell. [44]

Wittgenstein, L. (1998). *Culture and Value* (P. Winch, Trans. G. H. von Wright Ed.). Oxford: Blackwell. [146]

Wittgenstein, L. (2000). *Wittgenstein's Nachlass. The Bergen Electronic Edition*. Oxford: Oxford University Press. [52] [98] [106] [142] [144] [146] [160]

Wittgenstein, L. (2001). *Philosophical Investigations* (G. E. M. Anscombe, Trans. G. E. M. Anscombe and R. Rhees Eds.). Oxford: Blackwell. [60] [87] [94] [100] [104] [105] [106] [108] [109] [110] [113] [119] [122] [123] [124] [125] [126] [127] [128] [129] [130] [132] [133] [134] [153] [154] [155] [213] [218] [220] [221] [231]

Wittgenstein, L. (2009). *Philosophical Investigations. The German Text, with an English Translation* (G. E. M. Anscombe, P. M. S. Hacker and J. Schulte, Trans. P. M. S. Hacker and J. Schulte Eds.). Oxford: Wiley–Blackwell. [106] [120] [123] [126]

Yalowitz, S. (1997). Rationality and the Argument for Anomalous Monism. *Philosophical Studies, 83*, 235–258. [90]

INDEX OF NAMES