Università degli Studi di Cagliari

PhD DEGREE

Doctoral Program in

Epistemology, Philosophy and History of Culture

University of Cagliari

XXXIII Cycle

2017-2021

Coordinator of PhD Program Prof. GIUSEPPE SERGIOLI

A SITUATED AND TIMING-BASED APPROACH TO CONSCIOUSNESS IN LEARNING

Scientific Disciplinary Sector

M-FIL/02

Candidate: FABRIZIA GIULIA GARAVAGLIA

Supervisor: Prof. MARCO GIUNTI

Final exam. Academic Year 2019 – 2020

Thesis defence: July 2021 Session

“Each soul is a perpetual living mirror of the universe”

G.W. Leibniz - The Monadology, 1714

Abstract

What happens to us when we learn something new? How is consciousness of our agency linked with this process? This work investigates the notion of consciousness by focusing on the process of individual learning from two perspectives. On one side, the role played by intentionality and intersubjectivity in individual action. On the other one, the role played by motor conditioning. I distinguish between two different types of learning. The first, standard and voluntary learning, proceeds between intention, imitation (intersubjective) and consolidation of neural structures. The second, recursive, proceeds from consolidation of the motor side to promote some neural circuits, consolidating effective motor experiences among a wider range of ineffective experiences. The goal is to compare the level of consciousness associated with them, from a neural point of view and from a phenomenological point of view. I will also explore whether an interaction between neurophysiological and philosophical approaches could contribute to reinterpret some data, in order to have a broader notion of consciousness that could be associated with these learning processes. I will argue that an analysis of the timing processes involved in this kind of mental activity could help to shed a new light on these processes and could help to connect neural subjective activity to phenomenological experience, on one hand, and to intersubjective reality, on the other hand. Starting from an extended and dynamical approach (4E type) that underlines the relevance of the environment-system, the thesis proposes an explanatory paradigm focused on some aspects of the timing of signals that would characterize, in my opinion, some of the cognitive faculties most involved in learning, such as perception and memory. I therefore propose a redefinition of these faculties not locally focused on portions of space, but on their temporal architecture. I argue that this timing architecture offers a way to connect some outcomes of neurophysiological research with phenomenological aspects and helps to compare the awareness of agency in these two opposite kinds of learning processes. These studies also have significant ethical implications and practical applications. The intersubjective dimension and motor conditioning play an important role in learning and cognition, and also complicate the notion of personal identity, voluntary action and causation in agency.

Introduction

Cap 6 Self-generated memory: when the boundaries between what happened and what is happening collapse

Cap 7 Change in timing information. Learning and conscious agency

Introduction

This work aims to investigate the notion of consciousness by focusing on the process of individual learning from two points of view. On the one hand, the role played by intentionality and intersubjectivity in individuals' actions. On the other hand, the role played by motor conditioning. It is my contention that, to describe and explain the process of learning it is crucial to adopt a broader notion of consciousness, not limited to the subject's neuro-systemic activity. It is important to distinguish between different types of learning and to compare the notions of consciousness that could be associated with them, both from a neural point of view and from a phenomenological point of view.

Starting from an analysis of neural structure and its changes that seem to support learning processes, I will show that, in order to understand what happens in learning, we need a broader account of the notion of perception and consciousness. In particular, a notion linked to timing activity seems the most promising. Temporal structures of impulses and the complex spacetime architectures in which they are organized do not exclusively depend on the typology of their processing based on spatial location in the neural system. From this point of view, I analyze two different kinds of learning.

The first one is *standard or voluntary learning*, a process that proceeds from a voluntary intention to achieve a particular skill and involves repetition, imitation, intersubjective interaction, and parallel consolidation of neural structures.

The second is *afferent or recursive learning*, a process that proceeds in the opposite direction. It starts from motor consolidation to weaken certain neural networks, and it thus consolidates effective motor experiences between a wider range of ineffective experiences. In particular, I refer to some rehabilitation processes. Some orthoses allow a subject to carry out a motor task only according to certain degrees of freedom. This way, they are able to replace brain selection, and reeducate subjects with brain injuries to perform adequate actions.

From a phenomenological point of view, I will explore how subjects experience the resulting motor actions of these different kinds of learning. Among these aspects, the most interesting one concerns invisible, private and subjective learning, and it involves conscious and voluntary aspects, as well as unconscious and involuntary ones. The introjection of reality in a subjective and unique perspectival sphere generates an appropriate reaction. I think that an analysis of the timing process involved in this kind of mental activity could help to shed light on these processes and could connect neural activity to phenomenological experience, on the one hand, and to intersubjective reality, on the other one.

These studies also have significant ethical implications. Intersubjective aspects and motor conditioning play an important role in learning and cognition, and they strongly complicate the notion of personal identity, voluntary action, and responsibility.

Thematic distribution

As soon as we start a thorough investigation of learning processes, three main questions naturally arise.

1. What happens in our brain and in our neural system? This is the neurophysiological side
2. What happens to us? How do we experience the learning process from a subjective point of view? This is the phenomenological side.
3. What theories of consciousness are mostly relevant? What are the main notions of consciousness available from research in neurophysiology or philosophy? Is it possible to take advantage of some philosophical insights on consciousness to better interpret learning processes and some empirical findings about them?

In order to approach these three questions, my dissertation starts with a review of the current state of the art, and more specifically, of some fundamental concepts such as memory, learning, and neural communication. I think this is useful in order to be able to define the scope of this work. Its main target is to deepen the study of learning by taking into account both a physiological and a philosophical perspective, so as to find mutually useful aspects of these different approaches and to improve the interpretation of relevant scientific evidence. At the same time, it is also important to consider anatomical and functional studies of those areas of the neural system that seem to be more closely related to the functions of learning, memory, and perception. More specifically, some pathways that are crucial for perception, self-perception, and perception of time should be considered.

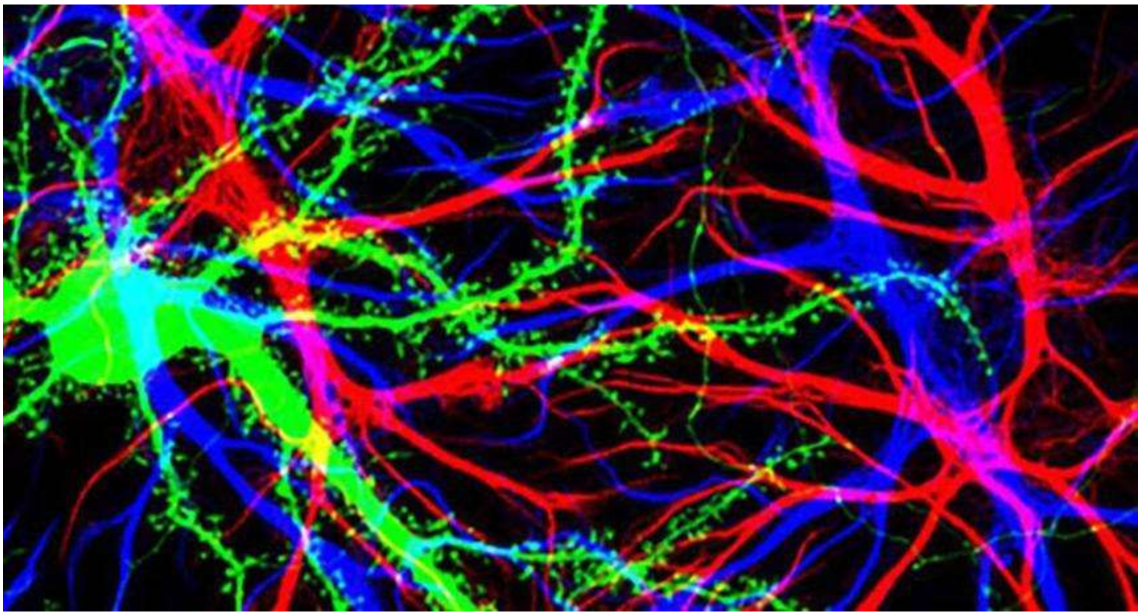
To this end, specific research concerning cerebral anatomy and physiology will be reviewed first, and then the most influential models of explanation in computational neuroscience will be accounted for. Some problematic aspects of these models will be pointed out, as well as possible ways out. We will then go into the properly philosophical side. The learning process and the notion of consciousness will be explored from the point of view of some contemporary philosophical accounts. Starting with the relevance of the role of the environment, we will compare theories that associate consciousness with a level of functional integration and theories that interpret the data in a different way. In order to do this, it will be necessary to investigate in more detail those cases in which, in a more or less controlled way, we intervene on the level of consciousness of a subject through the use of anesthetics or other drugs like psychedelics.

Having laid down the necessary foundations, I will then start building my theoretical proposal, which focuses on the relevance of spacetime architectures in the emergence of conscious experience. In particular, I will focus on timing aspects of synchronization, entrainment and coherence between certain neural structures and environment. I will refer to those phenomenological perspectives which consider perception as a multidimensional and extended process to motivate the use of an extended and timing-based interpretation of the evidence previously discussed. I surmise that this new dynamic paradigm can provide a deeper and more comprehensive interpretation of neural phenomena, and could be helpful to solve some of their problematic aspects. Preliminary results will then be evaluated and, in the last part of this work, possible applications of these reflections to improve learning processes will be considered.

Chapter 1

Learning from a neurophysiological point of view

Neural plasticity features and the role of the environment



Learning is a very rich notion in the wide field of cognitive science. In the present work the term "learning" is not used to solely mean formation of memories and regular detection. Learning is instead meant to be, more profoundly, a process of construction of possibilities of action, which involves the main faculties of the human mind, like perception, memory and imagination. There are many different levels of description of this process. In the present work I approach at first the neurophysiological level of explanation and I then address the phenomenological level. Finally, I will reconsider some observations concerning these two levels through the filter of timing processing.

What happens in our brain when we learn something new? What are the relevant features that we select in our environment and what is its role in our neural structure? These are the main questions that this chapter will try to address.

Learning, from a neurophysiological point of view, is a process of *remapping* of neural configurations that occurs thanks to neural plasticity**,** or ability to modify synaptic junctions. In the following part I will talk briefly about how this mechanism works, how neural plasticity shapes the whole neural system since our birth (during the development of the brain though the process of migration) and also during our entire life.

In Sect. 1.1.1, I will talk about the relevance of different aspects of the environment's influence on neural plasticity during the development of the brain.

In particular, I will investigate the role of the environment, from the dimension of cellular environment, those structures that surround neurons, up to a much broader dimension, the role of parents' behavior in pregnancy.

In Sect. 1.1.2, I will talk about the relevance of the environment in neural change throughout life will. I will focus on specific and more narrow aspects, namely, the change which characterizes learning processes, by examining the notion of evaluation of efficacy.

The target of this section is to highlight the active role of the environment in both these processes, its impact on the individual system, and how it is implemented. I intend to approach here the notion of plasticity from a morphological and situated point of view. In Ch. 7, I will talk more specifically of timing and the relevance of a timing-based perspective in neural changing that occurs in the brain during learning processes, and I will try to propose a paradigm of explanation that, taking into account the active role of the notion of synchronicity and coherence, could help to shed light on the differences between conditioned and voluntary learning and the different conscious perception associated to them.

The basic concept of new neuroscience is that over time the brain is always willing to reform and change. The brain, like life, is not a static system, but a process of self-creation known as *autopoiesis*. It was the great Polish neuroscientist Jerzy Konorski (1948) who used the term *plasticity* in 1948 to describe brain changes, which are due to the strength of connection between neurons expressed by the influence of experience. Previously, Ramòn y Cajal (1894) had argued that the ability of neurons to mature and their power to create new connections can explain learning.

In the early 1950s, several studies had also shown that "repeated administrations of an electrical stimulus to a nerve pathway were able to alter the synaptic transmission in that way generating a neural plasticity, that could be associated to learning improvement and emotional subjective responses” (Le Doux 2003)[1].

I cite this sentence that exposes a strong claim of the author, because it contains the most relevant elements that we are going to deal with. The possible connections between neural processes, learning and conscious perception is one the most complicated themes debated in the philosophy of mind. I will expose, in the following chapters, many different researches and theories that explore the problems involved in this connection, and I will try to explore it through a new situated and timing-based paradigm.

The work of Larrabee and Bronk (1947) and the one by Lloyd (1949) were also relevant in shedding light on neural plasticity. Thompson and Spencer (1966) found evidence that synaptic modifications could explain learning. John Eccles studied changes in the synaptic activation to motivate an interactionist dualism neurobiological perspective, which he developed together with Karl Popper in *The Self and Its Brain* (Popper and Eccles 1994)[2].

1.1 Neural plasticity during brain development

During development, after a period of intense proliferation, nerve cells emit extensions, the axons, which navigate under the action of recognition molecules, heading to a certain region and coming into contact with other cells.

Through the process of *migration* neurons develop and reach their final position in the different levels in the neural tissue assuming their morphology and function.

[1] Le Doux offers a fertile account based on an accurate analysis of neurophysiological processes in order to explain the construction of the self and the main features that we can ascribe to subjectivity. But his assumptions are limited to a descriptive level, so that they cannot have the causal powers that they are supposed to have in his account.

[2] This interesting perspective is carried out through a dialogue between the two authors, who argue for the existence of three different worlds, each one irreducible to the others. Eccles then describes the existence of functional modules, some closed and some open. The latter ones are able to share contents with other modules. This claim will be the base of global workspace theory, which will be explored in Ch. 3.

Morphological neural differentiation occurs during the development of the axon migration process (stretching and reconnaissance) to target neural cells that have receptors (NMDA) on their synaptic membrane that "approve" such link and is conditioned by the position they occupy within the brain stratification and neighbouring cells (Le Doux 2002).

The brain configurations that originate in this way are only partly genetic, since we do not have enough genes to codify for all the synaptic junctions we possess (this argument already provides a solid basis for a counterargument of a rigid materialistic innatism).

There is an active role of glia cells in structuring channels for neural migration. This is a very recent discovery because the protagonists of neural functions used to be associated only with neural structures. But now a more relevant role is attributed to the glial cells, and this demonstrates another relevant role of the environment, in this case of the neural cell and its surroundings.

These kind of glial cells are important for the nutrition and support of all neural structures, but they are also truly relevant in structuring channels and pathways during the development of the brain that neural cells go through in the process of migration.

Also, at a broader level we can consider the active role of environment: the context, the situation in which the subject is, impacts on the activity of glial cells and neurons through epigenetics processes (as we will see in next section), by modulating their gene expressions.

At this point electrical activity, partly intrinsic and partly driven by stimuli of the environment, becomes essential to strengthen some connections at the expense of others and thus build that fine architecture that allows our brains the most sophisticated performance. Functional and morphological adaptation driven by electrical activity is exposed briefly in the scheme in Fig 1.

This extreme ability to grow and shape the connections typical of development persists for the rest of life, albeit to a more limited extent. It consists of structural changes that underlie the processes of learning and memory, both motor and cognitive, and the processes of adaptation of the organism to changing environmental conditions: activity-dependent mechanisms that govern the formation and elimination of synapses.

One of the most important processes of change in the spacetime architecture is *pruning*. This process cut the overabundant neural structure to sculpt the neural architecture.

Many major mental illnesses start to emerge in adolescence and may be caused by aberrant synaptic pruning. Although we are only beginning to unravel the ramifications of synaptic pruning in the human brain, this process clearly has significant consequences for normal human brain function and may provide key insights into the causes of some devastating and mysterious neuropsychiatric diseases.

Modifications of this process seems to be involved in some cognitive diseases that concern learning and social abilities, like autism and schizophrenia. Disordered synaptic pruning could, in fact, explain the age of onset of schizophrenia; a group of researchers (Mc Carrol et al. 2016) published genetic and experimental evidence supporting this association. While schizophrenia seems to be characterized by over-effective pruning, autism disorders seem to be connected to a lack of effective pruning.

Coinciding with specific patterns of pre-synaptic and post-synaptic activity, functional changes in the effectiveness of synaptic transmission, known as enhancement (LTP) or long-term depression (LTD), are induced. These phenomena are associated with changes in gene expression in the neurons involved, followed by structural remodelling of connections: the number of contacts between neural cells increases as a result of LTP, decreases as a result of LTD.

Neural plasticity is believed to initiate initial changes in the effectiveness of synaptic transmission, induced by precise patterns of nerve activity. These functional modifications are followed by structural remodelling processes that lead to the new training or elimination of connections.

According to a theory put forward by Donald Hebb (1949)[3], synchronous activation of pre-post synaptic neurons induces a strengthening of connections, while the activity that is mismatched over time tends to reduce their effectiveness.

[3] This work is very important, not only for the study of neural plasticity, but also because it opens the possibility of a treatment of timing windows of activations of neural units, which will be central in the part on timing of the present work (see Ch. 7).

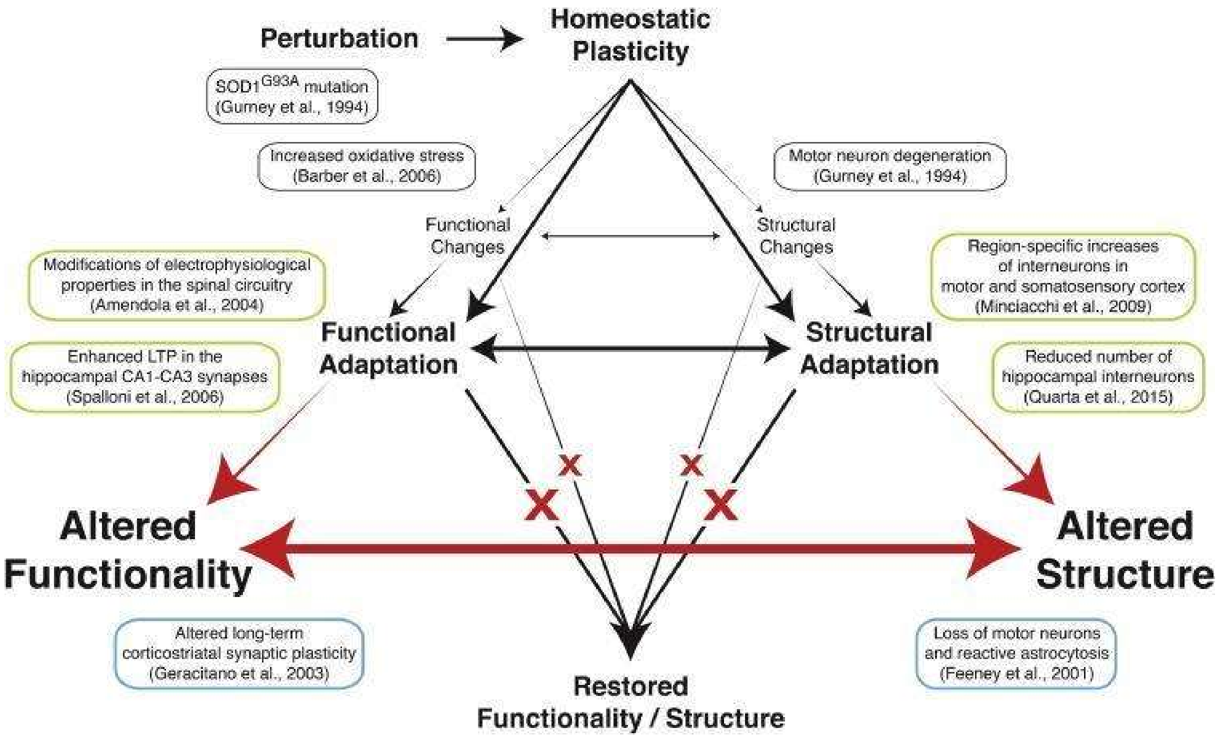


Figure 1 The scheme describes the difference between functional and structural adaptation and their main features.

1.1.1 Epigenetic factors: ecosystems and growth

Environment and experience in neural networks are very relevant during development of the brain, particularly because configurations that originate during the process of migration are only partly genetic, because we do not have enough genes that encode for all the synaptic junctions we possess.

The environment can modify function, more precisely it can amplify the activity of a gene or reduce it through biochemical mechanisms: methylation reduces the activity of a gene, acetylation amplifies it. Phenotypic aspects could be derived from previous cells without changings in genetic sequences.

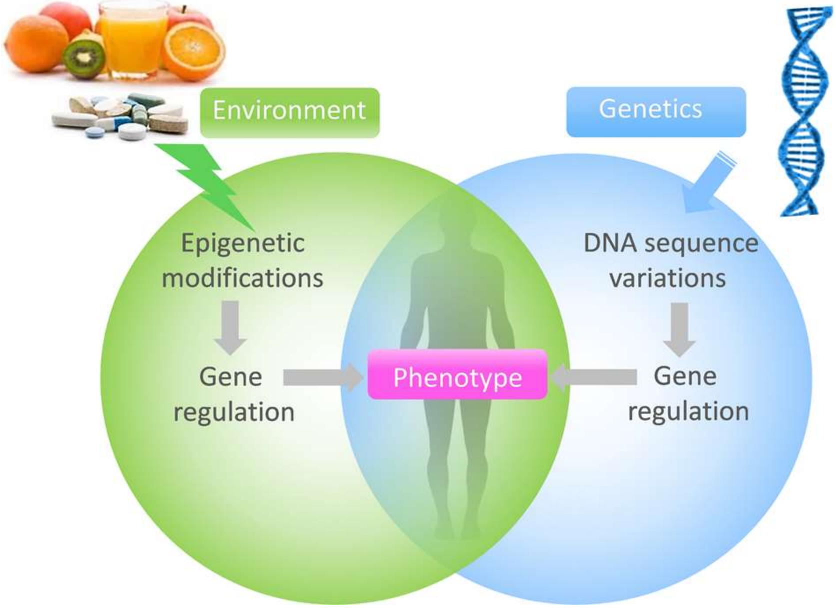


Figure 2 Phenotype is the effect of gene regulation that occurs for both genetics and epigenetics factors

Epigenetics is the research field that focuses on these processes and lots of recent studies investigate the deep meaning of this. This in fact implies a revolutionary perspective that has a strong impact on the models and methodology of neural investigation. One of the main claims of epigenetics is that there is a sort of footprint, or more precisely an epigenetic mark, which changes the activity of either a gene or a group of them, and this comes from the past of the subject. The molecular response to a specific situation could be responsible for the degree and way of activation of a specific gene. This means that, through epigenetic mechanisms it is possible to modify the activity of genes, giving them the input to generate one or more of those fragments that I mentioned before, the epigenetic marks, which could be between 250 and 400 thousand, an extraordinary number: the genes can therefore be regulated in this way.

Even if the complexity of the molecular processes that let this happen, is the root of a strong debate on this kind of explanation, the epigenetic impact is evident in many situations. In my opinion it is also a very interesting topic for the purpose of the present analysis, because phenotypic aspects, under this perspective, are one side of the environment, as they derive in some way from it and reflect some aspects of it. Therefore, a comprehensive understanding of subjective faculties must include an adequate analysis of this deep connection.

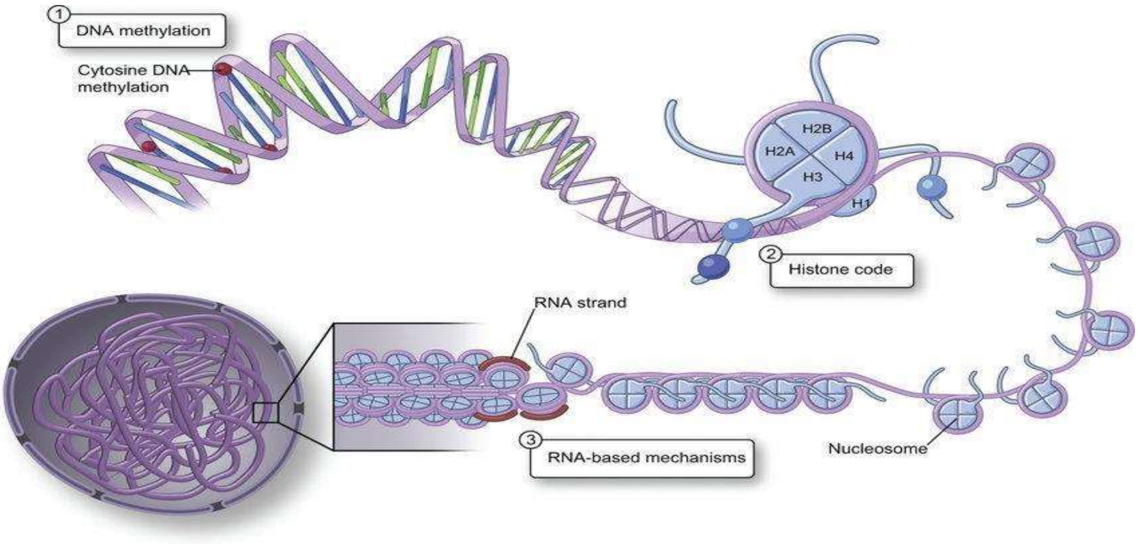


Figure 3 Methylation of DNA is an example of an epigenetic factor that marks specific sequences of activation, which in turn impact on the function of the cell.

An important factor of the environment that is crucial in the development of the brain is parents' behavior, as lots of epigenetic studies demonstrate. In particular, the mother's behavior during pregnancy plays a relevant role in structuring the development of the children's brain. The first nine months of life are very relevant to determine the fate of our entire lives. The fetus receives everything that the mother transmits, positive and negative, from food to substances that she can take, to the degree of work stress, for example through levels of glucocorticoids. (I speak of glucocorticoids because they have been marked better, but there are definitely other molecules that pass the placental barrier and reach the fetus going to genes.)

The fetus develops many cells of great variability, complexity and behavior, and the psychophysical state of the mother, for instance a depression or some other pathology, can cause alterations in these delicate codes.

Thus, during the development of the brain, experience and environment are allied with genetics to originate effective synaptic junctions and to determine learning.

1.2 Neural plasticity during life

During life neural plasticity occurs thanks to the environment too. In the adult, neural plasticity is mainly related to synapses. Here, as a result of experience there is a strengthening or weakening of the effectiveness of the transmission of the nervous impulse from one cell to another. These changes can be ranging in duration from milliseconds to months.

From a structural point of view, synapses, especially those that form on dendritic spines, can increase or decrease their surface area or vary in number. In addition, there may be variations in the number of synaptic receptors, or the molecules released into the synaptic space, like neurotransmitters.

Figure 4 shows different morphologies of dendritic spines, conjunctive structures that vary a lot during learning processes that we can observe in neural tissue after repetition of several trials of the same experience.

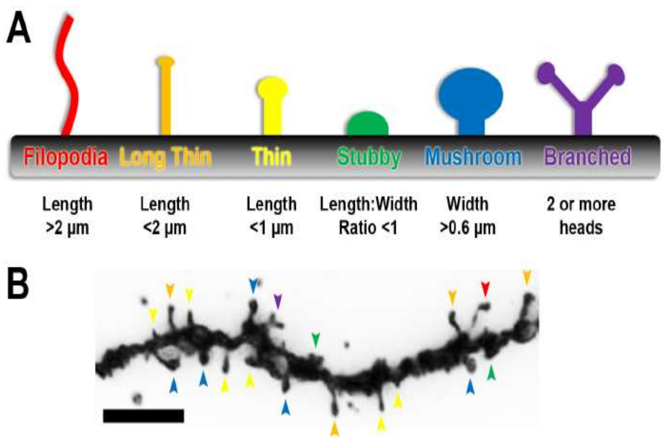


Figure 4
Dendritic spines, small neural membranous protrusions on a dendrite surface that typically receive input from a single axon. Dendritic spines contain neurotransmitters and receptors. Signalling systems essential for synaptic function and plasticity have different sizes and shapes. In this scheme we see the mainly relevant shapes. They appear in those circuits that are more used, so they are useful markers for learning effects.

Within a neurophysiological analysis we talk about neural circuits and systems. A circuit is defined as a group of neurons that connect through synaptic connections (observed from the outside to investigate certain functions).

Instead, a system is defined as a set of circuits that perform a given function. Neurophysiological investigations describe sequences of impulses and try to ascribe a functional meaning to them, indirectly connecting a class of activations to previous stimuli, occurred in a specific time window that allows for this causal explanation. I will consider in detail this aspect of causation, which methodologies of different disciplines presuppose, in Ch. 3. Here, I wish to point out that these methodologies have a problematic epistemological aspect. In fact, when we talk about a function, in neurophysiological analysis, we refer to a relation between some neural activations. This functional attribution is based on a perspective cut, which means to isolate activations (that we choose as relevant in the process that we are examining) and to consider some of them as origins or inputs and the others as responses or outputs.

In order to better understand this aspect of the enquiry in neural functions, we need to remember that the neural system is always active and works in synergy with many other systems. Therefore, the level of neural circuitry that undergoes certain functions depends on a functional and finalized delimitation of the observer. To explore a function and a target, the observer operates a cut, selecting a portion of the dynamical structure, the part that seems to be mostly involved in the process he is interested in. *Involved*, under this perspective, means that a group of neural cells is active in a time interval that can be considered causally determined by the stimulus that hit the network in the previous moment. This could be considered a coherent activation in a stream of activations that come after a specific stimulus (and in an experimental condition, stimulus structure is controlled, as the features of the output related to it). The intrinsic dynamicity of the neural system (*i.e.*, the constant activity that depends on the morphology of the neural network and the shape and type of function that each neural cell exhibits), as well as its extrinsic dynamicity (that depends on the correspondence between the structure and timing of stimuli), imposes an epistemological relativization.

The differentiated structure and morphology of neural cells, on the other hand, does not depend on a perspectival choice of the type just exposed. We can, in fact, observe and describe a lot of shapes of neural cells and their relative positions, and we can also observe the level and typology of neural tissue in which that group of cells is located. These descriptions of the features of neural cells, material connections, and neural tissue are based on static observations, and they don't depend on a perspective cut, like in the case of functional circuits, but they are objective descriptions of static components. So, we can see that the epistemological problem seems to be connected with the functional description of dynamical processes.

Thus, at this level of explanation, the diatribe concerns the weights that we can associate to the different functional units in the network that we are considering. This distribution implies a stance regarding the weight and the role of the different inputs from the environment that influence the individual network in a specific way. We could say that different neuro-computational approaches differ from each other with respect to how this distribution is construed. Computational explanations presuppose a distribution of weights on which the system computes, and these weights represent, in some way, the relevance of environment features for the individual system, in a specific context. This implies a stance about the boundaries of the subject and his identity. There are two basic neurophysiological theses in this regard, that I would like to briefly summarize here.

The first one is Neural Darwinism, the second one is strong neural Darwinism, which will be discussed later. In this model the self is not built from simple elements to form a complex. Rather, it is selected from a complex structure by the possibilities in the environment. This interesting theory of brain development during life was formulated by Gerald Edelman, in a book called *The Mindful Brain* (1978)[4]. It was then revised and extended in a subsequent famous work (1987) by the same author. The name Darwinism alludes to the fact that this theory is presented by the author as an extension of Darwin's natural selection theory. More specifically, it is an extension with respect to the process of neural development in the human system. The analogy is played on different levels and it starts from the previous studies of the author in the field of immunology. This explanation and its aspects of continuity with Darwin's theory have been subject to several criticisms which I will briefly review in Ch. 4. Here, I will just focus on one of the most interesting

[4] The theory exposed by Edelman here retraces the immunology functions of the body and applies a concept of selection and adaptation, taken from an evolutionary paradigm, to the neurophysiological functions. But this analogy especially in its extension to psychological function is nowadays very debated.

aspects of this theory, namely, its claim to be able to reconcile different levels of explanation. This is clearly shown in the next few lines taken from one of Edelman's most famous books:

> We need to be able to reconcile the structural and functional variability of the brain with the need to explain its ability to categorize. For this purpose we need a theory that has some indispensable characteristics: it must agree with the data of evolution and development; it must be coherent with the adaptive nature of responses to new situations; it must show how brain functions grow in harmony with the functions of the body, following their changes due to growth and experience; it must account for the existence of brain maps, their functionality and variability; it must explain how an integrated response can be reached from a multiplicity of maps and how perceptual responses can be generalized even in the absence of language. Finally, such a theory must account for the appearance, during evolution, of the various forms of perceptual and conceptual categorization, memory and consciousness. (Edelman 1978)

Even if it is necessary that different levels of explanation contribute to research on brain workings, neither is there a spiritual claim, nor a reification of language contents. Edelman refuses the linguists' explanation for the genesis of mental faculties. He argues that the integration of functions and responses, to be valid and correct, must also respond, without the involvement of language, to evolutionary requirements, as demonstrated by advanced neurophysiology. There must be some signs of the evolutionary process in the development of the brain, and some of them can be investigated.

Before approaching his perspective in more detail, we need to anticipate one of its crucial notions: the neural map. I will talk in more detail of the concept of neural map when I will examine later specific cognitive functions and their possible correspondence with portions of neural networks (see Ch.6). Here I will briefly introduce the concept of a neural map in order to better understand Neural Darwinism.

A neural map consists of an organized topological structure, and the activity of groups of neurons and fibers in the brain of an organism. The neural map is the functional unit of a neural network that corresponds to a function that manifests perceptual or motor output. We can talk of two hierarchical levels of neural maps. The first level, the most basic one, is the local neural map and it corresponds to an interconnection of smaller spatial and numerical portions of neurons that perform more specific and limited functions. The global neural map, more plastic, corresponds to an interconnection of different local maps.

We could think of the global map as the road network of a city that interconnects the local neighbourhood networks (local maps). The closure of a street in a neighbourhood involves the redefinition of the whole traffic at the city level.

In a nutshell, the groups neuronal selection theory is based on three main evolutionary principles:

1 The first principle refers to the fact that there is a neural selection in fetal development; on this basis, a primary repertoire of functional maps is constructed during development of the fetus, thanks to genetic and epigenetic factors.
2 The second principle refers to neural selection that depends on the animal's experiences during its existence; this provides for the construction of a secondary repertoire of functional configurations of maps and circuits.
3 The third principle, the most debated one, refers to the existence of the *re-entry mechanism*.

This means that the experience generates some individual perceptions in presence of some specific conditions of the external world. These perceptions correspond to the formation of specific signals about individual conditions that are sent back into the maps of neuronal networks and continuously contribute to update them, producing their continuous change at the level of synaptic connections.

If we consider all these principles together, we get a synthetic view of neural Darwinism theory. In the same book we have a brief and clear summary of this view:

> The whole process is based on selection and involves populations of neurons engaged in topobiological competition. A variable population of groups of neurons in a given region of the brain is defined as the primary repertoire, including networks of neurons that emerge through somatic selection processes. The genetic code does not provide a precise and detailed pattern to get to the formation of these repertoires, but rather imposes a set of constraints on the selection process. (Edelman 1978)

In order to explore more deeply the process of neural selection in the brain cells, we can refer to three fundamental assumptions of the theory:

- Redundancy: there is in every individual an overabundance of synaptic endowments
- Reinforcement: there are reinforcement processes of the synapses most commonly used
- Subtraction or degeneration: there is a removal process of unused synapses.

From this point of view the baby has an overabundance of synaptic connections; in this "synaptic cloud" every experience and learning process contributes to cut ineffective or useless ones and to reinforce the useful ones. Then, experience in the environment sculpts the subject, by selecting, among the many synaptic possibilities, only some of them and reinforcing the most used.

There is also a more rigid version of the neural Darwinism model, called *Strong Neural New-Darwinism,* which maintains that neural activity contributes to avoid neural death, which is to result in the death of those cells that do not receive input. Neural activity, therefore, in this model, preserves from death existing patterns. The focus is on the role of neural death in the structuring complexity of synaptic ramification and its contribution to the demarcation between cortical areas, and not on the reinforcement and re-entrant processes.

Neural plasticity, the changing in synaptic connections through dendritic spines, on a molecular level of explanation, occurs by simultaneous input and by the presence of neurotrophies (NGF, BNDF, N3, N4) or function stimulants and growth factors.

However, they act in an anterograde path, that means they contribute to the reinforcement of those junctions (which make up observable circuits) that have proven to have effective outputs.

So, the process develops according to the following steps: at first a stimulus occurs, an evaluation of the effectiveness of a useful output is required, the reiteration of some output in concomitance with a situation generates the reinforcement and production of neurotrophy. The result of this process is precisely what in neurophysiology is called *learning.*

More briefly, neural plasticity develops by virtue of neurotrophies, growth factors, which act in an anterograde sense: they contribute to the reinforcement of those junctions that had proven to have effective outputs. This means that there must be an integrated evaluation that judges whether an output related to a particular stimulus is better than another one with respect to the best functionality of the whole system.

This process is schematically shown in the following graph.

Stimulus → **Evaluation of efficacy** → Remapping ... Learning

Note the dots that I have inserted between remapping and learning. The reason for their insertion is the following. Even if remapping of the network through morphological changes is an observable effect, the causal connection from this to a learned outcome is not necessary, and it is the very theme of the present work.

So, even if the priority accorded to neural material connections contributes to offer a detailed description of neural phenomena, it seems lacking at the explanatory level.

1.2.1 Features of environment and complementary systems

The evaluation of efficacy appears to be a crucial aspect to have a satisfactory output in a specific context. What it means to be effective and what it means to evaluate the relevant parameters of a situation in order to perform an action is very debated, and I will take up this problem in the following sections, not only from a neurophysiological point of view, but also from a psychological point of view.

The role of hormones is another important aspect, which heavily influences neural changings since intrauterine life and during the whole subsequent life. These molecules work in synergy with some neurotransmitters to modulate the activity of neurons.

For example, the brain hemispheres of a healthy woman vary in volume during the menstrual cycle depending on hormone levels. It has been observed that hormones, both progesterone and estrogen, play an important role in regulating trophism: in women these neurons are going through cycles of hypertrophy.

Moreover, this is one of the reasons why, during pregnancy, some areas of women brain vary in dimensions as we can see in figure 5. This picture is taken from a very interesting work published on *Nature* in which the brain of a group of women in pregnancy condition was compared with non-pregnancy condition.

I chose this picture because it shows very clearly the variation that hormones and environment in a woman can cause on her neurophysiological system, and also the deep interconnection with the development of neural system of the baby, that I approached in Sect 1.1.

This variation in the brain morphology could be related to some extraordinary abilities that new mums develop and improve. For example, the auditory system becomes extremely sensitive in order to let the mother ear the baby cry, even if it is far or weak.

The increasing of alertness level, the capacity to resist sleep deprivation, the expansion of imagination to foresee potential aversive events in order to better protect the baby, the increasing of the responsibility feeling, are only some of the numerous changes that occur in the psychological world of mothers.

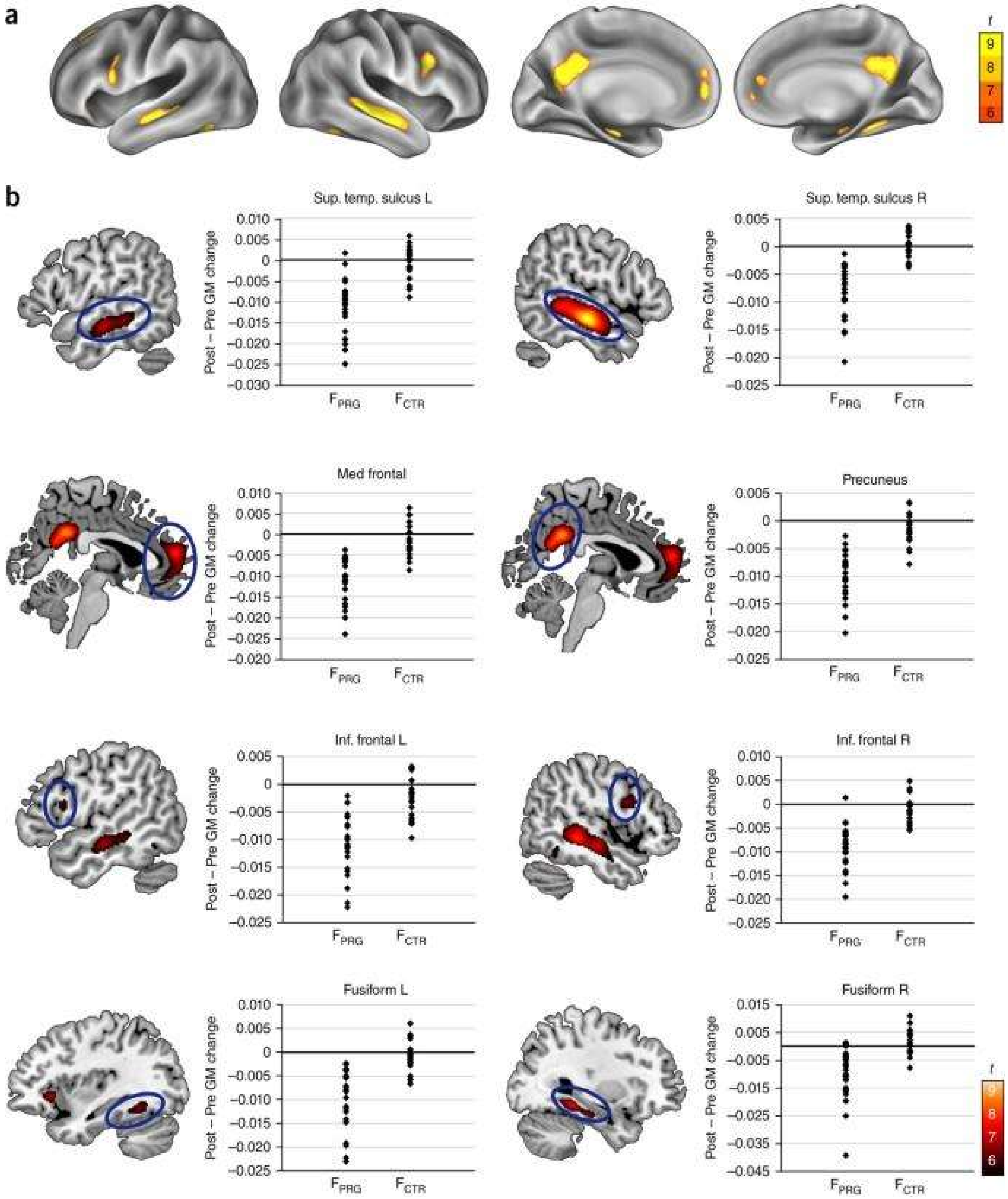


Figure 5 The dimension and of some areas vary in a visible way during pregnancy. These pictures are fMRI are taken from: Hoekzema E. et al (2017) Pregnancy leads to long-lasting changes in human brain structure *Nature Neuroscience* 20, pp. 287–296

There is another relevant system which is a complementary element to the work of the neural system. This is less studied, and it has a lot to do with pleasure and satisfaction. It is known as *endocannabinoid system,* and it refers to the trade of endocannabinoids and molecules.

Endogenous cannabinoids work in synergy with many other molecules, and especially with hormones. Endocannabinoids are lipid-based retrograde neurotransmitters that bind to cannabinoid receptors (CBRs) and cannabinoid receptor proteins that are expressed throughout the vertebrate brain and peripheral nervous system.

The endocannabinoid system seems to be involved in regulating physiological and cognitive processes, including fertility, pregnancy, pre- and post-natal development, various activities of the immune system, appetite, pain-sensation, mood, and memory.

This system is also implicated in mediating the pharmacological effects of cannabis, because cannabis contains THC and CBD, two molecules which are able to mimic endogenous elements, by replacing the link with cannabinoid receptors or remaining in the linking space of receptor for a longer period of time.

As we will see more in detail in Ch. 7, where I will explore the timing role in these processes, we will talk a lot about time experience alteration, that occurs in subjects under the assumption of these kind of substances. This affects perception of the self and the circumstances. In the environment humans selected, since their very origin, many different elements, which impact heavily on their experience of self and world.

They experience the change occurring in some relevant aspects of perception, as a consequence of the consumption of some substances, which could last for a small or very long period of time. I will go deep in this aspect talking about the role of drugs, like psychedelics or cannabinoids, on consciousness and learning processes. But here I just want to mention to highlight that there is a strong impact of the environment on the individual neural system, and this occurs by the modulation of parallel and complementary other systems. How this modulation is possible will be the topic of the next section.

1.2.2 Hippocampal trophism: neurobiology of aging and environmental factors

Hippocampus and its connections are a crucial neural structure for the topic of this work. In fact, talking about learning means also talking about memory. Hippocampus is involved in memory, location in space and time and the initial phase of action performance, so it is a truly relevant structure for the learning process. During life, as we saw earlier, our brain constantly changes its structure. Studying the neural features of age range allows us to highlight and reinforce the strong connection between synaptic changings and efficacy in learning performance. Many features characterized a lot of old subjects: decreasing of white and grey matter, hyper-activity of glial cells, changes in vascularization, decreasing of trophism of some brain areas, especially hippocampus. Alzheimer disease exhibits a dramatic volume loss in this area that is one of the most relevant roots of the symptoms.

The decrease in dimension of the brain in old age is due to three main factors: loss of neural cells, decrease of dendritic arborizations and reduction of dendritic spines.

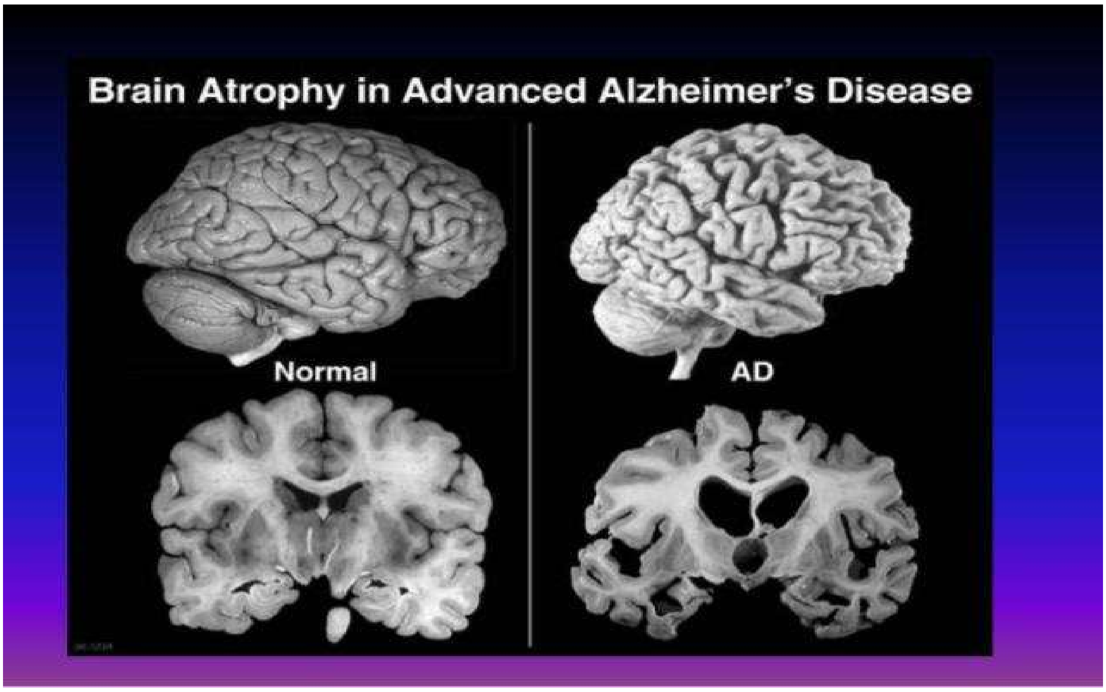


Figure 6 Representation of half dx side of the brain and, lower, a transversal slide that shows the comparison between a normal brain and the brain of a patient affected by Alzheimer Disease, with the same age. This pathology shows up with aging and becomes step by step worst. It affects the trophism of the whole brain particularly Hippocampus area.

If in the last part of our life changes in the brain trophism are physiologically normal, other factors can determine a decrease in the volume of the hippocampus even earlier in life. We talked before about the role that some kind of hormones play during our life in brain activity.

Stress is one of the most relevant environmental factors in hormones balance. Under stressful conditions the structural problem that occurs in the hippocampus could be seen as similar to what happens in the age range case. Even if the primary cause is different, the process which leads to decrease in synaptic arborization is more or less the same. In the hippocampus there are a huge number or glucocorticoid receptors. Under stressful conditions the number of these receptors sharply decreases, and so the hippocampus becomes unable to metabolize cortisol. A considerable increase of cortisol levels, which is evident also in elderly people, determines problems in the synaptic arborization and, on the behavioral level, an evident worsening of learning abilities and memory performance. All these aspects of the change in neural plasticity are well studied in post-traumatic stress disease (PSTS), and to shed light on these processes becomes relevant for treating patients successfully.

1.3 Neural plasticity and synchronicity

The role of timing and, especially, the research on synchronic aspects and coherence between sequences of neural activation, seem to be truly relevant to shed light on brain function and subjective perception. This is one of the main topics of this research. We will go deep in the instruments that neurophysiological research uses and the critical aspects that they host in the third chapter. Here I want to focus on the meaning that concept synchronicity has in neurophysiological research at synaptic level. The first meaning of this concept is linked to the timing of the environment's aspects. This means that there is a deep link between stimuli of the environment that are presented together and the architecture of subsequent neural activations in a subject. This aspect is explained in the Hebbian model.

Hebbian plasticity establishes reinforcement of synaptic junction linked to synchrony. The main idea of Hebbian synapse is that neurons that fire together wire together.

The hypothesis that the modification of synaptic transmission by experience mediates associative learning dates back to the elaboration of the concept of the synapse itself (Cajal 1894). Hebb's influential statement of the hypothesis was that if a presynaptic neuron repeatedly played a role in firing a postsynaptic neuron, there is an enduring modification which results in an improvement in the quality and timing of postsynaptic response. Here, when I talk about quality, I mean the coherence of the postsynaptic activation with the context and, when I talk about timing, I mean that there is a stronger sensibility in firing after a repeated signal under specific context conditions. In Hebbian synaptic structure, the activity in the presynaptic neuron becomes more likely to excite activity in the postsynaptic neuron, under specific conditions.

Martin et al. (2002) reviews the arguments in favor of this hypothesis, which is widely accepted by psychologists, cognitive scientists, and neuroscientists. The neurobiological process or phenomenon now most often identified with the Hebbian synapse is long-term potentiation (LTP). Recently, interest has focused on a form of LTP called spike timing–dependent plasticity (STDP); for a recent review (see Caporale and Dan 2008[5]). In a variety of neural circuits, an enduring modification of synaptic transmission is produced by varying the timing of weak and strong synaptic inputs over a range of a few tens of milliseconds. The sign of the modification depends critically on the relative strength of stimulation and the timing of the two inputs. For some parameter values, transmission increases; that is, a presynaptic spike now produces a "potentiated" (larger amplitude or shorter latency) postsynaptic response. For other combinations, transmission decreases; that is, a presynaptic spike now produces a reduced postsynaptic response. Most of the neurobiological literature on LTP focuses on its cellular and molecular mechanism. The relevance of this research to the neuroscience of learning depends on the hypothesis that links LTP to associative learning and to memory. The evidence for this link would be strong if the properties of LTP aligned closely with those of the associative learning process as revealed by behavioral experimentation. Here we review those properties and conclude

[5] They offer here a fertile interpretation of timing structures modification considering small time windows of activation

that the alignment is poor. Effects of interstimulus interval and intertrial interval behaviorally measured association formation depend on time parameters.

I will explore all these aspects at the end of next chapter. Before approaching this topic, we need to consider what synchrony means. Stimuli can be considered to have different speeds only by an external observer.

Perception is an overly complex and debated topic. Some approaches see it as a ***reconstructive*** process where it *seems* that the mind collects different inputs from its environment and reconstructs them in a meaningful universe. I will try to clarify those aspects of this topic relevant for the purpose of the present work in the first part of next chapter.

Chapter 2

Mental reconstruction of synchrony

Its role in multimodal integration

How can the mind synchronize signals, which reach the brain from different modes of reception, to form a meaningful whole? This problem, often named *Multimodal Integration*, begins with the interest in the process that stands between stimulus and response, and it refers specifically to temporal architecture of subjective perception. Stimuli have a temporal structure that differentiates them, a temporal distribution, and they impact on the spatial construction of the body that processes them depending on this aspect. The correspondence that we could construct between the shape of a stimuli, the shape of the output, and its causal dependence from temporal and spatial body aspects will be the theme of this chapter. In this part, I am going to address this question by first focusing on recent important evidence from neurosciences, and then making explicit their relevant philosophical implications.

From the phenomenological point of view, perceptual reality is a result of a complex process of decomposition and mental elaboration, which gives us a representation of an object mediated by our own potentialities. The Multimodal Integration problem owes its explicit formulation to the work of Sumby and Pollack (1954). Perceptual reality is thus a punctual representation of a stimulus that is the attentional focus of a subject, but it is not the intuitive and unconditioned apprehension of something external to the subject. Thus, the sense of continuity that accompanies the phenomenal experience of external reality and at the same time of the very sense of personal identity, stems from the re-composition of different and discontinuous signals, such as images, sounds, and somatic sensations, to produce a subjectively meaningful experience. We know for example that, physically, the sound and light waves propagate at different times and therefore at the cerebral level these inputs undergo a process of fine rehash.

This reconstruction process seems to act according to a Principle of Semantic Ascription. This basic rule could be inferred from the famous article of McGurk and MacDonald (1976), which emphasized the conditioning that the cognitive side operates on the perceptual one. Therefore, the central problem is to understand how this process of perceptual synchronization is possible in the light of this principle.

Sect. 2 discusses the structure and function of the convergence areas most involved in the synchronization process. I will elaborate on the neurophysiological mechanisms that underlie the cerebral timing of the signals and will discuss the informative power of this aspect, compared to others (for example the type and intensity of signal). I will then outline the processes of semantic ascription involved in the work of these areas, by analyzing in particular some reconstructive and illusory processes.

Sect. 3 will focus on the timing aspect, with respect to both its relevance for the coding process of signals, and its relationship with the cerebral rhythms that differentiate cognitive activities. The evidence reported in this analysis are aimed to justify perception as a reconstructive process and to emphasize the level of integration between different systems, which is necessary for the emergence of conscious experience. These tools will finally lead us to question Edelman's widely received distinction (Edelman 2000) between primary consciousness—the consciousness of the punctual present and short retention—and secondary consciousness—the consciousness of one's being present to himself and of his own perceptual acts—which rests on the primary one. That distinction is not supported by the extant evidence and, in my opinion, it is also phenomenologically unrealistic. My analysis shows that it is impossible to isolate the content of a punctual perception, because perception is always the result of an integrated and semantically driven process.

The principle of semantic ascription will be analyzed in the following two subsections by considering the functioning of multimodal areas (Sect. 2.1) and some examples of semantically reconstructive processes (Sect. 2.2).

2.1 Integration and timing. Neurophysiological observations

Until the eighties it was thought that different stimuli were processed by different cerebral parts (Unimodal regions that would compute according to a feedforward scheme), but subsequent research has developed a new systemic approach, which emphasizes how the possibility of perception is the result of an integration of many functional units that are able to overlap their functions or modify them thanks to plasticity, following a much more complex logic, which unifies in a complex dynamics simple feedforward schemes and feedback loops. Let me recall that the activity of each cell is first dictated by its position within this intricate network. Its morphology, linked to its DNA, establishes a possible range of activations in a specific system of variables and a specific temporality of response.

Temporal configurations of firing are so generated that, by synchronically acting, they can recruit different brain parts and synchronize different experiences. There are areas, Convergence Centers, or areas of Multidimensional integration, as defined by LeDoux (2003), which collect different sensory data and synchronize their responses through Re-entry circuits in the same maps they constitute. Multimodal neurons are activated if their Receptive field (receptive junction areas) are combined with different types of stimulation. What is really amazing is that this kind of integration seems to have a semantic nature. In the next section the meaning of this statement will be clarified.

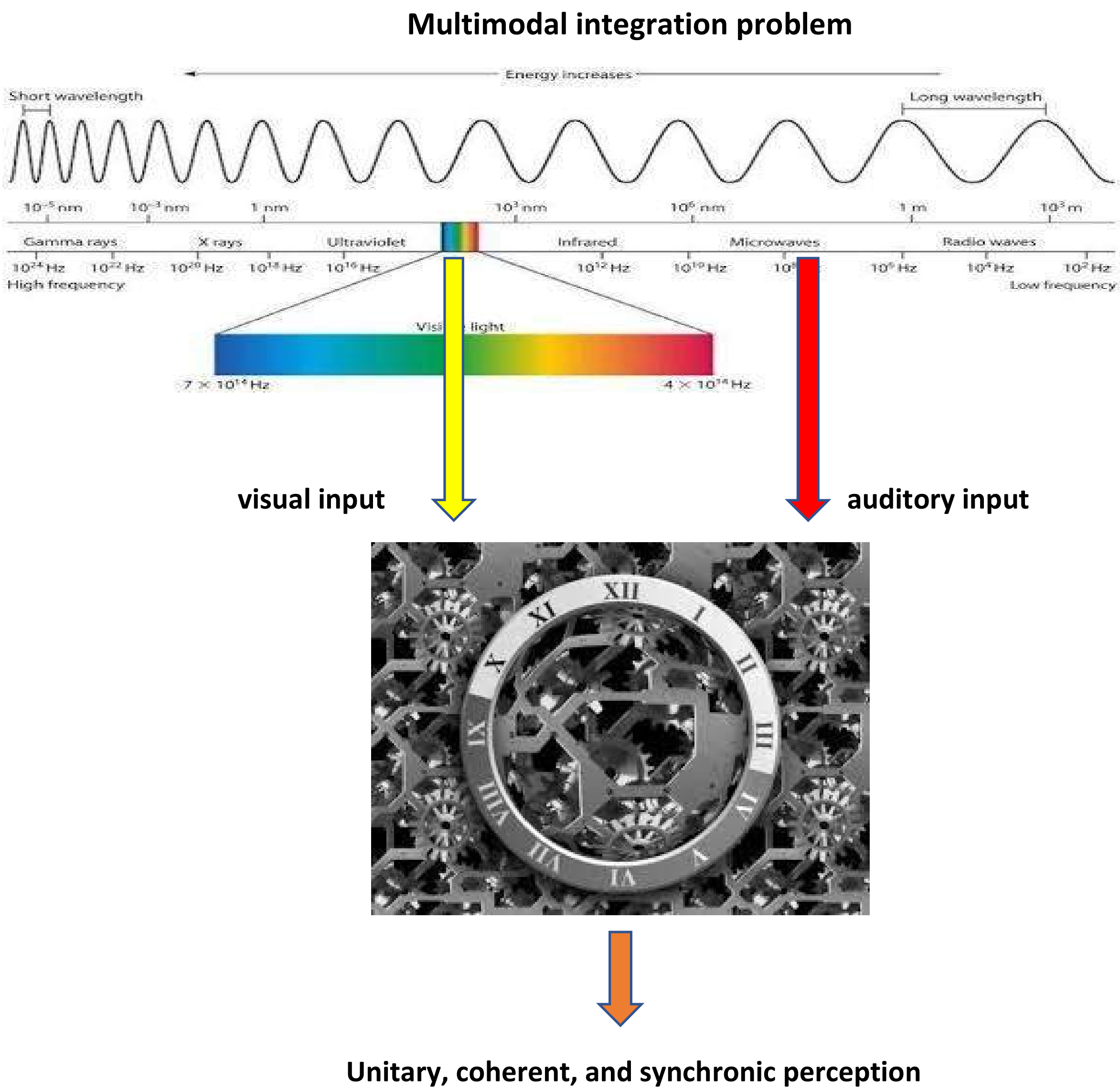


Figure 5

This scheme synthetizes the problem of multimodal integration which has two faces. First, we have different stimuli that have different features (for instance visual waves have different properties with respect to sound waves. Second, we have differences in the shape and length of our peripheral neural system, this means that it takes more time to process inputs that come from different parts of the body.

How is it possible that we have a perception of a whole speaker if voice sounds and mouth movements are inputs of such a different kind? If we assume that the mind is representational and it recollects different inputs that come from outside into a unity, this implies a strong assumption with respect to the environment role on subject perception.

Moreover, under this perspective, we must accept that there are distinct functional modules that are specialized in a specific activity; this means that they are sensitive and compute some features of the environment in the same way through different conditions. From this point of view, the brain harmonizes the different products of different computations of the main functional modules involved in a specific perception. This modular perspective is the root of many different contemporary approaches to consciousness, like global workspace theory (Dehaene and Naddachè 2001). I think this is not a good enough explanation of the nature of perception and the mind's working, but I will consider these critical aspects in more details when I am going to present a different model of explanation in Ch. 4.

2.1.1 Perceptual illusions and plasticity. Evidence of reconstructive processes: semantic ascription

If you observe the brain areas active in blind people while they are reading a braille text, you can see that the visual areas are strongly active instead of the tactile areas and, analogously, in deaf people who display sign language the auditory cortex is active. This extraordinary ability of the brain, in accordance with neural plasticity, is the basis of an effective recovery of cognitive faculties ascribed to an injured area. The multimodal zones are in fact spatially associated with different functional units and hebbian plasticity, as we have seen in the first chapter, prescribes the reinforcement of synchronically-activated structures. These two coupled principles make it possible to discriminate more accurately compound stimuli (one of which is perceived with less clarity) or to cope with cerebral lesions. An alternative pathway, to process signals and give back the same semantic content, can in this way be organized through zones linked with the same convergence-areas of the damaged ones. This shows that the integration operations that lead to perception are not modular and formal, but sensitive to the semantic meaning of the inputs and can be implemented according to different paths.

Considering these observations, it is possible to formulate a *Principle of Semantic Ascription: Perception relies on semantic structure, which depends on previous timing activity*.

This principle is useful for describing the activity of multimodal integration zones, which prove to be sensitive, more than to the form of a signal, to its semantic appearance, interpreted in the light of a horizon of sense formerly constituted. This means that the horizon of meaning that each subjects builds is able to guide and fill the meaning of the next perception, which in turn integrates and modulates that very horizon created through timing activity.

The Principle of Semantic Ascription might be reformulated according to the changed dispositions of the subject. The subject could in fact modify his understanding of the environment or the attribution of meaning to it; nevertheless, the physical external reality would still be known through a reconstructive process. Perceptual illusions show interference between the cognitive side and the receptive one, which consists in the reworking of a signal to make it compatible with the universe of sense of the subject. I will further elaborate on this aspect in Ch. 3, where I will consider *predictive coding*, which is a theory that has its roots in the claim that cognitive understanding and expectations about what is supposed to happen next strongly impact on neural activity and also a present perception. The predictive coding approach tries to model this aspect of mind functioning though a prediction error suppression dynamics. Here I limit myself to the claim that expectation of a certain content can condition its actual perception; accordingly, also the effects of priming and masking are to be considered examples of perceptual conditioning.

The phenomenological approach has largely reflected on the intrinsically integrative aspect of our perception, in the sense that, according to Husserl (1913) every intentional object contains an unperceived virtual part, connected to our possibilities of observation (this aspect will be the main topic of the first part of Ch. 4). This dialogue between virtual and perceptual reality and between actuality and possibility drives the mental process of integration and it is reflected in a modification of the weights ascribed to the inputs and, consequently, in the timing and contextual comprehension of the situation. Therefore, it seems that perceptual reality is the result of a process of reconstruction carried out according to a semantic line. This hypothesis is corroborated by the interpretation of many indirect data about the neurophysiology of the nervous system. In the next section I will elaborate on the codification aspect.

2.2 Informativeness in timing

The nervous system elaborates the environment in which it is immersed by the translation of stimuli into codes through molecules that, depending on the system in which they are inserted, change the electrical trim. Coding occurs on an electrochemical basis. This translation is more than a serial computation. Stimuli that come to our system have different frequencies and it is impossible to establish a proportional or two-way match between the input and output frequencies. Important experimental evidence (Rossi 2020) emphasizes that much of the informativeness of the signal resides, as well as in the frequency, in synchronization and timing of the cerebral signals. According to this evidence, information coding not only employs the recognition of the active units, but also the temporal structures of impulses and the complex spacetime architectures in which they are organized. To probe the relationships between timing and consciousness it can be useful to look into new evidence emerged in the medical field, particularly in the anesthesia process. Narcosis and the simultaneous loss of neural integration might be due to a desynchronization of the signals with a consequent loss of the possibility of their propagation. In the next section we discuss these issues.

2.2.1 The relationship between timing and consciousness

Anesthesiology offers a fertile field for a kind of reverse engineering investigation on consciousness. The mechanism that leads to suspension of consciousness is still obscure, but a progress has been made on the interpretation of signals and data available so far. During anesthesia an adequate level of drugs must ensure the control of the following kinds of components (which are isolated according to Woodbrige 1957[6]): mental

[6] Anesthesia aims to obtain unconsciousness or loss of awareness; analgesia, loss of response to pain; amnesia, loss of memory; immobility, loss of motor reflexes; paralysis, skeletal muscle relaxation and normal

(consciousness), motor, sensory (various afferent impulses) and reflexive (harmful reflexes).

The hypothesis advanced by Tononi and Massimini (2013) ascribes the consequent loss of consciousness to the reduction of the rate of neural integration. Their studies show that the wave of propagation generated by a stimulation through the complex neural architecture is considerably weakened under the influence of anesthetics.

However, other interpretations of this data are possible, for instance the one provided by Marshour (2013). According to this author, the suppression of consciousness seems to be primarily attributable to a desynchronization of signal processing rhythms. This desynchronization would be the basis of a slowdown in the propagation of the signal with consequent loss of integration. Consciousness and its suspension would therefore be primarily and more closely related to timing and synchronization in the process of coding and decoding.

2.2.2 Rhythms, differentiation and synchrony

If timing of the signals condenses much of the information of a circuitry, then each neural structure works as both a receiver and an emitter of signals, as well as a timer and a bearer of possibilities. Every neural cell has in fact a time window of possible activation. This means that each cell receives different inputs at different times that can generate an output with a variable probability level (a stimulus can never generate a certain answer). Timing activity is also linked to this probabilistic aspect.

New interpretations of behavior of cortical neurons (Desmaisons et al. 1999)[7] pave the way for a new evaluation of activity between axons and dendrites, greatly complicating the picture. These data do not reveal the axonal spike, but a complex activity that arises from recursive and modulating dynamics.

---

muscle relaxation. There are three different classes of drugs directed to the last four targets, but it is their interaction that helps to reach the first one.

[7] The authors highlight that one of the main methods to investigate neural activity, EEG, conveys many pieces of information that often are not properly considered. Noise that can be ascribed to dendrites activity could be as relevant as the information that we can detect from axonal spikes.

Conscious experience is not only linked to the generalized and strongly integrated activity of the nervous system as described by Tononi and Massimini (2013), but also to the probabilistic and temporal relationships that are established between these activations, so that the level of integration or modularity of functional maps can also be expressed according to the entropy of the system.

Timing is strongly involved in the modification of the cerebral rhythms that characterize our different levels of alertness and attention (see Fig. 6). In the waking state the neural discharge is continuous or tonic, but with very low oscillatory peaks; in slow-wave sleep oscillatory patterns show high frequency bursts followed by silent pauses. With the reduction of the repertoire of selectable cerebral states, consciousness is lost. It seems therefore that consciousness requires an activity that incessantly changes, differentiated in space over time. The conscious processes are thus typically associated with distributed variations of the cortical-thalamic activity, rapid and effective at the same time.

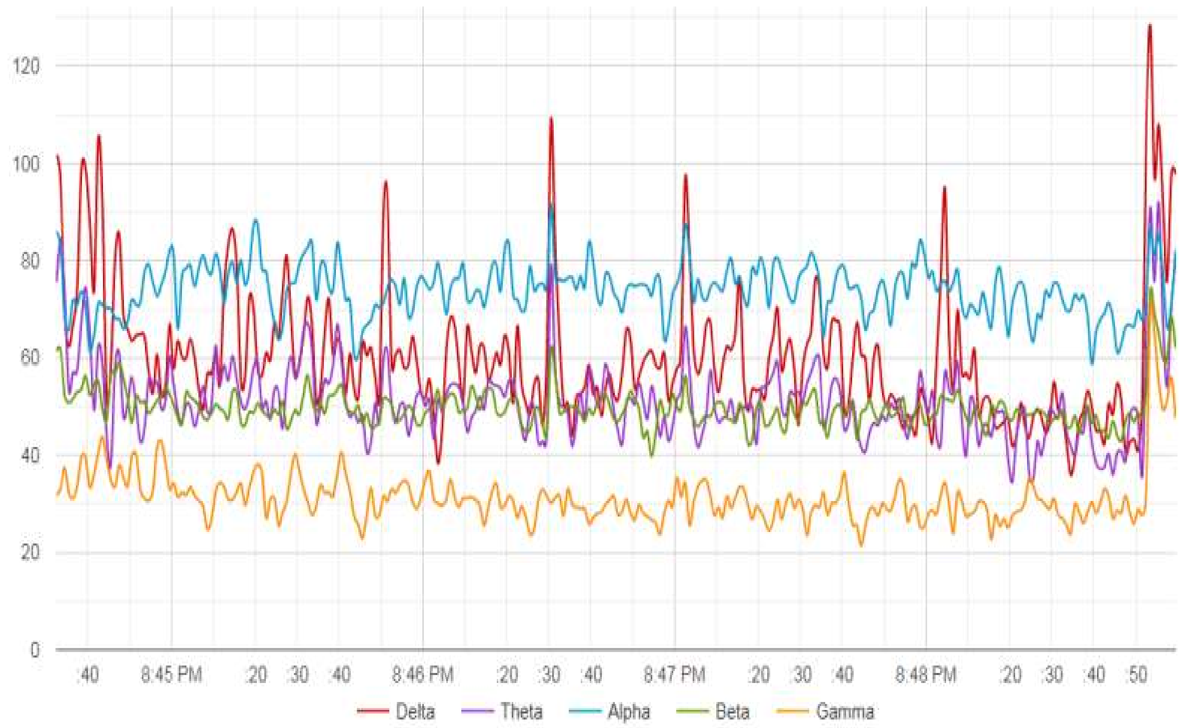


Figure 7 The scheme shows the main cerebral rhythms (waves characterized by different range of frequency) that are associated with different mental activity and level of consciousness. The shape of the wave represents the distribution of groups of action potentials in a time unity.

At this point, we need a short excursus on the problem of the meaning of neural rhythms. When we study and model the nervous system, we constantly deal with wave phenomena. First, spike trains (a sequence of neural activations that we can detect through EEG) typically assume wave form. Second, synchronization between multiple neural cells generates signal spikes that can be detected through the evaluation of brain rhythms with different tools.

Consciousness is associated with a differentiated and changing strong activity, that manifests as a multi-differentiated wave, not synchronized and with very low picks. In contrast, states of not full consciousness manifest with slower and more synchronic tracks. But we must also consider the relationships between the waves that represent the state of alert and those of the meditative states or the not fully conscious ones. What kind of synchronicities can be detected?

Here we can see the problematic nucleus in the debate on the meaning and role of cerebral rhythms, which divides neuroscientific researchers in two opposing camps. Let us briefly see what this debate is all about.

Are the brain waves, that we detect in the EEG, a simple product causally determined by unknown activity of the brain, or are they a tool which the brain uses to encode information? This question hides a deeper question about the nature of brain activity. In the first case, we accept that the brain has an autonomous activity, and we detect that activity through more or less synchronized and fast signals that become visible on the track. Everything is inscribed in a causal chain of impulses and responses.

However, if we consider the second horn of the problem, according to which waves are means that the brain uses to encode and transport information (waves cross the brain and are modulated by it), then we should consider a different dimension of influence that the individual has on the environment, beyond the visible material action and behavior. In fact, we should also consider that electrical activity could have an impact on the environment through the simple interaction between the information flow in the brain and the flow which constitutes the environmental stimuli.

These different approaches to cerebral rhythms and their implications (that will be explored in detail in Ch. 3, where timing-based models will be discussed), show an important point: the most relevant aspect in the brain system concerns the undulatory nature of information processes. This aspect seems to be more relevant than discrete analysis of successive brain events, which can be modelled through more or less complex networks. It thus seems necessary to analyze levels of pulse synchronization and desynchronization in order to better understand the connection between brain activity and subjective perception.

According to this line of thought, we then need to focus on the space-time architecture that occurs between neural individual phenomena, as well as between the neural system and its environment. This thought also leads us to reconsider causality relationships between these phenomena.

From a phenomenological point of view, conscious experience is highly integrated, each state of consciousness comprises a single scene that cannot be reduced to independent elements. It is also highly differentiated, in the sense that an immense number of different states of consciousness can be experienced within a short period of time. Consciousness discriminates between billions of different situations, each of these can generate different behaviors or consequences, and it is thus extremely informative for the organism.

Let us explore this aspect a bit further. Conscious perception of time relates to the possibility to compare different perceptions, and the basis of perception lies on the timing process, so that macroscopic perception of temporality is also rooted in this type of activity. Consciousness is therefore a selective, differentiated, continuous and, at the same time, continually changing process. Internal coherence and privacy are general phenomenological characters of consciousness, almost universally agreed. Instead, its unitary character has often been questioned. Edelman, like many other authors, proposes a stratification of consciousness on two levels. This critical aspect will be discussed in the next section.

2.3 A preliminary conclusion: reconstructive processes

The relevance of timing has been discussed, not only for the sleep-wake cycle or in the transition between consciousness and unconsciousness, but also for the formation of the content of conscious perception itself. This, through multiple evidence, has been delineated as a highly integrated and semantically oriented reconstructive process.

A high level of integration in the central and peripheral nervous system has been proved to be a necessary condition for the emergence of conscious experience. In this sense, even the intersubjective dimension becomes very important, because the neural resonance, generated by the "mirror mechanism", strongly interferes with the subject's elaboration and distribution of weights within his/her system, but also because the external observer registers the physical discrepancy between different inputs and contributes to the normalization of discordant or contradictory perceptions.

This assumption of reciprocal integration between systems is justified in the light of the neurophysiological analyses of timing, and also by appealing to theories that underlie the aspects of continuity and the dynamics of the mind (Maturana and Varela, 1980)[8].

Having set these theoretical tools, I propose to question the distinction put forth by Edelman (Edelman 2000; also see Damasio 2000), between primary and secondary consciousness. This kind of distinction, which does not seem to account for the integrative nature of perception, is, in my opinion, phenomenologically unrealistic.

By the analysis exposed above, it is impossible to isolate a content of a punctual perception, because perception is always the result of an integrated and semantically driven process. In the distinction between primary and secondary consciousness, which is thus not tenable at the ontological level, it is possible to recognize a linguistic genesis.

At the phenomenological level, there does not seem to be a detectable separation between a punctual content of consciousness and the flow of self-consciousness wedged on it.

[8] Maturana H. and Varela F.J. (1980). The mutual reconstruction between the individual and its environment is analyzed here from a dynamical point of view which considers both polarities. In Ch. 3, I will explore a different kind of homeostatic relationship with the environment by analyzing the role of the very body in the construction of the perception of both objects and actions perception.

The content of a perception, as claimed in Sect. 2.1, is already mediated by a semantic horizon formed by the possibility of the mind to transcend the direction of the temporal axis. The mind can in fact compose contents of previous perceptions, expectations and imaginations aimed at the future, in a significant universe. In addition, the reconstructive possibility of the mind implemented in multimodal areas with parallel synchronous and bidirectional communications, exhibits a situational and semantic character that makes it inconsistent to talk about objects and singular contents separated by their "apperception" in the Kantian sense.

Timing concerns both individual neural units and general connections of the system. It makes conscious experience inherently distributed, integrated and intricate with content that it is possible to objectify only by a linguistic act aimed to an objective analysis of conscious experience.

Consciousness is given to the subject as a multifaceted reality, with blurred boundaries, and distributed on different levels, which could be linked with some neural responses. The proposal is not to consider consciousness as a private and stratified reality pertaining to a subject, but existing as a possible and temporally-structured relationship between impulses. The perception of matter and body is the result of an intense process of integrative reconstruction, which is directed, at the semantic level, by the displacement of the attentional focus, a multidimensional fluid encounter of signals partially synchronized in a mutual reconstitution of the system and its very limit.

2.4 A multi-level stance in cognitive explanations

As a conclusion of this first part, I want to briefly clarify some aspects. The notion of "levels of explanation" presupposes the existence of different levels that collectively contribute to shed light on the same reality. According to interactionists positions, which reject autonomous levels of explanation, reality seems to be the limit between environmental change and the dynamical properties of a system, where either side is constitutive of the other. The concept of level helps us understand the fine line laying between the two sides, but we need to take a specific stance to evaluate internal/external and top-down/bottom-up distinctions. I show here that different levels of description are necessarily and mutually interrelated in cognitive explanations. Hence, a multi-level stance is always required for a complete understanding of cognition.

The first problem I focus on relates to learning. I maintain that the notions of “circuit” and “synchrony” (Hebb 1949), which are required for the explanation of synaptic reinforcement, cannot be properly understood if we only look at the neurophysiological level, for the full significance of these phenomena can only be understood at a higher cognitive/behavioral level. The same problem arises in connection with the notions of convergence zones, exposed when I talked about LeDoux in the first chapter, and in the field of perception, and value systems exposed in Edelman’s and Tononi’s approaches to explain attention processes. The explanation of attention processes is problematic even at a higher cognitive level, because attention represents a sort of superlevel placed at the same level as the other psychological mechanisms it controls.

I will show, then, that analogous problems may be found in the computational theory of mind. Here, we have a very general explanation of the mind, whose functioning is conceived in terms of algorithmic operations over symbolic representations. But this single level explanation is subject to severe difficulties. I maintain that these problems can be fixed once an interactionist theoretical background is assumed.

My general conclusion here is that describing phenomena at different levels is epistemically useful in many scientific fields, such as physics, biology and psychology. However, none of these levels belongs to reality per se. Rather, they rest on the teleological interests of the observers and their explanatory value can be properly understood only through continuous and reciprocal cross references.

In the same vein, splitting psychological reality by describing at different levels the processes of memory, attention, perception, imagination, and so on is utterly useful to understand cognitive behavior.

However, the selection of the relevant variables at any level is the result of an abstraction from some real phenomena that are always given in their global complexity, rather than in some abstract local level at which we choose to focus our point of view.

Chapter 3

Fragmented or continuous perspectives on learning methodological roots and limits

A discussion on causation

This chapter aims to make explicit the different notions of causation that are presupposed by different research methodologies, and the limits of their applicability. In particular, I analyze the notion of causality assumed by neurophysiological enquiry, and I maintain that it involves a fragmented and modular understanding of its object. This method proves to be fruitful for an analysis of certain specific motor faculties, but it fails in its application to the understanding of consciousness. There are several notions of consciousness examined by neurology that are indirectly derived from this method of enquiry. First, I will show the links between the assumptions of a modular realism, a one-way temporal concept, and autonomy ascription, which are implicated by the neurophysiological or, more generically, materialistic methodology.

In particular, I discuss the notions of *integration* and *interaction between fields,* associated with consciousness. I highlight the implications of these concepts with respect to a typological and morphological categorization of stimuli, and their limitations. I then outline a different conception of the notion of integration that considers a fundamental property of consciousness: its being linked with time and continuity. I will argue that consciousness cannot be understood within the fragmented conception of cause and time typical of the neurophysiological method of enquiry. This opens up the possibility of a different paradigm which, in order to better understand these fundamental aspects of consciousness, is based on some reflections that belong to the theoretical apparatus of phenomenology. In particular, I explore synchronicity in order to propose a re-reading of the waves phenomena that occur in the brain, the role that they can play in encoding information, and their correlation with conscious experience.

There are many theories that try to approach the fascinating theme of the emergence of consciousness, its possibility conditions, its neural correlates and the links it has with other phenomena. An exhaustive picture of them is not the purpose of this analysis and there is no claim here of completeness in the examples that will be addressed; the target of the analysis is rather to highlight some of their assumptions and discuss their implications.

The first section examines the causal assumptions on which some paradigmatic models of neurophysiological enquiry are based, and their implication of a modular conception of mental faculties, which is identified as the foundation of conscious experience. At the same time, these models show a dependence on a granular, fragmented conception of time and conscious content. Some critical circular aspects, implicit in these accounts, are here analyzed and, by referring to the tools of neurophysiological enquiry, I review some notions of consciousness implied by a causal and mechanistic interpretation of these analyses. Among these approaches, I first analyze the notion of consciousness as *integration.* I emphasize its dependence on a typological conception of information based on the spatial configuration of signaling. It also shows a dependence on a strict delimitation of the boundaries of the individual system and, at the same time, of the boundaries between different contents that seem to constitute subjective conscious experience.

Notwithstanding the explicit claims of the theory, this account does not seem to capture the most essential aspects of conscious experience.

Integration, as understood here, seems to be an instantiation of that interpretation of emergentism, which we call “vertical”. A second interpretation of emergence might be called “horizontal”, and the vision of consciousness as a global workspace depends on it. This will be the theme of the second paragraph. Also this second approach is dependent on a modular and representational concept of mental contents

A third way to think about interaction and fundamental entities is possible. Even if this view has many limitations, it opens the way for a consideration of the interaction between wave phenomena. The aspects of synchronicity and consonance will then be reconsidered in the second part of this chapter, where a new interpretation of consciousness will be proposed, which is based on philosophical approaches to phenomenology, as well as 4E approaches to cognition. This takes into account the difficulties of the models described above, in order to soften the boundaries between the subjective and the objective polarity, as well as among the contours of different contents of consciousness, to provide a more coherent account of subjective experience.

3.1 Modular dynamics and consequentiality: an epistemological evaluation

The traditional research paradigm that drives neurophysiological enquiry operates according to a strong directional causality assumption: it is based on the perturbation of the neuronal system through precise and controllable stimulations, which are considered causally decisive with respect to changes that take place in the individual system, identifiable through Brain imaging techniques. Central to this method is, on the one hand, the assessment of sequential neural activation changes in the individual system, supposed to be the result of the consequent elaboration of a stimulus. On the other hand, it is also important to evaluate the relationship that can be established between the different steps of this process. There are different theoretical models that adopt this method.

The *rate model* endorses the centrality of certain blocks of activations in encoding information. An alternative approach is taken by the spike model, both at the ontological level (with respect to which units are taken to be fundamental) and at the application level. This second approach will be explored in detail in Sect. 3, which deals with models whose main focus is temporality, consonance, and synchronization.

The firing rate is an abstract concept that is defined in a limit that involves an infinite number of spikes (of course, experimental measures of firing rate necessarily involve finite numbers of spikes). For example, it can be defined for a single neuron as a temporal average: the inverse of the mean inter-spike interval.

There are other definitions, but in all cases, rate is an average quantity defined by taking into account the timing of spikes. It is important to distinguish two different concepts: spike timing is what defines spike trains, whereas firing rate is an abstract mathematical construction on spike trains. Therefore, the rate vs. timing debate is not about which one is the relevant quantity, but whether firing rate is a sufficiently good description of neural activity or not. Spike-based theories do not necessarily claim that rate does not matter, but they question the claim that firing rate is the most essential quantity for such a description.

Starting with the work of Beurle (1956) a popular modeling approach has developed a continuum of models aimed to describe the average activity of large populations of neurons (Wilson and Cowan 1973, Nunez 1974). This paradigm is anchored on two implicit assumptions. First, that there is a strong and constant ***differentiation*** between an individual polarity and an external objective polarity. Second, it requires a ***modular dynamic,*** structured in the following way: from the external side, stimulations involve the peripheral system, this transmits information to the central system, in which appropriate computations generate corresponding outputs, which emanate through the peripheral system and produce perceptual or motor responses.

This modelling approach also involves a ***typological differentiation*** of stimuli, a property that depends on the area of processing involved. For example, a visual stimulus is considered a kind of stimulus that generates activations in the visual cortex and, in a similar way, an auditory stimulus is processed in a distinct area. This conception of causation proceeds from the external stimulus towards internal processing, and from internal processing to the causal response generated from the incoming stimulation.

By tracing the path of signals in this way, many neurophysiologists have argued that the emergence of conscious experience mostly depends on the cortex, which would have a predominant role in the organization of perceptual content. In this perspective, consciousness is taken to be connected to a spatial and causally determined localization. This is precisely one of the main points that are questioned in this section.

As we will try to emphasize, the cortex is certainly implicated in the emergence of conscious experience, in the sense that it constitutes its conditions of possibility, but it does not exhaust its contents. The possibility of deferring attention and not consciously perceiving stimuli that nonetheless are correctly detected by the cortical system supports this intuition. A new focus for the parameters that remain unchanged through stimulation will then be proposed.

As it is well known, neurons are distributed in a dense network of connections, and the information transition between them is studied at several levels. Either the activity of individual neuronal cells, or the connection between cellular complexes, can be considered.

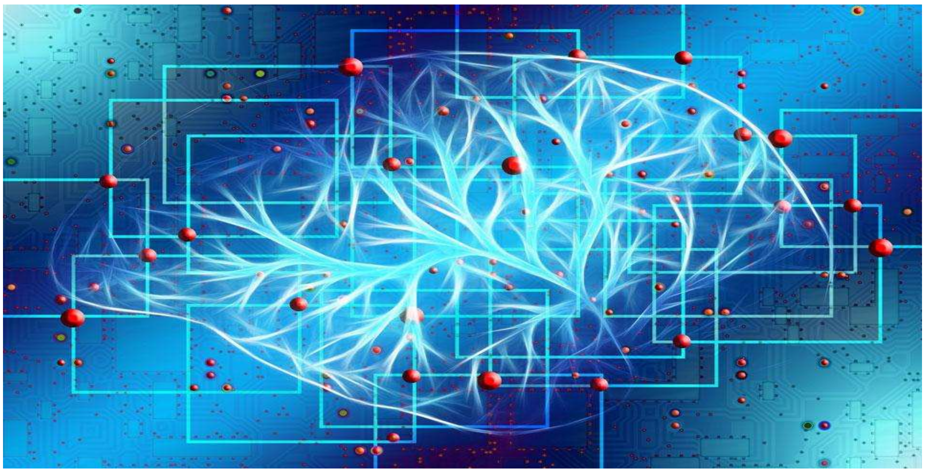


Figure 1 This picture shows how neural nets theory models the activity of neural system in human body. The red dots are functional units distributed in peripheral and central system. Their different dimension symbolizes the different relevance and gradual impact that they have on the network. The bleu lines symbolize the connections between different functional units

The informational exchange between them can be evaluated at the molecular level, as synthesis and trade of neurotransmitters that contribute to the modulation of signals, or at a different level, as electrical potentials that move along the axons and that reflect membrane depolarization and the related trade of neurotransmitters.

Each synapse is associated with a weight that is determined by the impulses from the previous neuron, so we can see as input of a neuronal cell the weighted sum of signals from the axons of other cells. Importantly, in these models, coherency, which means simultaneous activation of multiple complexes, is often considered to be a measure of the morphological and functional connection.

This link between action time, spatial placement in the neuronal network and functional connection will be analyzed analytically in the next few paragraphs.

At this point it is necessary to focus on a crucial point: neurons can have an exciting or an inhibitory nature. We know that there are cells that are responsible for the activation of the postsynaptic cells they are connected to and, on the contrary, there are cells aimed at the inhibition of postsynaptic cells. This depends on the synthesized neurotransmitter which increases or decreases the probability that the following neuron fires. However, the complexity of the neuronal system makes it very difficult to fully comprehend the activity of these different types of cells. Therefore, stimulating the system means having an impact on cells of both types, excitatory and inhibitory. So, tracing the signal only as successive activations, leaves out of the explanation a very important inhibitory component.

It is useful here to point out that typological differentiation is implied by a modular conception of the functional units linked to an equal time fragmentation into equal sequential portions corresponding to the changes in the state of the aforementioned functional units. This process develops according to a directional time axis and is described through the isolation of a stimulus at time t_1, the observation of a modification that takes place in the neuronal structure at t_2, (modifications which trace the differences between a previous state), and therefore through the isolation of an event that is considered to be a consequence of this path. Determining a response is seen as a mathematical function of incoming stimuli.

If this neurophysiological model is complicated by using neural networks that can model even multiple concurrent stimulations, the output is seen as an integrated function of such stimulation composition through the attribution of different weights distributed in the neural network as a whole[9] , as it is shown graphically in fig.2. Here it is showed a quite simple scheme that explicates the most relevant components of the causal neurophysiological enquiry and explanation.

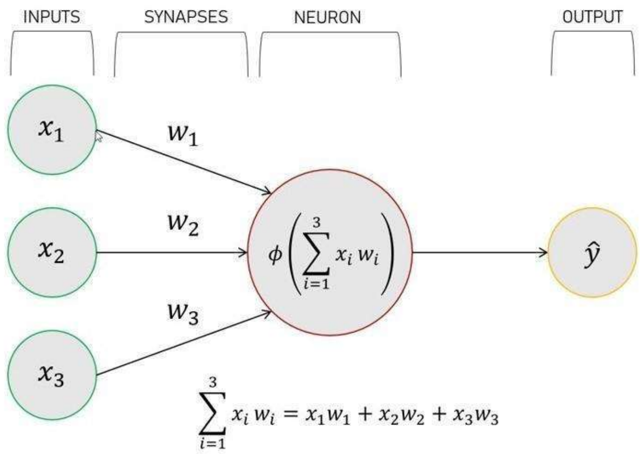


Figure 2 Output as a result of a linear sum of weighted inputs

The model, based on the standard neurophysiological research method, does not explain why one output occurs instead of another, and it doesn't explain a specific reason for this but avoid to talk about this, talking only of a rate of probability of a given response to stimulations. It describes what a group of cells can do, under certain conditions, not what they actually do. The focus of this kind of approach is the description of what is happening while the explanatory power seems lacking, and this impact also on the predictive capacity of this kind of model.

[9] We can model neurons as non-linear functions that transform inputs (x_1, ..., x_i , ... x_n) into an output y. Every input is associated to a weight (w_1, ... w_n), For the neuron to fire, the weighted sum $\sum_i x_i w_i$. of the different inputs should reach a threshold, which depends on the specific morphology of the network. In recursive networks, the output serves as a further input that provides information about the error. Each input, besides a weight, has another parameter, w, which can be seen as the weight (the relevance for the whole network functioning) of the input .The output, 0 or 1, is determined by whether the weighted sum of the incoming stimuli reaches or not the firing threshold.

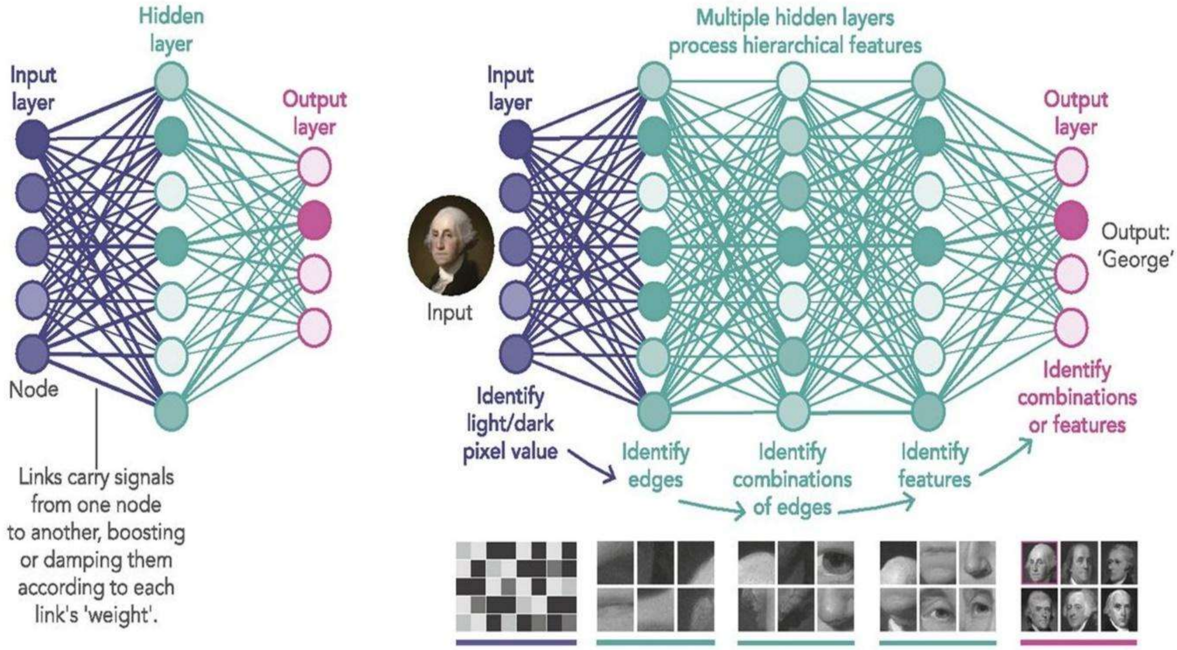


Figure 3 A connectionist model of neural networks that shows the link between an input and a output through the process of hidden layers. The input is a picture of George. The first layer, the blue one, corresponds to the distribution of gradation of pixel colour. Multiple hidden layers, green, process hierarchical features, like the edges or the organization in some kind of particular shapes. The final output, pink, corresponds to the output layer which identify combinations of features of the previous layers, giving a perceptive output.

A cognitive activity performed by the brain is seen, from a connectionist perspective, as the result of a complex dynamical elaboration of an external input, that could be pictured as shown in Fig. 3. The first layer corresponds to the distribution of gradation of pixel colour. These incoming data are then processed though many other layers, multiple hidden layers, that process hierarchical features. Each of them, is devoted to discrimination of specific features, like edges, or specific kinds of shapes. The final output corresponds to the output layer, which identifies combinations of features of the previous layers, giving a perceptual output.

Deep learning is a type of machine learning that has received increasing attention in recent years. With deep learning, an algorithm does not need to be told about the important features of the data. Instead, it is able to discover relevant features from data on its own using a neural network. Artificial neurons are arranged in layers, and deep learning involves a deep neural network (DNN) that has many layers of artificial neurons.

As the name *hidden* denotes, a neurocomputational approach modelled by networks like the one exposed before, seems to be more interested in the descriptive level than in the explanation of the reasons for the occurrence of a specific output. This is the biggest difference with the spike-model which, conversely, aims to investigate the shape and structure of the single response, conceiving it as the fundamental information unit, and searches for an explanation of the probability of a single unit response and its causal power over collective neuronal responses.

The rate model, considering the average of activations in the unit of time, makes an approximation on the probability that a neuron fires. The model establishes a unit of time, which is supposed to be relevant to the research, and evaluates in it a content that is supposed to represent a stimulus. In order to see a response as a function of the stimulus, it needs to check that a structure isomorphic to the stimulus occurs in such a unit of time, so that a causal assignment can be made. This presupposes a delimitation in space and time that makes such isomorphic relationship possible.

An important difference emerges here with respect to how the rate model or the spike model interprets information coding. Both models see information coding as a representational structure, or a map of the stimulus, that leads to a satisfactory response. But one, the rate model, considers it as a block of activations in relation to an external occurrence; the other one, the spike model, instead looks at a more subtle ontological level, by blurring the boundary between these two polarities. We will consider these aspects in more detail in the next section.

Notwithstanding this difference, both models see this internal/external interchange as being aimed at the preservation of the individual system, which would be achieved by a homeostatic response to the environment, in order to preserve balance and integrity. However, it is very difficult for either model to explain the notion of integrity and homeostatic balance, as well as many cases that seem to endanger the preservation of the system, like aberrant activations, the mechanism of addiction, the one of chronic pain, or activations due to expected, but not actually occurring, repeated stimuli. These cases will be addressed in chapter 5, in support of a different interpretation of these notions.

To sum up, the strong assumption shared by these models is that the same laws apply to both the individual polarity and the external one, while the limit between them is kept stable. These laws involve a one-way causation, which is implied by a modular conception of stimuli and functional units of elaboration, and the granularity of time determined by them.

In the next section, some interpretations of the notion of consciousness, which stem from the methodological assumptions described above, will be analyzed, and some of their limitations will be highlighted. In fact, even if the model of neurophysiological enquiry just described proves to be useful to simplify and predict some motor and perceptual responses that concern specific human faculties, it does not seem to be appropriate to understand the phenomenon of consciousness, because its requests are too rigid for an object of study with so many different and subtly interlinked properties. Integration is the first notion that we are now going to discuss.

3.1.1 Neural integration: the dilemma of composition

The concept of neural integration has been variously used to sum up in a word the notion of a theoretical system that makes specific assumptions about the fundamental properties of consciousness and the conditions under which a system can give rise to phenomenological experience. According to Edelman and Tononi (2008)[10] consciousness depends on a specific mode of organization and interaction of the constituent components of the system.

[10] Edelman. G and Tononi. G. (2008) *A Universe Of Consciousness: How Matter Becomes Imagination*. Basic Books. This rich work makes a strong claim about what brain waves detected by EEG show. Different levels of neural integration are represented by different tracks corresponding to different subject conditions.

It is expressed by a high level of entropy that is connected to the kind of waves that would characterize conscious states, which manifest a low level of synchronization and consonance, unlike those that characterize states of unconsciousness. Let us consider Tononi's definition of integration (Tononi & al. 2014):

> At the level of an individual mechanism, integration postulates that only mechanisms that specify integrated information can contribute to consciousness. Integrated information is information generated by the entire mechanism above and beyond the information generated by its parts. This means that, as far as information is concerned, the mechanism is irreducible. The integrated information f is calculated as the D distance between two probability distributions: the cause-and-effect repertoire specified by the entire mechanism is compared with the cause-and-effect repertoire of the partitioned mechanism divided into its components. (Tononi 2014)

This theory of consciousness, according to the author, requires the acceptance of fundamental axioms drawn from the observation of phenomenological properties of conscious experience and from empirical evidence in the neurophysiological field.

The properties of consciousness that the theory of integration isolates as essential are the following:

- Existence: Consciousness exists – it is an undeniable aspect of reality.

- Composition: Consciousness is composite and structured: each experience consists of multiple aspects in various combinations. Within the same experience, you can see, for example, left and right, red and blue, a triangle and a square, a red triangle on the left, a blue square on the right.

- Information: consciousness is informative: each experience differs in its particular way from other possible experiences. For example, an experience of pure darkness is what it is in that it stands out, in its particular way, from an immense number of other possible experiences.

- Integration: consciousness is integrated: every experience is (strongly) irreducible to non-interdependent components. For example, the experience of the word is irreducible to its component letters, the experience is a whole. Similarly, seeing a red triangle is irreducible to see a triangle but no red color, plus a red spot but no shape.

- Exclusion: Consciousness is exclusive. Each experience excludes all others – at any given time there is only one experience that has its full content, rather than an overlap of multiple partial experiences; every experience has defined boundaries – certain things can be experienced, and others cannot; each experience has a particular spatial and temporal grain – it flows at a particular speed, and has a certain resolution in such a way that some distinctions are possible and finer or more coarse distinctions are not.

This fascinating theory exhibits, like any theory, some limitations. In fact, at least some of its axioms may be questioned, for they imply a fragmented concept of the mind.

As I represented in the schema in Fig. 4, some fragments intersect their content and contribute to the appearance of a composed ones that is precisely the content of consciousness, represented here in blue.

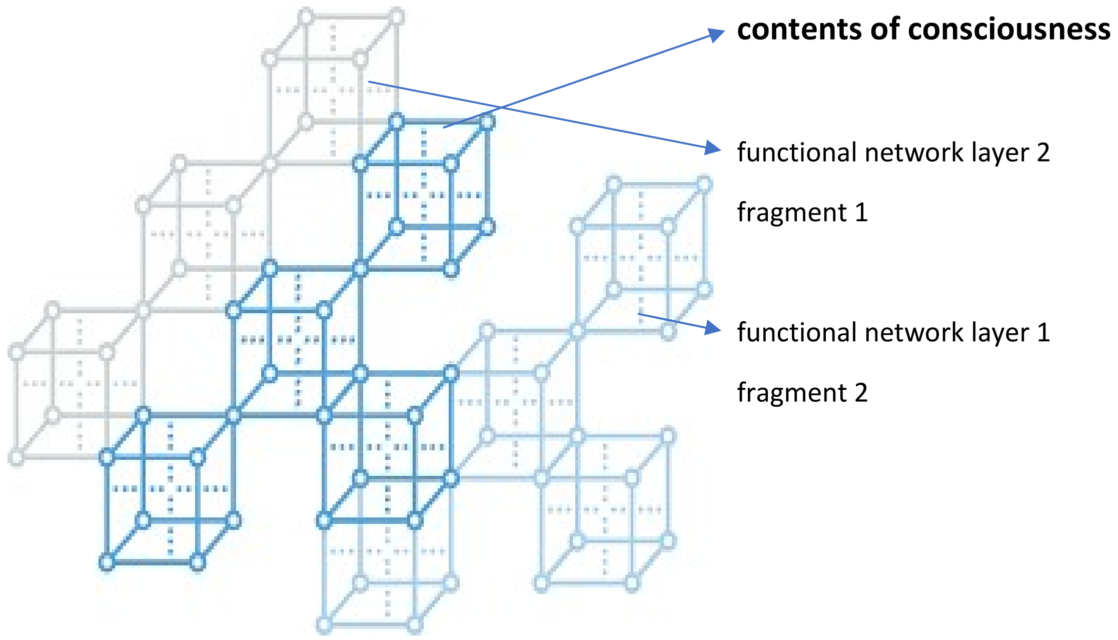


Figure 4 This scheme represents graphically one of the aspects of the theory exposed above: two different functional networks (layer 1, the grey one and layer 2, light blue). Some fragments of them contribute to the appearance of a composed ones that is precisely the content of consciousness, represented here in blue.

In my opinion, in stating the fundamental axioms of this approach, the observational level and the phenomenological ones are confused. In particular, the point concerning exclusion is extremely problematic for two reasons.

First, it is intended to represent the main characteristics of phenomenal conscious experience. But phenomenal experience does not display sharply limited edges, or a well-defined space-temporal grain. Experience consists of multiple factors that can be more or less vivid and relevant in space and time, depending on the shift of the attentional focus

Second, the relationship between each single experience and the infinite possible ones that could arise from the same configuration of events, is not clear.

If it is true that, to be conscious, an experience must differentiate itself from other possible experiences, as stated in information principle, the perceptual content does not contain the representation of the possible alternatives to it.

This kind of integrated approach, as long as it emphasizes the importance of integration development as defined above, considers consciousness as an emerging property of a complex dynamic system. Emergence is meant here as an ascending process that stems from the integrated organization of multiple systems that operationally interact.

3.1.2 Integration and emergence: between abstraction and intersection

A way to understand an emergent phenomenon, which is shared by several different approaches to consciousness, is to see it as a different and sequential aspect that arises from a specific organization of discrete events. This vision of emergence was first put forth, over a century ago, by Huxley (1893), and then by Morgan (1923) and Broad (1925).

This way of explanation proves to be effective in describing many dynamic systems, but, in my opinion, it is not sufficient to describe the phenomenal and functional properties of consciousness. For it subscribes to a directional temporal conception, in which interacting events and the succession of their space-time specific architectures involve the emergence of consciousness.

Even within the different approaches that read consciousness in this way, there is disagreement about the relationships between the fundamental events, which are located at the same ontological level, and the ones that occupy a different ontological level and are constitutive of consciousness.

Consciousness under this perspective emerges from a particular organization of more fundamental elements, which, depending on the context of explanation, could be seen as physical or mental properties. We could try to visualize this perspective as an abstraction from down to upper ontological levels as shown in Figure 6. Many phenomena are generally used as examples of emergent properties, like diamonds, snowflakes or special dances of birds in the sky. Here I prefer to use geometrical examples because, visually, I think it could be helpful to think about two different aspects of the same notion. I distinguish a vertical perspective, which derives from an abstraction of aspects of lower levels of composition, from a horizontal perspective, which derives from an intersection between components.

Figure 5

A drop in the water could help to visualize a vertical interpretation of emergence. Components of the same kind originates different structures, and they have a two-way constitutive relationship.

Although this ontological separation between different levels is precisely what the emergence theory aims to cancel, it is evident that its description of the organizational relationships presupposes a given perspective, and therefore operates from time to time some caesuras on the dynamic system, which have a certain investigational function.

On closer inspection, many theories of consciousness that come from the neurophysiological field implicitly assume a notion of emergence, when they postulate that certain parts of our brain, such as the cortex, are more involved in the emergence of conscious experience and are associated with higher perceptual functions. Models of this kind put fundamental elements at a lower ontological level, where they perform unconscious functions, and the emergence of phenomenal consciousness is then abstracted from a particular organization of their interactions and relations with the cerebral context

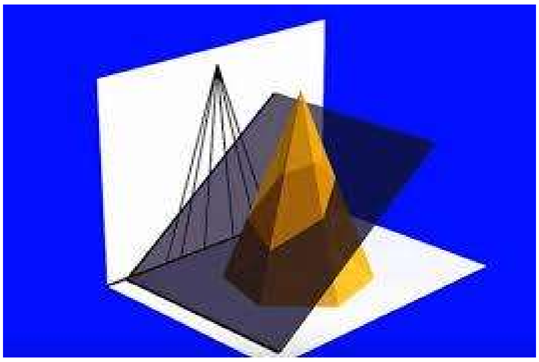
Causation between stimulus and response is assessed in the light of a directional temporality and arranged spatially in a vertical line that connects the nuclear structures of the central brain system with the cortical structures. Emergence has so far been interpreted in an ascensional sense as an integrated extraction of information from lower-level organizations.

Another reading of emergence, alternative to the process of integrated ascensional abstraction just presented, is possible, and it seems to circumvent some of the difficulties presented above.

Figure 6 exhibits geometrically the distinction between these two approaches.

Figure 6

This geometrical representation could help to consider the distinction between vertical interpretation of emergence and horizontal. The first is here represented by the upper part of the pyramid, weather the horizontal is represented by the brown hexagon, intersection between plane and pyramid



In fact, emergent phenomena can also be thought to be at the same ontological level as the organizational structures of the fundamental basis: the model that best represents this concept of emergence is the one of *intersection*.

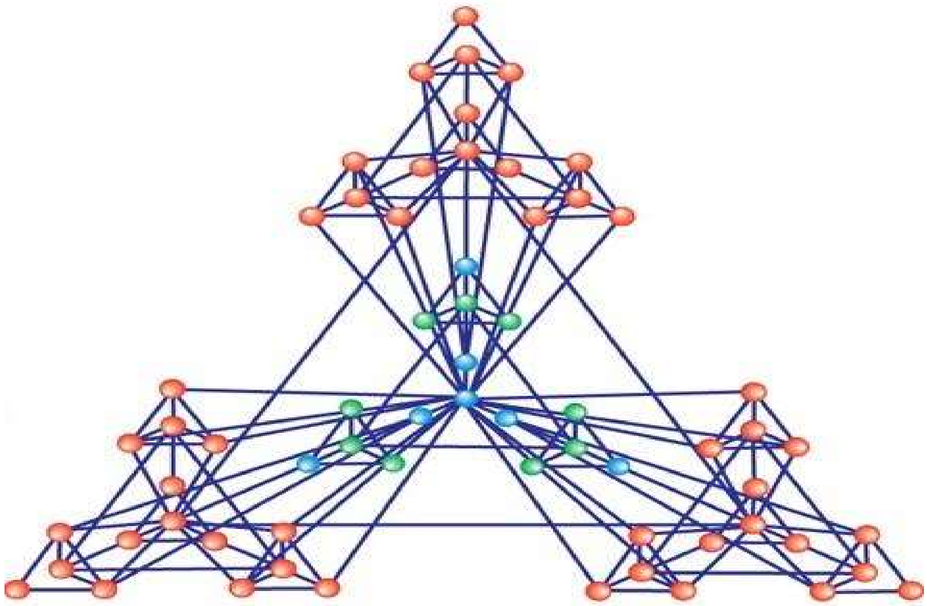


Figure 7 Representation of emergent properties in a neural network schema

This interpretation, which is somehow reminiscent of the Platonic concept of participation of sensible things in ideas, seems to be more effective in describing the phenomenological properties of consciousness, and it will be explored in Sect. 3.1.2.

According to Bergson (1896)[11] as well as many other philosophers after him, the most original and obvious aspect of consciousness is the tension between its continuity and its ability to transcend temporal directionality. Taking this suggestion seriously, we should therefore look more carefully at those factors that are not directly involved in the stimulus-response directionality.

If we analyze the metabolism of some molecules, which have an impact on the central nervous system, we can say that they act on certain receptors, by modulating the presence of certain neurotransmitters. This is correct from an observational and empirical point of view. What is problematic is the inference from this observation to the claim that this process affects consciousness in a specific way. It is true that certain molecules generate a specific variation of behavior with a higher or lower statistical incidence, but a phenomenal experience is not always associated in the same way to such variations, and it could be better explained in the light of the dynamics of what does not come directly from stimulation. The phenomenal side, or the "invariant property", could thus be better understood by not considering the consequential activations that are causally associated with a precise stimulus. Some aspects of philosophical phenomenology that might be relevant for this kind of reading of neurophysiological processes will be discussed in Sect. 3.2.

3.2 Sharing between functional modules

One theory that moves along the emergentist line that we can define "horizontal", has been proposed by Dehaene and Naccachè (2001) and reads consciousness as a shared global workspace. According to this perspective, consciousness emerges from, or rather coincides with, a content shared among multiple functional modules for a sufficient period of time, which is needed to reach the level of conscious perception. This is the description of global workspace that Dehaene gives, that I report precisely because I would like to discuss a little in this section about this claim:

[11] Bergson provided insightful phenomenological observations about time perception and the famous distinction between subjective time and objective time

> Besides specialized processors, the architecture of the human brain also comprises a distributed neural system or 'workspace' with long-distance connectivity that can potentially interconnect multiple specialized brain areas in a coordinated, though variable manner Through the workspace, modular systems that do not directly exchange information in an automatic mode can nevertheless gain access to each other's content. The global workspace thus provides a common 'communication protocol' through which a particularly large potential for the combination of multiple input, output, and internal systems becomes available. (Dehaene, Kerszberg & Changeux 1998)

The authors are proposing here a new interpretation of a famous theory developed by Baars J.B. in 1983, *Global Workspace Theory*. With this theory Baars provided an explanation of consciousness as a space of interaction between different systems that operate on different features. In his first formulation this space, or this central interchange, was called *global data base* (Baars, 1983). He used the metaphor of a bright spot on a theatre stage to explain his view. In the dark side of the backstage a lot of systems work, shaping the contents of consciousness without themselves entering consciousness (Baars, 1997). Consciousness is the show that appears on the stage as the product of cooperation between audience and backstage workers. In 1998 he worked with two other researchers to highlight some common aspects of their view about workspace functioning, especially shaping the mathematical aspects of this model (Baars et al. 1998).

Deheane and Naddachè, in more recent years have proposed a new interpretation of the global workspace theory. According to this perspective several functional modules, that process different cognitive and motor activities, intersect contents in a global workspace through specific long-range anatomical connections, which are able to connect distant areas, and to determine which specific modular systems make their contents globally available to others the workspace, is just only an empirical issue.

It is thus not necessary that shared information be represented elsewhere, sharing itself gives rise to consciousness. This is why I have chosen to refer to this theory as a horizontal view of consciousness emergence. Despite this theory tries to avoid a representational content grounding, it has some problematic through aspects that I review below.

The possibility for a content to become conscious is subordinated to the type of cell units and the specific morphology of a given circuit, as well as to the time units in which it remains active in that circuitry. Forms that remain active for too short a time, or which are disconnected, do not originate conscious content.

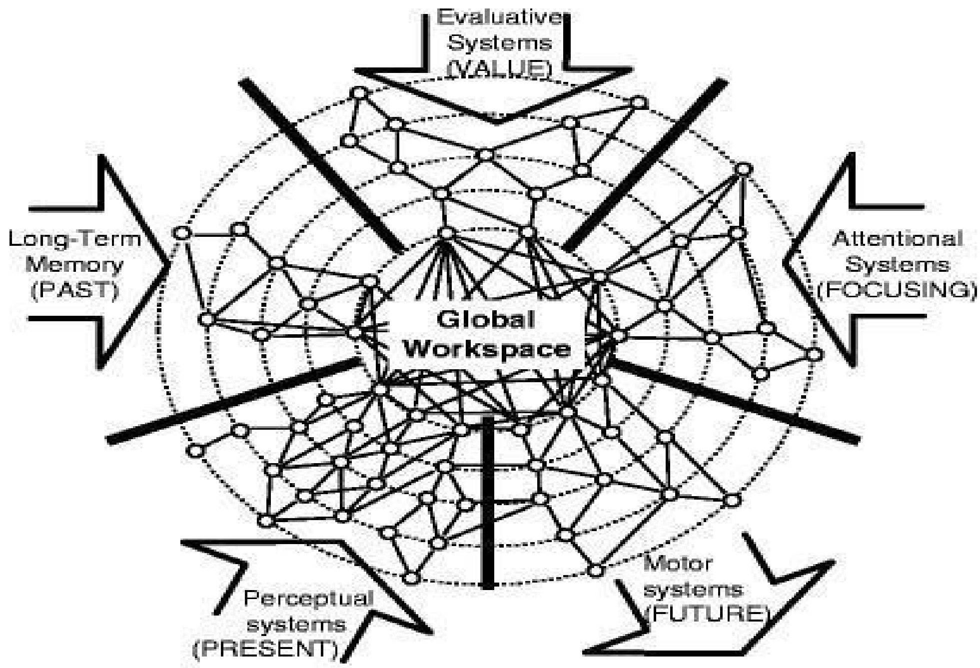


Figure 8 The scheme represents the intersection between different functional units in the global workspace. It reminds the horizontal emergence schema shown in fig 7**.**

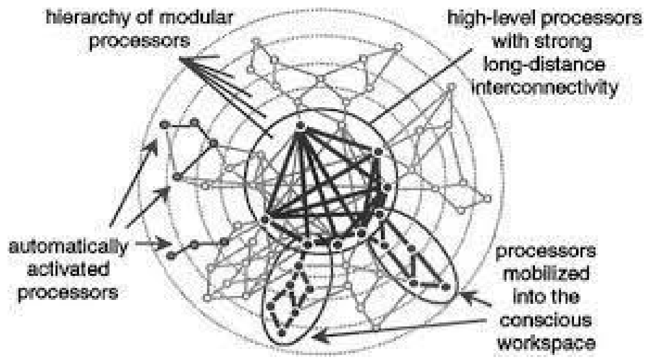


Figure 9 The scheme represents the morphology and way of functioning of the cooperation between different functional units in the global workspace.

However, according to this theory, for a content to become conscious it is not sufficient that it remains active for long enough, as it is shown by the many processes that take place at the subconscious level.

The large amounts of information that we can consciously process would also suggest, according to this theory, that at least five main categories of neural functional modules participate in the workspace: perceptual circuits that inform about the current state of the environment; motor circuits that allow for the preparation and controlled execution of actions; long-term memory circuits capable of restoring the past states of the workspace; circuits that give them value in relation to previous experience; attention or top-down circuits that selectively guide the focus of interest. The global interconnectedness of these systems would explain the subjective unitary nature of consciousness.

Nevertheless, this subdivision is arbitrary and misleading. First, the very definition associated with these processes is as questionable as speaking of five separate systems. This seems to be an operational distinction that is not coherent with the idea that the global workspace is structured precisely thanks to an interconnection and sharing of different overlapping processes, which gives rise to phenomenological experience. If information is portioned according to the situation, and its very content arises from this sharing process, it seems that a strict distinction between circuits has to be discarded.

Moreover, this theory shows a problem of circularity, which is important for the sake of our analysis. In the same paper mentioned above, in fact, we read:

> We suggest that, in many cases, the ability to keep representations in a focus for a period of time in the absence of stimulation seems to require consciousness. By 'in an active state', we mean that the information is encoded in the discharge patterns of active populations of neurons and is therefore immediately available to influence the systems with which they connect. (Dahaene and Naddachè 2001, p. 9)

It is evident here that this account is strongly linked to the concept of representation as a certain configuration of firings that represents a closed and limited content. If we accept this definition, however, it is problematic to understand the second part of the statement. Maintaining an active representation implies that there is a dynamic structure that persists equal to itself through the passage of time. This is problematic because, for this theory, a specific and defined content depends on a complex of fragmented components, which belong to different functional systems. These systems make such components available in the shared workspace from time to time. It can therefore be seen that this approach is based on a granular conception of time. And this claim does not seem to be consistent with a reproduction of representational contents that identically persists through time.

Furthermore, there is a connection between this aspect and the differentiation between two different modes of cognitive processing: the active neurons of the workspace area are in fact, according to the previous quotation, supposed to send top-down amplification signals that increase the activity of hierarchically lower neurons, whose bottom-up signals in turn help to maintain workspace area activity. The establishment of this self-subsistence closed circuit requires a minimum duration, thus imposing a temporal "granularity" to the subsequent neural states that form the flow of consciousness. This dynamic constraint suggests the existence of two boundaries in the processing of human information, a lower level one that originates content and a narrower one, which is a partition of them characterized by the possibility of becoming conscious.

Summing up, also the notion of consciousness as a shared workspace rests on a fundamental principle of modularity of functional processes. This gives rise to a hierarchical differentiation of the contents of consciousness and implies a granular conception of objective and perceived time.

3.2.1 Global workspace theory and active state of representation

According to this theory, several functional modules intersect contents in a global workspace through specific long-range anatomical connections, able to connect distant areas. Consciousness is a content shared among multiple functional modules for a sufficient period of time (needed to reach the level of conscious perception).

> Distributed neural system or 'workspace' with long-distance connectivity can potentially interconnect multiple specialized brain areas in a coordinated, though variable manner. Through the workspace, modular systems that do not directly exchange information in an automatic mode can nevertheless gain access to each other's content. (Dahaene and Naddachè 2001, p. 13)

The possibility for a content to become conscious is subordinated to:

1. Specific long-range anatomical connections, able to connect distant areas.
2. A sufficient period of time needed to reach the level of conscious perception.
3. Attentional focus.
4. It is not necessary that shared information be represented elsewhere, for sharing itself gives rise to consciousness.

In my opinion, points 3 and 4 are problematic. Regarding point 3, an explicit prerequisite for the onset of conscious content is the attentional focus. Attention is here supposed to guide the selection of information from different modules. In this way a specific representation has access to the global workspace. However, this move seems only to shift the problem. In fact it does not clarify what it is meant by attentional focus, what the causal relationship between it and the underlying modules on which it operates is, what kind of neural correlates could be linked to attention processes, what relation there is between attentional focus and consciousness content.

Moreover, point 4 seems to make even more difficult to interpret what we read in the previously quoted passage:

> We suggest that, in many cases, the ability to keep representations in a focus for a period of time in the absence of stimulation seems to require consciousness. By 'in an active state', we mean that the information is encoded in firing patterns of active populations of neurons and is therefore immediately available to influence the systems with which they connect. (Dahaene and Naddachè 2001, p.9)

How can the process of “maintaining representations in an active state” require consciousness, if consciousness itself is supposed to coincide with this process? Here the authors endorse a representational claim: representation is understood as a certain configuration of firing that represents a specific content. This can be traced back to a source (the stimulus), which is assumed to correspond to the specific content of consciousness. If a representation must be maintained active for a certain amount of time, then a neural activity configuration, which represents a specific content, must recur. However, for a representation to exist, it is necessary that it is available to different modules and is sensitive to a top-down reworking mechanism as well. This generates a paradoxical consequence.

Representations, as I pointed out, root their existence on dynamic sharing, so how can they re-present “dynamically as identical” if they depend on the changing state of the whole system they belong to, through the passage of time? This is even more problematic because it seems that this theory is aimed to deconstruct the concept of consciousness-specific content, and to replace it with a variety of different components of different functional systems that make them available in the shared workspace. This problem is linked to another problematic aspect: the arbitrary division, that they postulate, of *five* different functional systems involved in the workspace.

Some developments in neural research, in particular those concerning comparisons of neural activation during passive or active tasks (see for example Cao, Thut & Gros 2017) seem to be coherent with some aspects of the predicting coding framework and, in my opinion, could provide a useful perspective for an alternative approach to these problems.

A different proposal for an interpretation of the phenomenon of consciousness, which transcends morphology of network sharing in terms of a more plastic notion, is the one proposed by Pockett (2012)[12]. This account is centered on the interaction between the electromagnetic fields generated by neuronal firing. The focus of this explanation is shifted to the properties of patterns that contribute to the interactions between electromagnetic fields responsible for conscious experience. Consciousness is affected by the distribution and morphology of the constituent units, but it is not limited to this. This kind of emergentism, more flexible in the determination of fundamental units, focuses more on the undulating nature of the interactions between them, rather than on their molecular biological basis and, although it still has many limitations, it leads us to the central theme of the second part of our critical analysis.

As a conclusion of this first part, we may say that the approaches directly related to the method of neurophysiological enquiry are linked to a discrete conception of either time or neural functional units.

Given the limitations of these approaches, an analysis of the most fundamental properties of conscious experience is needed. Is there something we can borrow from philosophy? Some aspects of phenomenological analyses of perception, for example Merleau-Ponty theory, emphasize the continuity and mutual influence of cognitive processes in creating conscious perceptions, and some aspects of enactivist theories like the one proposed by Maturana and Varela (1980) and emphasize the role of the environment and the embodied and embedded character of any response. The following sections are aimed to explore these aspects and identify other central properties of

[12] In Pockett's theory, the starting point is the physical fact that information is codified through an electrical exchange between neurons. Every electrical current implies the formation of a magnetic field, so around fibers there are magnetic fields that interact.

consciousness On the one hand, the capacity of consciousness of transcending space, time and causal relations will be analyzed. On the other hand, a notion of synchronicity (among internal events, as well as between stimulus and phenomenological response) will be explored.

Concluding, if we take seriously the limits due to the local interpretation and to the temporal resolution of the instruments of the standard neurophysiological methodology and we grant the philosophical teachings in the field of consciousness, we can propose a research paradigm that does not renounce the tools proper to its analysis, but take into account these clarifications. This will be the theme of the second part.

3.3 Learning and perception: suggestion from philosophy of duration

The aim of the second part of this chapter is to consider some aspects of two philosophical approaches in order to see whether they may be useful to interpret physiological data. There are many philosophers who can be inscribed in the phenomenological and enactivist schools. For reasons of space, I review in this section a few paradigmatic examples from either approach.

More specifically, the main goal of Sects. 3.2.1 and 3.2.2 is to focus on the two themes of causal connection and temporality. In order to do this, I will first address the notions of temporality and temporal consciousness as conceived by Husserl, in particular, the concept of duration extension and its relationship to retention and memory. Husserl helps us to delineate, in alternative to the fragmented view of consciousness typical of many neurophysiological models, a unified conception of consciousness, which includes past contents as causally determined in a circle of mutual reconstruction between perception and retention. I will then argue that this perspective might be applied to some neurophysiological evidence about the functioning of memory and the brain regions involved in it. I will discuss this point in detail in chapter 6.

The following Sect. 3.2.3 is devoted to some relevant aspects of the philosophical analysis of perception by Merleau-Ponty (1945). His analysis allows us to avoid separation between the different sensory modes that give rise to conscious perception as a whole. He also analyzes the active role of the environment in our perception, and the enactive approach finds here its roots.

I then consider in some detail the theory of perception in this approach, by focusing in particular on some contributions by Noè (2004), which may turn out to be useful in reinterpreting neurophysiological computational theories of perception. This aspect will be analyzed in Sect. 3.2.3.

3.3.1 Extended present

The traditional model of time assumed by neurophysiological explanations, considers time as an enigmatic, invisible entity, that extends above and beyond objects. This entity possesses a punctual "here and now", a sort of moving cursor that flows from instant to instant. As a result, everything present must be compressed into such an instantaneous snapshot, devoid of width. This view contrasts with the way in which our brain works, as well as with our experience, which is not narrow and instantaneous at all. It is also very difficult to determine the position of this kind of present on the time line.

In contrast to this model, some theories of mind consider change and becoming as the basis of reality. The present moment, in this kind of approach, is considered to be an entity held together by causal processes. The present moment is not a container but, as suggested by Manzotti (2014), it widens to incorporate everything we have experience of, and this is consistent with the fact that our experience extends through time to include past events.

So, what is the extent of the present if it is not instantaneous? It has a depth that varies depending on the object on which we focus attention. As my gaze glides across my field of vision, temporal depth varies, sometimes slightly, but sometimes in a drastic and sudden way. No metaphysical gap keeps past, present and future separate, like no metaphysical gap separates a nearby object from a distant one, both are parts of our present.

Past and present are practical concepts based on the conditions of the subject; something is considered part of the present if we can interact with it, and so past is part of the present. This is demonstrated by the fact that we have experience of the past, which in turn has causal power over what we consider to be our present.

The first step we take, is to describe the properties of conscious experience as they present themselves to the subject. This type of analysis is the very method of phenomenology. If it is true that experience can be misleading, however, such unreliability only emerges in comparison with intersubjective reality. We can explain the nature of an illusory perception by a deductive process, based on our stronger intersubjective beliefs that contradict such a perception. But, from a purely subjective point of view, hallucinatory experience is indistinguishable from veridical one. It may have different or unusual features, but it shows up to the subject with the same evidence. The first feature that I would like to examine, because it is particularly important for the purpose of this analysis, is the relationship that conscious experience has with time.

3.3.2 Husserl and the Extended-Duration Structure

When phenomenology refers to temporal consciousness, it implicitly takes it as something specific. In it and thanks to it we build our concept of time. This intuition can already be found in Agostino (398) who, in his *Confessiones* written as far as 398, talked of time as a "distention of spirit": time exists relatively to consciousness. Husserl takes up this aspect in his famous text *Phenomenology of internal time consciousness* (1893). Time is equivalent, for him, to temporal (non-individual) consciousness.

The connection of different forms of objectivity with different forms of temporality is built with an emphasis on the contribution of temporality with respect to the problem of identification. Different regions of being – real, unreal, ideal – are correlated with different forms of temporality. The phases of the temporal object are described as 'modes of being' of a noematic object, that is, of a phenomenological object taken in its objective configuration. Objective time, that is, the time of the clock, presents a rigidity in its temporal order that is not present, past, or future from an ontological perspective.

Husserl insists on a kind of original dynamism of time, which requires us to conceive the present not as a static point, but as an act characterized by a constant flow. Consciousness is made up of phases, defined by Husserl with the terms of *impression,* *retention* and *protention*. *Retention* is the notion by which Husserl tries to capture the permanence of a newly-perceived within the current perceptual phase. His famous example (Husserl, 1913) to explain this point is about music: the melody is composed by several notes (A–B–C–D–…), performed sequentially. From a phenomenological point of view, our perception of it will begin with an original impression of the first note (A). The impression is initially present in person or in the flesh, that is, in a living way. However, it soon leaves the place, in the next phase, to the note (B) impression. While we perceive in person the note B, the note A does not disappear at all, indeed, it resonates again. Although it has passed, the original impression of A is maintained in the form of retention (Ar). This mechanism, according to Husserl, ensures the presence of the past in the present, constituting the perception of an ongoing unity in experience. In the next phase, with the appearance of the impression of the note C, we see a further change in consciousness. There will be an original impression of the C note, present "in person", and a retention that includes the previous steps, which will be perceived as even more distant in time. It is, to use a metaphor by Husserl, a kind of echo.

But how do you define the identity relationship between retention and the original impression, and what kind of differences are there between retention and memory? If we look more closely, we find that retention is distinguished by the following characteristics:

- Unlike remembrance, it is not an act. Instead, it is a constitutive phase of an intentional experience (along with impression and protention).
- It is necessarily passive.
- Retention does not lead to reification. It does not return an object to us in the past, but it rather contributes to the creation of both, the object and the present.
- It is the presentation of a phase (the impression) that has just passed, as it has passed.
- Retention, like all phases of absolute consciousness, is not temporal. Although it constitutes time, according to Husserl, it is not *in* time.

Each of these features would require a further and deeper investigation, but I want to focus on this aspect: retention, as a phase of absolute flow, is a real perception of the past, which is presently *given* in it. In retention, the past can be, in Husserl's words, “directly looked at” (Husserl 1893), for retention holds back in the present the past just spent. The past, in this case, is not the result of the apprehension of a temporal content. It is instead a content itself, if we can say so, which is *the* past. Husserl seems to explicitly conceive of the past as a way of original presentation. The role of retention, in this context, is to ensure a conception of an authentic past. Retention has the explicit function of preserving the same in the other structuring the continuity needed to perceive events consciously.

The *current now* memorably encompasses a *continuum* of the past. Each new now once again encompasses this continuity of remembrance, the new now encloses the new continuous memory, and so on. Memory, intended as an active reminder of a past event, is seen to play a key role in the identification processes. If there were no remembrance there would be no object for the self, who would certainly lack the consciousness of what can be grasped in multiple acts, something to which one can always return, and which can always be recognized as the same. The causal power of the past is not seen here as a close event that causes close events in the present. The whole past is materially in the present. The present is a collection of things, all present as causes of joint effects. Present and past are relative and extended concepts.

3.3.3 The dynamic structure of perception

The extended concept of present, which concerns time consciousness, has interesting implications for an analysis of perception in general. I will start with Merleau-Ponty's philosophy, which represents an approach that has been the subject to countless interpretations and commentaries. Some have called him a "structuralist" because of his use of the concept of "structure"; others have called him a "phenomenologist", because he was identified as the worthy French heir of Husserl’s phenomenological school. But in his work there is a very interesting fusion of the two currents, so I will talk a little about both.

The theme that runs through all the work of Merleau-Pointy is the relationship between the organism (understood as the body of the individual), and the environment that surrounds it (understood as the world in which the organism is). This relationship is understood as a relationship between consciousness and nature, between the subject who perceives and the perceived world. By virtue of the fact that there is a living body that perceives, perception becomes the mechanism that immediately connects consciousness to the world, preceding every form of scientific objectification.

The body has a central place in the development of Merleau-Ponty's concept of perception. In *Phenomenology of Perception* (1945) the author dedicates a conspicuous part to the body and seeing. He writes:

> The own body is in the world as the heart in the organism: it continuously keeps the visible spectacle alive, animates it and feeds it internally, forms with it a system. When I walk into my apartment, the different aspects under which it offers me could not appear to me as the profiles of the same thing, if I did not know that each of them represents the apartment seen from here or from there, if I had no awareness of my own movement, and my body as identical through the stages of this movement. (Merleau-Ponty 1962)

It is the body that keeps the show of the world visible, in the sense that the broad process of perception happens in it, and the subject gives a meaning to real things through it. The basis of every knowledge is precisely *embodied*. Through our body the cognitive power of the perceptual world is established. The body is the instrument through which we are in the world, we perceive and know it. (A similar concept can be found in the Heideggerian notion of *Leib*. In Merleau-Ponty the body represents the synthesis between subjectivity and objectivity, and it thus acquires the status of an ambiguous entity. The body, understood for example as a hand that touches, represents both sides: the subject that feels touching, and the felt object that is touched.

Moreover, all the abilities that we have in our body (hearing, seeing, touching, etc.) cooperate to give rise to a conscious perception, but they do not remain differentiated in that conscious perception. Every perceptual content is strongly modulated by the situation in which we are, and by the cooperation of all the involved parts of our body.

This aspect is interestingly investigated by Alva Noë (2004)[13], whose theory assumes an enactivist perspective. This perspective highlights that we are familiar, as a general rule, with the ways our sensory experience changes as we move. Moving the eyes, blinking, turning the head, or moving the body, we produce familiar kinds of sensory change. Perceivers know how to gain access and make contact with the environment around them.

This is the key to the problem of perceptual presence: our sense of the perceptual presence of the detailed world does not consist of our representation of all the details consciously present. Rather, it consists of our present ability to access all details, and of our knowledge (itself practically embodied) that we have this access. This knowledge takes the form of our comfortable mastery of the rules of sensorimotor dependence that mediate our relation to the world. In Noë's words[14]:

> My sense of the presence of the whole cat behind the fence consists precisely in my knowledge, my implicit understanding, that by a movement of the eye or the head or the body I can bring bits of the cat into view that are now hidden. (Noë 2012, p. 68)

3.3.4 Perceptual liquidity

To sum up, we have seen that, according to Merleau-Ponty's and Noë's accounts, perceiving is an embodied relationship to the world. The body and its movements are strongly involved in the generation of perception. Suppose we perceive an object and we want to reach it. We could explain this situation in a naive way, by saying that we represent the object as the target and the action is the pathway to achieve it. But if we take seriously what we said about the enactivist and phenomenological perspective, and also with respect to the role of the very body in every perception, it is possible that the difference between

[13] Noë A. (2004). *Action in perception*. Cambridge, Mass: MIT Press. The strong involvement of a bodily component in every perception contributes to make the boundaries between action and perception fuzzy. This requires to conceive a new stratification of presence.

[14] Noë A. (2012). *Varieties of Presence*. Harvard University Press. The presence of objects and the representation of specific targets is always relative to our position and more strongly to our attitudes and competence to manage sensory motor dependence, formed during our life in synergy with our environment.

the target and its achievement (or a learned output as well) becomes more faded. If perception already contains a motor component, which provides meaning to the perceived, then the motor achievement of the target is a change in the body state that simply involves the activation of a bigger or stronger motor component.

There is no time contrast between discrete states on a timeline and a fluid structure that contains different possibilities of action. We have seen that a central aspect of the phenomenological reflection on temporal consciousness is rooted in an equally dynamic and extended interpretation of perception.

Perception is a complex structure that hosts in itself more or less extensive and intricate retentions and remembrances that are causally responsible for the onset of perception as it is given to the subject. And the bodily dimension is essential to scaffold perception.

We can in fact use this approach to reread the mechanism of perception at the neurophysiological level. The stimulus, from this perspective, is not a determined and measurable reality that encounters an equally determined structure (in the nervous system) which, by virtue of its complexity, performs mechanical computations and generates an output. Rather, the stimulus is a dynamic structure which encounters another highly dynamic structure, already constantly active and in constant dynamic relationship with the circumstances in which it is in.

What becomes relevant, is to explore this dynamic through a specific and controllable stimulation. When we hear a sound, it is not just a sound wave that hits a specific portion of a system. It is rather the result of the encounter of the sound wave with a strongly plastic structure that is already active.

The next part of this work is expressly dedicated to these reflections, in order to enrich the notion of this intrinsic dynamics by taking into account some neurophysiological evidence about anticipation perception and memory. More specifically, I will analyze the mechanisms of stimulus expectation, and the one of the perceptions of expected, but missing stimuli. This analysis implies a discussion of the notion of memory.

Chapter 4

Ongoing activity of the brain:

perception between prediction and errors

If we take seriously what we have said in the previous part and consider the phenomenological properties of consciousness, we should also consider expectations about future, memories, and their affective values. Our perception of what we usually call present is structurally extended and composed of different degrees of existence and presence. One of the most extraordinary abilities of the human brain is that it is constantly imagining the future. This means two different things or, better, two different degrees of the same thing. First, we figure out possible conditions, expected because we have an available set of data that could be coherent with a set of outputs. But also, on a finer scale, during the performance of an action, it seems that we constantly form a model of the world, and our perception tries to anticipate the perception that will be connected to a specific change in the conditions of our activity.

Carey, Tanaka and Van der Meer (2019) maintain that, in the field of computational neuroscience, the opposition between behavioral and computational explanations of the mind presents itself in the new form of a model-free vs. a model-based approach. The first one is exemplified by *temporal difference reinforcement learning* (TDRL). It explains the activity of the brain as the resulting learned output which is automatically produced by a simple comparison between different weights that are put on some elements of previous situations.

The network assumed by this framework is composed by certain targets and possible actions in order to reach them. There are different possible actions that can be taken to reach a target, and there is no previous model of the world that guides us in performing an action or choosing the best one. Everything is a mechanical sequence of events, part of them are outside the boundaries of the body and part of them are inside, the encounter between them either produces a reward and a consequent ascription of a good mark, or generates a pain, or some unpleasant sensation, with a consequent ascription of a negative mark. In this case, a prediction error is sent back to the upper hierarchical levels, in order to correct the values of the elements and consequently reshape the network. Thus, according to the TDRL account, during experience the individual system associates a mark to one of the actions we perform; the more positive the mark, the more effective is that specific action in bringing a reward.

In this model, there are also backward connections that inform the previous step about the mark value, in order to reshape and reorganize the network in the better way. This model is good enough to describe simple situations, but it shows a limited explanatory power to approach more complex activities of the brain, especially to cover situations where different choices are on a par, or when the subject changes her mind during performance of the chosen action.

A more recent and complex pattern of explanation is the model-based one: *planning with transition model* (BWTM). This framework approaches the brain as if it is a predictor; this means that there is a previous model of different expected situations connected to the performance of the possible different actions. The subject has foresight, in the sense that during performance she considers, not necessarily in a conscious way, possible states of affairs as possible consequences of her choice. This model focuses on the ability of the brain to perform different actions with different degrees of effectiveness. The learning process, under this perspective, is not concerned with a comparison of state values and the application of an error prediction rule, but it is about the strength of the connection between subjective actions and consequent states of affairs. In this way, the model is able to predict possible action outcomes before actually performing them. This idea is the basis of the predictive coding approach, which we are going to discuss in the next section.

4.1 Mind and predictions

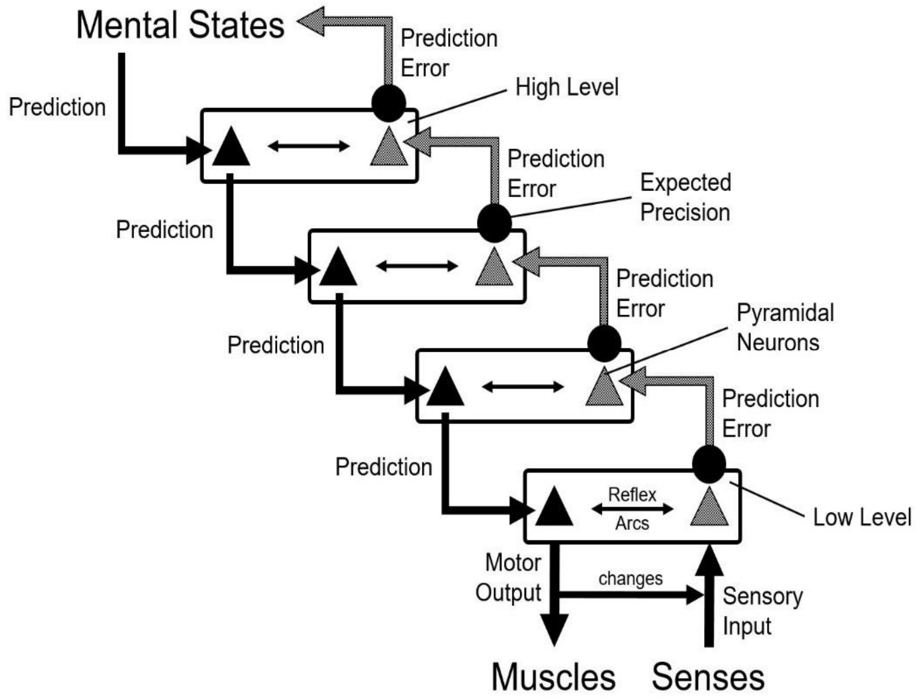
The predictive coding approach is a framework for explanation of mental activity that takes as its starting point a specific aspect of our daily lives: the effectiveness of our actions seems to require anticipation of those world conditions that different possible actions will produce. Under this perspective, in order to act and reach a desired target, it seems that something that drives our movements is anticipated by predictions of our brains. When we are thirsty and we move in the direction of a glass of water, the brain not only responds to a stimulus, but it uses previous information about the position of the body, the environment and its features, as well as relevant laws, to model the possible trajectory in order to perform a successful grip. The same happens when we walk. The brain evaluates the structure of the pathway and the relative positions of legs and body to predict, at every step, when the foot reaches the ground. If the prediction fails because we encounter an obstacle, the brain immediately activates to correct the unexpected situation and reorganize the movement.

The brain is constantly predicting the future to compare it with the past, and it uses old experiences to anticipate what is supposed to come next. For decades, in a computational perspective, it has been believed that the brain processes inputs by reacting to stimuli that comes from outside reality through the sense organs. In the last two decades, however, the brain has started to be understood not in its reactive action, but more in its predictive internal ongoing and creative activity.

Prediction is intrinsically linked to the reward system, as we will see better in the next chapter, where the role of pleasure in the construction of perception will be examined. One of the most important neuronal mechanisms that influence behavior is studied through activation of the dopamine circuitry. Metabolism and correct levels of dopamine are able to modify the effectiveness of a learning process based on reward, as shown in the opposite cases of turettic patients and parkinsonian ones, as I will explain in more detail later. On the behavioral and the practical side, one of the most evident implications of prediction involvement, is Pavlovian conditioned learning: the connection between stimuli, pleasure and behavior proposed by Ivan Pavlov when he announced in 1903 that he had discovered the conditioned reflex in dogs. But, at that time, he could not know that he had actually anticipated the discovery of dopamine circuits, which occurred only in 1958.

Subsequent studies of primates have found that once they have mastered the mechanism for obtaining food (e.g. pressing a certain number of times a button) the dopamine discharge arrives before the monkeys press the button in question. Contrary to what was once thought, the reward is not at the end, but before you take the action. Computational neuroscience, which studies brain functions in terms of statistical data processing, considers the brain as a Bayesian machine that produces constant inferences about the world and readjusts them based on actual sensory perceptions. The neuronal prediction that acts without our knowledge, however, needs another important resource. To compare the future and the past the brain also needs an embodied memory.

One of the central ideas of the predictive coding approach, is that the brain is constantly involved in an ongoing activity of generation of models. These models are constructed on a set of incoming variables. The one that will be used to generate a conscious content of a perception and to perform the action, is the one that results compatible with the highest number of incoming variables. The brain works on an internal, or generative, model of the world to actively predict incoming information. The deviation of this predicted information from actual information, the prediction error, is used to drive the inferential process towards a best estimate of the state of the world in which the subject is involved. Numerous algorithms have been proposed for predictive coding, which differ in fundamental aspects such as the form of the generative model, the criteria used to drive inference, and the nature of the information being transmitted for further processing. There are also many hypotheses about the neural implementation of predictive coding.



Fugure1 the picture represents a schematic hierarchical bidirectional cascade where predictions move from high layers to the inside f the brain and prediction errors move from lower layers to update following prediction

4.2 Direction of prediction

Whereas some models suggest a bottom-up form of predictive coding, one of the most influential forms of predictive coding proposes a hierarchically organized system in which predictions are generated at higher levels to be fed back for comparison with inputs at lower levels. After this comparison, processing of correctly predicted information is suppressed, whereas prediction errors are transmitted to the next level for further processing and to drive inference in order to achieve the overall goal of prediction-error minimization. This has been the focus of some recent research by K. J. Friston (2012, 2017)[15] Although suppression, or dampening, of correctly predicted information is emphasized in this model of predictive coding, recent psychophysical evidence suggests that prediction can serve to make the perceptual representations of predicted events more distinct than those of unpredicted events. This sharpening is not necessarily inconsistent with predictive coding theory, which suggests that activity in units that generate predictions rapidly converges on a fine-tuned representation. Nevertheless, other computational approaches that include predictive components place a stronger emphasis on sharpening, and might prove to be useful in helping to explain it. Neuroimaging and electrophysiological findings are not conclusive, as some studies suggest that prediction leads to dampening of predicted information, while others suggest it leads to sharpening.

4.2.1 A possible interpretation of predictive coding: two opposite prediction streams

One of the most interesting interpretations of predictive coding is the one proposed by Teufel and Flecher (2020). They try to give more relevance to bottom-up processes, by differentiating between general prediction and specific prediction, or prediction about general laws, which can be applied to different situations, and predictions about details of a specific context. Both structures indicated by these authors, are dynamic and always active.

Predictive coding, for them, has to be integrated with an extended view. It is the working process of the brain and every specific functional system that contributes to updating the ongoing model with its outcoming data but, from their perspective, this process is not unique, as it is made of two different processes that capture different features of the environment and occur through different cerebral computations.

[15] Friston has proposed a philosophical approach on the predictive activity of the brain, and he has tried to find empirical evidence for it by studying brain functional circuits and their properties and effects on subjective conscious perception and action.

Even if I think that this interpretation is not enough powerful, because the boundaries between the two processes are not easy to define and also because it is liable to the general critics that it is possible to address to many other interpretations of predictive coding, critics that I will explore in the next section, I think that it is worth discussing some crucial points of this approach in some depth.

Teufel and Flecher (2020) argument is simple: critical design features of an agent's information processing mechanisms can be understood by observing the structure of his/her environment. This basic idea is not new and has its roots in an 4E approach to cognition, which emphasizes the role of both body and environment in cognitive processes. In my view, an application of this principle to the role of prediction in brain functions points at an innovative and useful perspective. To achieve a deeper understanding of the brain as a "prediction machine", they suggest that it is necessary to recognize that prediction in the nervous system comes in different forms and is about different kinds of regularities. As they explain in their latest work:

> Regularities of an environment are mirrored in corresponding context-independent predictive mechanisms that act on bottom-up processing, mechanisms that we have referred to as constraints. Likewise, fluctuating, context-dependent regularities in the world point towards flexible, context-dependent predictive mechanisms implemented by top-down processes, which we refer to as expectations. (Teufel, Flecher 2020, p. 240)

They argue that, even if there are several possible distinctions in the kinds of prediction that the brain activates, which could be useful to be drawn in order to explore different faculties, the distinction between context-independent regularities and context dependent ones seems to be primary and fundamental.

The very interesting point, that I would like to highlight here, is that this approach presupposes a distinction between singular events and more complex ones. The causal connections that characterize the environment and the laws that the subject perceives to govern it seem to be processed in different ways. The object is also constituted by elements causally connected in a specific way. The identity of a perceived object comes out from this organization. An object, a subject, or an event are all dynamical structures characterized by specific causal relationships, internal to each of them. However, it seems that this continuity must stop at some point, in order to describe two different pathways of computation that the brain performs, two streams about two different classes of environment features.

Although we have explored these ideas mainly with respect to neural information processing, a comprehensive view of the role of prediction in information processing must ultimately include an appreciation of the whole organism interaction with its environment. Predictive information is present not only in neural mechanisms, but also in the morphology of the organism as a whole, as it is illustrated in numerous examples of sensory and behavioral ecology.

Ultimately, the predictive processing framework should aim to incorporate the many different ways in which prediction is part of biological information processing, in order to offer more comprehensive insights into how we interact with our world in either healthy or pathological conditions.

4.3 Some critical remarks

In this section, I explore some aspects that I find problematic for a predictive coding account. In particular, I will consider some difficulties in the definitions of degrees of prediction and priorities (Sect. 4.3.1), and the implications of prediction for a representational model of the mind (Sect. 4.3.2). I will then briefly review the debate about the role of prediction error and its connection with erroneous perception (Sect. 4.3.3). Finally, I will discuss the notion of error, partiality and the perception of unexpected stimuli, and I will highlight the distinction between a predictive computational view and a different philosophically based approach to brain prediction (Sect. 4.3.4).

4.3.1 Degrees of predictions

One of the fundamental problems that predictive coding leaves unsolved is the criteria used in the division between predictions, and the reason why we isolate a particular content. More precisely, it is not clear what the relationship is between the procedural sequence of actions and their purpose, or where the boundary is between the trip, the performance of an action, and the goal. It is true that the mind presents itself, from a phenomenological point of view, with a strong imaginative activity that could be about states of complex and future things, as they were crystallized in a time that is only a name or a label. But there are also predictions about the near future and simpler circumstances of action. For example, I can imagine (and desire) to play the piano in a classical concert, and I can enrich this imaginary scene with lots of details. But also, I can imagine and predict many actions necessary to achieve this goal: I should practice in a proper way, study several hours every day, and postpone many other desires that compete with the goal of playing in a piano concert. Moreover, on a finer scale, during my piano practice, I have to predict every movement of my fingers on the keyboard in a procedural sequence of actions; I call these kinds of predictions *procedural prediction*s. This complex dynamical relationship between different degrees and durations of predictions, what prediction exactly means, and how a prediction error could be defined for each different kind of prediction, are all crucial points that remain unexplained by the predictive coding account.

The predictive coding account describes, through the notion of error, a discrepancy between a data network and an inconsistent perception, but it does not explain how the subjective system ascribes priorities to one choice at the expense of another. How do you choose to organize the next actions in the actual dynamics that leads to the realization of a goal? Many mental illnesses might be linked to problems in the ascription of desirability to certain actions or could be seen as involving a problem in the distribution of priorities and in understanding the causal link between them and the achievement of the target.

Here, the crucial question becomes: Why do we delay certain needs taking some actions instead of others? How do we differentiate between what is necessary, or is not, in a sequence of actions? This topic could belong to the mereology of action, or to the analysis of the relationship that exists between a specific event, which is the target, and a sequence of actions with different degrees of relevance, that are performed in order to achieve that target. How do we set the scale of our priorities? How do we define the boundaries between movements that compose actions directed to something else and movements that coincide with a supposed represented target? This is a crucial point in the present research about learning because it is a process that, through an appropriate action makes it possible to reach a target.

By analyzing these aspects more deeply, we find again the preponderance of timing and temporal relationships: predictions about future circumstances and procedural predictions differ in their timing scale and their role in timing organization. The predictive coding approach, in many of its varieties, seems to identify a granularity of imaginative experience and predictive activity, which is difficult to inscribe in an ongoing activity, thus being subject to the same problems that I mentioned for other theories in Ch. 3. And the aim to grasp the complex dynamics of perceptive processes, typical of this approach, seems difficult to be achieved if brain workings are analyzed only in terms of predictions and error signal propagations.

4.3.2 Predictions and representations

One criticism that is strongly connected to the previous one and can be addressed to some interpretations of predictive coding, is that talking about predictive processing implies the existence of a kind of mental representation. This, as earlier argued (Sect. 4.3.2), involves a granular view of experience, which, in my opinion, is not adequate for a comprehensive account of perception. In fact, if this working mode of the brain is based on model construction and its constant updating with incoming data, there are two main interpretations of this process, which are then declined with different nuances, which in turn give rise to different readings of predictive coding.

The first one ascribes prominence to top-down processes that are responsible for a cognitive evaluation of incoming data. This evaluation is about the coherence between incoming data and previous acquisitions and it aims to minimize prediction error.

The second interpretation highlights the relevance of the bottom-up processes and the role of peripheral and unconscious computations. Under this perspective prediction error is the most relevant part of what becomes conscious. Obviously, there are many other interpretations of this framework that try to put together top down and bottom-up processes, which are always dynamical and simultaneously active. We saw an example of this kind in Sect. 4.2.1.

Moreover, a predictive processing framework describes brain activity as being organized as follows: the brain has a perception during a temporal interval (this interval may have a variable duration, depending on context and attention) which generates a representation of a target and, at the same time, an expected perception (coherent with the available dataset). A computation compares the expected perception and the incoming flow of information and thus generates an action, which contributes, from time to time, to the refinement of the prediction algorithm associated to that performance sequence.

The essentially dynamic structure of this paradigm of explanation, in some interpretations, is still tightly linked to a representational conception of the mind, which however seems to contrast with its foundational assumption. The dynamic predictive structure of the brain activity, in fact, characterizes the very object of perception, and also the target of a possible action. Perception, for the predictive approach, constantly depends on a comparative evaluation between models with different degrees of adequacy with respect to incoming data, so that enucleating a representational content that corresponds to a specific and static content seems contradictory. It is true that, in general, a dynamical system may involve representations. But the kind of dynamics described by the predictive approach also involves perception.

Perception is a construction that comes from a constant ongoing activity, so it is constantly changing, as it always depends on incoming data. Hence, it is difficult to establish the boundaries that identify a stable and identical content, and this seems to be inconsistent with a fixed representational content of perception.

4.3.3 Prediction error and erroneous perception

Finally, the most urgent problem for this perspective is the case of perceptions that radically differ from previous predictions. Prediction error plays a fundamental role with respect to the contents of consciousness, but it does not exhaust them. Predictive coding is not able to provide explanations for many situations in which the subject is overwhelmed by totally unexpected perceptions. By this account, for example, we will be able to explain a simple walk by saying that our brain predicts the trajectory and position of our body in the environment and activates motion sequences that are consistent with incoming data about the ground morphology. But what about situations in which our perception is subverted?

If we walk barefoot on a ground exposed to the sun, our perception is completely absorbed by the burning perception produced by the contact of our feet with the searing ground, which instantly suppresses any other content of consciousness. All the data necessary to form the prediction consistent with a searing ground were there, so why not to expect that perception? And how to account for the different relevance of the prediction error?

The problem here is that predictive coding not only accounts for the brain's prediction of possible future state of affairs in order to organize our actions, but it also explains the very content of our perception in the present moment, not about what I expect in the future. If conscious perception is the effect of a comparison between a group of perceptions, predictive coding can't provide an explanation for surprising contents of the present perception.

Moreover, every set of incoming data generates multiple prediction errors with different degrees of relevance, but here there is not a conclusive answer for the reason why one results the winner. The winner is the most coherent with the majority of incoming data, but the very set of incoming data is the result of previous comparisons of the same kind.

There are endless discrepancies between what can be considered a small prediction error that, within a set of consistently predicted data, results in inconsistent retracing the hierarchical layers to alert the consciousness that sends a correction response. However, in many situations one is completely overwhelmed by a perceptual flow. I therefore find that the predictive coding model is properly descriptive of certain situations, but the explanatory power seems weakened by some of its implications.

4.3.4 Error, partiality and the unexpected

Predictive coding involves a hard, constant, and precise work of considering different features, comparing their weights to previous or future state of things, and deciding step by step which model is the one that ensures the most adequate perception.

As we have seen above (Sect. 4.2), in the predictive coding perspective, higher areas of the cortical hierarchy are involved in generation of prediction, cognitively informed by previous experience. Top-down signals are in this way sent to networks of lower levels, where they induce an expected pattern of activation. The lower area receives sensory input from the unconscious network and confronts it with the expected activity pattern constructed before. Any mismatch between the predicted pattern and the one evoked by the incoming input is sent to the higher area as a prediction error. The ongoing model is constantly updated correcting errors of discrepancy. The same procedure is repeated in multiple hierarchical cortical levels, in a circularity of prediction errors and corrections, and as suggested by Friston (2005) and later by Clark (2016),[16] this could be identified as the general neural coding strategy.

The notion of error is the most relevant aspect, in my opinion, in the distinction between the predictive coding approach and some philosophical perspectives, in which the error is considered to be a different thing and play a different role. Prediction error, in the predictive coding framework, is a discrepancy between perceptions, the incoherence between the set of incoming data and the expected perception due to the model.

As we saw in the previous chapter, in the phenomenological approach, error is interpreted as a structural part of the whole perception. We have a relative perspective and, in order to become perceivable, something has to be hidden in some of its parts. So, strictly speaking, there is no proper error in our perception, but only partiality. Regularities that govern the world appear through the relationship to what is hidden or undisclosed in the perceptual subjective universe. This way of interpreting the error of perception leads

[16] Clark A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind.* Oxford University Press. In this book, Clark extends the predictive coding approach to an embodied and situated theory of mind, which explores the possibility of action and describes the mutual reconstructive dynamics between perception, action, and environment.

to a different model of the mental and brain activity, where erroneous perception is integrated in a complex structure of different sensory modalities.

In sum, I have briefly presented the main characteristics of a framework in which predictions are coded by selectively modulating the neural units whose response properties match the predicted characteristics of the stimulus, and I have shown some limits of this framework. The final consideration I want to propose here concerns a doubt on how we process and manage unexpected stimuli and situations. This will introduce us to the topic of the next chapter.

Humans are able to make unspecific predictions, knowing that something is going to happen now, but not knowing exactly what it is. For example, we can certainly notice the difference between the adequate termination of a song and its undue interruption, even if we don't know exactly how the song would have continued otherwise. This kind of prediction is difficult to reconcile with the predictive coding models: If we do not have a specific sensory representation to predictively activate, then how do we predict?

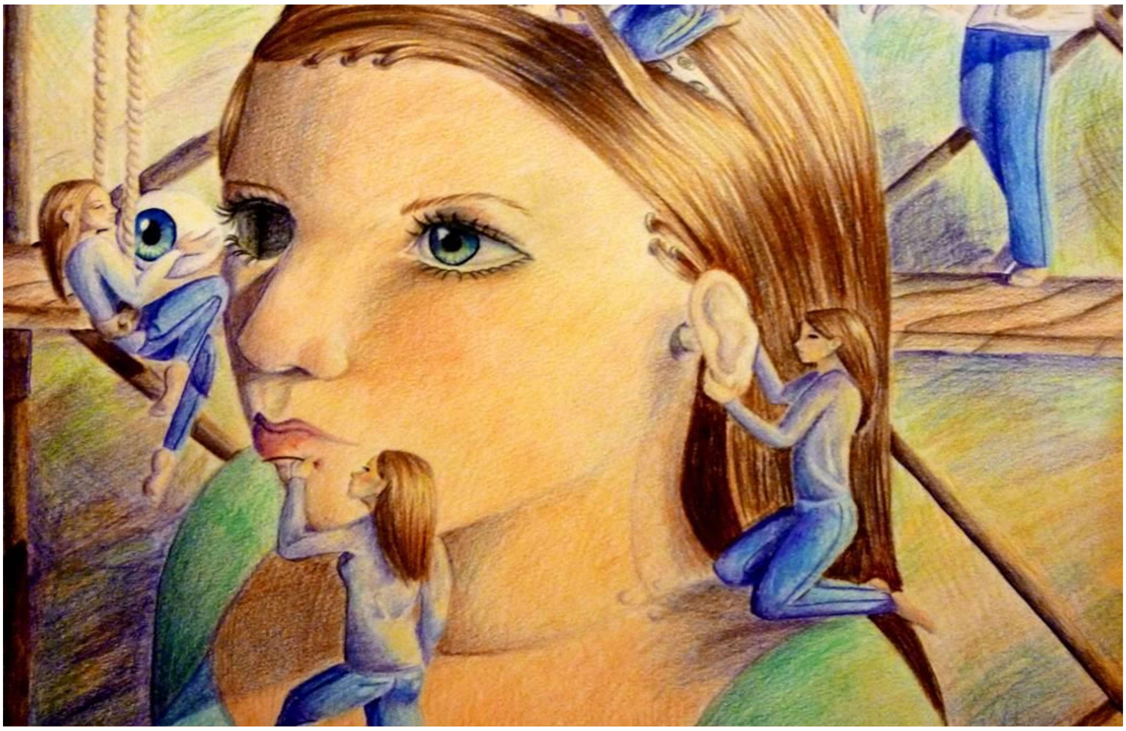
There are different features in the environment that are relevant for the subject in a specific moment and context. It is very hard to explore the relationship between interest for specific features, attention process and the updating of the ongoing model. It is hard to understand how a specific functional sequence could be modulated in different contexts. A similar problematic aspect was found earlier (Ch. 3, Sect. 3.1.2) talking about workspace theory.

In my opinion, and for the purpose of the present work, an interesting aspect of predictive coding is that it focuses on the relevance of what is unexpected. It is exactly what differs from what is expected that passes through hierarchical levels, arrives at being conscious, and has the strongest impact on future model-based predictions.

This aspect is very interesting because it has important connections with unexpected stimuli, and the way the brain processes them on the basis of context dependent time scales and relevance scales. Brain activations coherent with unexpected stimuli testify to the existence of an internal and subjective timing processor that is always involved in every computation and perception. This is the tool that the brain uses to establish connections between repetitive events and their distance in time, and to formulate hypotheses on future happenings. In fact, this is one of the most powerful and interesting aspects of the brain workings. It is fundamental for memory perception and imagination, and its connection with self-construction and perception will be the main theme of the next chapter.

Chapter 5

Self between construction, intuition and timing



This part is dedicated to the notion of self. Starting from an analysis of interoceptive interference, I discuss a predictive approach to this notion and I highlight some evidence in its favour, as well as some limits. Then, I explore the possibility that this predictive coding framework might have some methodological roots in German idealism. I consider some aspects that Fichtian idealism shares with the constructive likelihood principle. Specifically, this will be argued with respect to Fichte's first and second dialectic step. I will then explore another methodological base, which idealism, in its third steps, offers to the development of enactivism and 4E cognition theories. With these tools, I explore the theme of reflexivity, or the role of action in self-perception and identity ascription.

As we saw before (Sect. 4.2), the concept of the brain as a predictive machine employs Bayesian tools. This means that the brain activates a cognitive evaluation of circumstances through a comparison between previous similar conditions, incoming data, and probable states of affairs. We saw predictive processing theory generically and primarily applied to exteroceptive perception, because it is the most studied. But one interesting and more recent application, especially to self-construction, is the one proposed by Anil Seth (2013).

5.1 Predictive self through Interoceptive Perception

Seth describes a predictive, inferential perspective on a particular kind of perception: the one that is about our bodily states as integrated in a unity and continuity. This kind of perception seems to be primary for any other kind of perception and it is the most complex and challenging because it is the most composite. In fact self-perception is in part interoception, which is a bottom up process that proceeds from internal regions of the body to cortical networks, and which constantly updates information about the internal state of the body, what is happening in it, and what is good or not. But with this term we also refer to the sense of self between agency, sensitivity, and passivity.

In every perception interoceptive inference is always active. This is a deductive inferential process, which, by taking incoming data as premises, derives a possible ascription of value to some events, and thus predicts a possible coherent perception. From this point of view, subjective feeling states (emotions) are not primary and immediate, but they arise from actively-inferred generative (predictive) models.

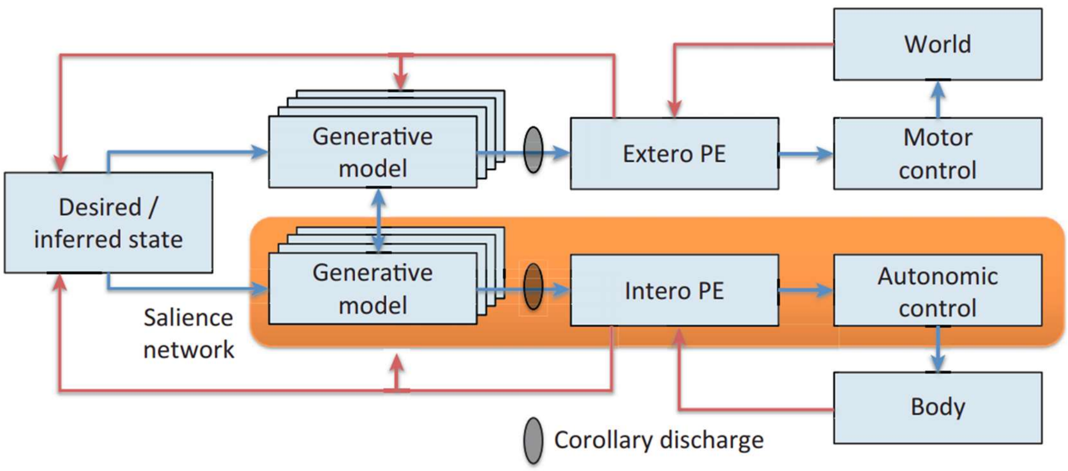


Figure 1 Model of interoceptive inference proposed by Seth A.
In the model, emotional responses depend on continually- updated predictions. Prediction errors (PE) could be interoceptive or exteroceptive and contribute to adapt adequate responses.

Emotional responses, under this perspective depend on continually-updated predictions of the causes of interoceptive input. The first element of the model is a desired or inferred physiological state and is itself subject to updating, based on higher-level motivational and goal-directed factors. From this element generative models are engaged, which can predict two kinds of signals, interoceptive signals or exteroceptive ones. Applying active inference, prediction errors are transcribed into actions via engagement of classical reflex arcs (motor control) and autonomic reflexes (autonomic control). The resulting prediction error signals are used to update the (functionally coupled) generative models and the inferred/desired state of the organism. Interoceptive predictions are proposed to be generated, compared, and updated within a salience network (Fig 1, orange shading) anchored on the anterior insular and anterior cingulate cortices, which engages brainstem regions as targets for visceromotor control and relays of afferent interoceptive signals. Sympathetic and parasympathetic outflow from these brain areas are in the form of interoceptive predictions that enslave autonomic reflexes (e.g., heart/respiratory rate, smooth muscle behavior), just as proprioceptive predictions enslave classical motor reflexes formulations of motor control. This process depends on the transient attenuation of the precision of interoceptive (and proprioceptive) PE signals. Blue or red arrows signify top-down or bottom-up connections, respectively. (Seth 2013)

The model generalizes *appraisal theories* (Clore and Ortony 2008), which view emotions as emerging from cognitive evaluations of physiological changes, and it sheds new light on the neurocognitive mechanisms that underlie the experience of body ownership and conscious selfhood, both in health and in neuropsychiatric illness.

> Emotions and embodied selfhood are grounded in active inference of those signals most likely to be 'me' across interoceptive and exteroceptive domains. In humans, self-related predictive coding simultaneously engages multiple levels of self-representation, including physiological homeostasis, physical bodily integrity, morphology, and position, and – more speculatively – the metacognitive and narrative 'I'. Subjective feeling states (emotional experiences) arise from active

> interoceptive inference, extending previous theories based on cognitive appraisal of perceived physiological changes and contemporary frameworks that emphasize bottom-up elaboration of interoceptive representations with perception and motivation. (Seth 2013, p. 570)

5.2 Likelihood principles for ascription of identity

The most interesting point in the previous quotation that I would like to discuss is that self is a construction with a probabilistic structure. This is very different from theories that take the self as a primary evidence, contraposed to the outside that is the object of perception. In between there are several possible data, some of which are more likely to be considered part of the self.

The mind is a machine that is completely in the dark with respect to the outside real world. All it has is a set of data about the environment, and it is constantly involved in an activity of composition of data into a prediction model, which means that it gives rise to the subjective perception that minimizes error and maximizes intra-perception coherence.

Bodily changes, emotions and the sense of our ownership and agency are not something simple and solid, on the contrary, they are the product of a guess about what part of the outside world could, under certain conditions, be considered to belong to the self. Usually, we do not have any perception about the inside of our body, our organs and processes that are always active to let us live. They may become consciously present to us only when something hurts, or when we focus our attention on a specific part. All these processes contribute to the construction of a model of which they only are one half.

We do not perceive the inside of our body, but we do perceive our environment, and we constantly construct the limit and the boundaries of our identity, including elements in our sense of ownership that maximize the self-function with the set of incoming data.

Some more recent interpretations of this predictive coding account to the self (Barret 2017) stress the dependence of perception on top-down processes. The strong interaction between interpretation and incoming data implies, in his view, that sensory responses are inevitably shaped by cognitive evaluation and sensory context.

If this first implication is correct, there is a second supposed implication that in fact does not follow from the premises of the argument. This is the claim that perceptual scenes that evoke interoceptive predictions are always affectively coloured. This claim requires to be able to trace a clean boundary between affective content and cognitive processes which drive perception. But separating a conscious perceptual content from an affective colour is not such an easy matter.

In another paper about a similar issue (Seth & al. 2012) there are some passages that are worth discussing here:

> Although the detailed neuroanatomy that underlies interoceptive inference remains to be elucidated, accumulating evidence implicates the AIC as a key comparator mechanism sensitive to interoceptive prediction error signals, as informing visceromotor control, and as underpinning conscious access to emotional states (emotional awareness). A predictive self is further supported by emerging paradigms that combine virtual/augmented reality and physiological monitoring, where the data so far suggest that the experience of body ownership, a key aspect of selfhood, is modulated by predictive multisensory integration of precision-weighted interoceptive and exteroceptive signals. (Seth, Keisuke Suzuki & Crithley 2012, p. 12)

The Illusory aspect of perception has been already analysed in Ch. 2. There I have explored some evidence that supports the idea that perception of the external intersubjective world is reconstructive. Here we see another possibility of illusion that concerns identity and self-perception. The author refers to a reedition of the famous rubber hand illusion to show that the self is far from a clear and immediate intuition, real and indubitable. In this famous experiment patients are subject to multisensory stimulation, visual and tactile, on one of their hands and on a fake hand located nearby. The experiment shows that, under some environmental and perceptual conditions, the fake hand can be perceived as the proper one.

To be perceived, in this context, means that not only is the hand believed to be the proper one, but it is really felt as it is so. But this is not the only case in which our perception is deceptive, we know many other sadder cases. This is the condition of some amputated patients that feel their limb present even if it is not there anymore.

On the opposite side we can find the condition of patients that lost their ability to perceive their bodies and their spatial positions, because a neurological disease destroyed the peripheral receptors responsible for the transmission of signals about the body state. This condition generates extreme difficulties in moving in the environment and it also changes the profound sense of identity of the patients.

Understanding the main aspects that sustain and motivate the individual ascription of self hopefully should help to approach psychiatric disturbances of selfhood and emotional awareness. Dysfunctions in inference about the self and its position in the environment could underlie a range of pathologies, especially those that involve dissociative symptoms, such as anxiety or schizophrenia.

5.3 Dialectic moves in self explanation

I propose here that the move that Seth uses to describe self-construction reproduces a dynamical structure that has ancient origins. This process can in fact be traced back to the core of idealism—the dialectical process—especially in the approach formulated by G. Fichte (1794). In the next section, I explore this possibility by reviewing the main characteristics of Fichte's ontological construction, from the self to the object.

5.3.1 Self-intuition in idealism

This is how idealism is traditionally presented: phenomenal reality is a product of the thinking subject. By contrast, realism maintains that objects exist independently of a subject who knows them. Fichte reinterprets Kant's transcendental notion of self as the formal possibility not only of knowledge, but also of being. The I sets an ontological limit to affirm its freedom and its infinite dimension. At the origin of consciousness Fichte assumes the intuition of the I, assimilating it to the Kantian 'I think' and the intuition of Kantian moral law. This, as self-intuition, must be an unconditional act, otherwise it would not be the first principle. It is therefore at the same time a self-knowledge and an action. In Fichte's words:

> I am not if not activity. I must in my thought start from the pure self and think of it as in itself absolutely active: not as determined by things, but a decisive thing, that determines things. (Fichte 1794)

The I is at the point where thought and thing thought are present as the same reality. Subject and object coincide and no longer have a connotation that differentiates them: this is the heart of Fichte's Idealism. From such a coincidence, Fichte will gradually come to the conclusion that all reality ends up being resolved in the absolute self. Fichte eliminates the need for Kant's thing in itself (noumenon), for it makes no sense to admit the existence of a reality that lies beyond our cognitive limits.

In order to talk about something, it is necessary to have a mental representation, that is, a transcendental scheme; how can we say, therefore, that there is an object if I cannot reduce it to the a priori forms of an acquaintance? It follows that the phenomenon is no longer a limit caused by the unknowability of the noumenon, but it becomes a creation of the subject itself.

The categories of the intellect will also assume a different role: while for Kant they were intended to unify multiplicity, for Fichte, instead, they aim to multiply the I in its uniqueness. He then outlines the three fundamental principles that govern this reciprocal relationship between subject and object: Thesis, antithesis and synthesis.

5.3.2 Self and its world of possibilities

Object is nothing more than the result of an activity. In classical metaphysics this was expressed by the motto operari sequitur esse (action follows being). The first who proposed it was Tommaso d'Aquino in his Liber de Veritate Catholicae Fidei contra errores infidelium (1264), but then many other philosophers used it to talk about ethical matters concerning freedom. One of them was Spinoza in his Ethics (1667).

Fichte reverses this statement by pointing out the essential dynamism of the construction of the world by the individual self: esse sequitur operari (being follows action).

This attention to action is the aspect of this theory that can be considered a historical root of enactivism, according to which the object assumes relevance for a specific subject in a dynamic connection of mutual re-constitution. And moreover, it also seems to be a precursor of the theory of affordances (Gibson 1979), in which the object assumes relevance only for a specific subject.

Under this perspective, the world, with its object, represents the whole variety of possibilities of action for a subject. We do not see an intersubjectively identical world. Everyone sees the objects as meaningful, and meaning is precisely the possibility of the relationships that a specific individual can entertain with that object. The world with its objects represents the possibility of action for a subject.

The term *affordance* means this subject oriented feature that all the things in the environment exhibit. The higher the affordance something has, the more automatic and intuitive will be its use as a device or a tool. For example, the appearance of a handle should make you better and automatically understand how the door should be opened.

But in the world there are so many different objects that exhibit so many different features and everyone in the world sees different possibilities of action. Obviously, then, it is debated what we should consider as an affordance, mainly because it seems difficult to draw the boundaries of an object and the features that it exhibits, that are so different depending on the subject skills.

In the last forty years several interpretations of this approach have been developed. We can think about objects like bottles and glasses that we approach more or less intersubjectively in the same way. We see them as meaningful tools to satisfy primary needs. But there are so many things in the world that we transform by adapting them to the environment and to specific situations. Moreover, a difficult step for this theory is the individual talent, for instance of a sculptor that is able to see a face inside a marble block.

An interesting and new perspective on this theme is offered by Giovanna Colombetti, who highlighted how the notion of incorporation could be helpful to reinterpret the connection between world and affectivity. In her situated approach to affectivity, she explains that some objects can be incorporated into one’s experienced and expressive body, and this affects the way in which the subject sees the world and the way in which he perceives it, always affectively. There are many different artefacts in the world and some of them could scaffold our affectivity and shape our relationship with our environment, is this the case of musical instruments in the musician performance, for example. (Colombetti -2013).

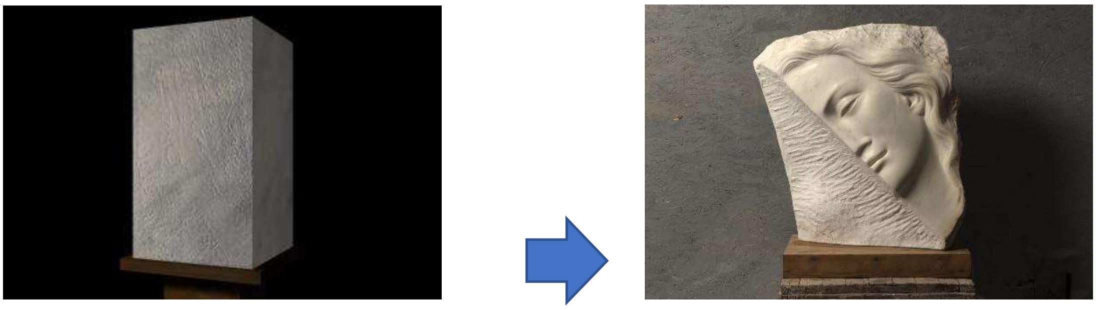


Fig 2 There is a strong difference between what an object can represent and the affordances that it exhibits depending on the subject.

Coming back to idealism, the essence of the self consists precisely in an activity of a self-conscious nature: this activity is to pose, in the sense of creating, the material world, according to the titanic romantic conception. The Fichtian I is, like self-consciousness, an intellectual intuition. Since now we have talked about individual empirical self.

As we will see in the following section, this is only the first step of the Fichtian explanation. At the end of the dialectic process, we will see that the I, from which everything comes, does not coincide with the single empirical self, but it is the absolute I. This absolute I is a cosmic principle, it is the apex of the dialectic process and the connection between the individual and everything else. This thesis will be articulated in two other principles that show the multiplicity of individuals and the non-existence of an external world.

3.3 From the Self to the I

It seems impossible to think of nothing, a thought has always a content. I do not want to explore here the development of this principle, which is the theory of intentionality that will be developed by Brentano's phenomenology. The claim that a conscious thought can only be in relation to thought objects is used by Fichte to move to a second formulation, the antithesis of the first one. The Not-I represents everything that is opposed to the I. The reason for its existence is that something external to the subject is needed to activate any possible knowledge. So, the intentionality of thought finds the necessity of the existence of an outside world. And it is also an epistemological foundation. Such an external reality, however, cannot even be something absolutely independent of the subject, because otherwise it would fall back into the Kantian dogmatism of the thing in itself (noumenon). Thus, the second principle serves to give meaning to human knowledge that, without a logical reference to the object, would become vacuous and inconsistent. But the Not-I, even if it makes subjective action possible, limits the self from an ontological point of view, generating the need for conciliation.

The opposition between self and Not-I does not happen as such, but in a dialectical way, so that while they are limiting each other, they also determine each other. While the second principle merely brought the Not-I back into the I, leaving them in a state of pure opposition, the third principle represents the moment of synthesis, giving rise to their mediation. The I becomes aware that it is not only opposed to the Not-I, but also limited by the latter, thus dividing itself into multiplicity. The mutual limitation of the I and the Not-I allows to explain both, the mechanisms of cognitive and moral activity. In particular: the self, as determined by the Not-I, is the basis of theoretical activity. The Not-I, as determined by the I, on the other hand, grounds practical activity. While in knowledge the object precedes the subject, in the action the subject precedes and determines the object, which arises to make itself a tool for his freedom. As we saw in more detail Ch 2, these central themes of Fichte's thought have been revived and developed in more recent years by enactivist and 4E and predictive coding perspectives, in the form of the mutual constitutive relation between the environment and the very subject.

5.4 Self-construction and reflexivity

As we saw in Ch. 2 and Sect. 5.2, contemporary reflection puts an increasing emphasis on the embedded and environmental dimension of an individual. The interaction between these levels is investigated through the integration of neurophysiological and psychological models that try to capture the mutual representation and mutual influence between them. This analysis is intrinsically linked to the analysis of intersubjectivity where, similarly, subject/subjects levels overlap in a mutual representation and mutual influence. Such relationships can be investigated on the mechanical or psychological side and it is impossible to ignore either side.

5.4.1 From outside to inside

4E cognition and predictive coding approaches stress that the boundaries between these two sides cannot be cleanly set. These perspectives are characterized by the refusal of the classic view of mind-body relationships, which can be stated as follows: *The mind is the inner form of the body, the body is the outer form of the mind*. This assertion presupposes two strong assumptions. First, that there is a line of demarcation between the inside and the outside that can always be uniquely defined. Second, as it encapsulates the mind into the individual's body, the mind is directly affected only by individual bodily changes, and not by environmental changes.

4E cognition approaches try to solve these difficulties, by extending the boundaries of the mind and its causal connections beyond the limits of the individual body. We could sum up this perspective by reversing the previous statement: *The body is the inner form of the mind, the mind is the outer form of the body*. This reversal, in my opinion, better highlights the two-way relationships and the constant dialogue between mind, body, subject, and objects, oppositions that are not defined once and for all.

From a phenomenal point of view, we have evidence from the outside: what perceptually lies there, outside us, seems to be fixed and stable as a possible object of our attentional focus and description. By contrast, our inner world seems to be constantly elusive, unfathomable: our internal body although responsible for our states, keeps the most part of its activity unconscious, and it is difficult to explore it, not only from an external or objective viewpoint, but also for the very subject that is experiencing her bodily changes. Even though the interior body does not explain the whole spectrum of our emotions, contrary to William James' (1890) claim that emotions are just feelings of bodily changings, our body is nevertheless strongly involved in our mental activity.

5.4.2 Causality in actions

The contradictory aspects of self-perception, as if we were spectators of the projection (what we see) of a motion-picture of which we are invisible actors, is a classic topic of philosophical reflection, which in the last decades has undergone a renewed interest. In a way, we are always with us and without us. The short moments in front of a mirror shows us an illusory reality, because of its partiality of extension. The image reflected in a mirror is at the same time an exasperation of presence, an intuitive evidence accompanied by virtuality, which opens many possibilities of different plans of existence.

Before the 80's, many models of explanation in neurophysiology or in the philosophy of mind, started from the assumption of the existence of an individual consciousness, but today the need to consider the complex situation in which subjective perception occurs becomes more and more evident. What we perceive is the encounter with our sensible structures and the environment. Thus, by its nature, it is dynamical, limited and ever changing.

The I seems to be this continuous activity of resistance to the world, of self-affirmation, of fulfilment of a form that external reality continues to call into question. The I is the limit that does not belong to reality and at the same time has its roots in it. However, the discrimination between these two sides, in continuous dialogue and mutual reconstitution, is the main node on which to work, and the principal reference of all classifications and level distinctions in the philosophy of mind. The prominence of the I and the claim that the outside world does not exist without an I that perceives, that started with the solipsistic doubt and had its apex in subjective Idealism, no longer is a foundation for our enquiry of the mental. Perhaps, it should be overturned or, better, completed with a more comprehensive view. Even the I does not exist without the outside world and other subjects.

It is for this reason that the analysis of consciousness is necessarily intertwined with the enquiry about time. In order to consciously perceive, it seems to be necessary to focus attention and establish some fixity but, as I will show in the next chapter (that is precisely about the relationship between self, time and memory), things are much more complicated, for the root of personal identity lies in continuous change.

I try to explain this perspective through the metaphor of reflection, as an illustration of this constructive aspect of the self. In the mirror there is a change in the virtual scene, which occurs concurrently with an action made by the subject. This is consistent with what we say in Ch. 3 Sect. 3.3.3 about Alva Noë's approach to perception. The acquisition of the competence of sensorimotor dependence, which connects a sensory experience to an experienced desire, is progressive.

5.4.3 The others in the self. Intersubjectivity as a function of perceptive normalization

By these considerations, it is plain that the essence of personal identity should not be searched merely inside the subject. Phenomenal knowledge (which is meant to be indubitable because it is intrinsically perspectival and subjective) arises from reflection, in the sense of being both sensory extrinsic and dependent on an inseparable interconnection with the other and the environment.

In a more narrative sense, personal identity contains a multiplicity of aspects that also concern the outside of the very subject. In nineteenth century, literature there are many works that are centred around the reflection that we constantly change mental states, shape, mood and thoughts. Identity is something that refers to how others see us and what we see of ourselves. If the essence of personal identity lies in change, then we should better explore this aspect and its connection with time. Timing and memory will be the subject of next chapter. Our experiences are a constitutive part of personal identity, but at the same time everything we represent for the rest of the world and for the other people is part of it.

All our actions are mediated by the exchange with the outside world. What we are constantly doing is a continuous exchange with the outside world and our action changes a field of possibilities for action around us. We carve out our personal identity within possibilities of action that we learn by reflection and mental representation of our limited corporal extent, and the actions of others on which we model our actions.

Not only the environment, therefore, but also other subjects, are crucial in the construction of the self and in the construction of the possibilities of action. Concepts of personal identity, consciousness and self are so complex and so interrelated, that is exceedingly difficult to describe them properly. There is here no presumption of completeness about their definition. I point out here some differences between these concepts to try to explain how I use these notions in the context of this work..

Personal identity is a multi-layered concept that consists in how others see us and how we see us, through the filter of the encounter with the world. Consciousness is an intuitive feeling about some events. These events are perceived as internal or external, depending on the causal sequence of their appearance. The self is here meant as the consciousness of personal identity. The self arises from the intuition that something is happening in the present, it happened in the past or never happened.

More important, the others also constitute the perceptual normalization that regulates and guides individual perception. Within the predictive coding framework, the other represents a set of input variables that involves certain predictive computations. The intersection of levels, the matching made by the areas of convergence between perceptions that arrive physically out of phase, is possible only through experience, through an external observer who informs us of perceptual synchrony.

Temporal synchronicity, so understood, is a perspectival problem. In conclusion, dissecting different levels of reality is useful for studying physical, biological, and psychological phenomena. Even dissecting psychic reality in different levels such as the processes of memory, perception, attention, and imagination is certainly useful to understand behaviors and reflexive acts. We will see in the next chapter that an analysis of temporality and temporization as a means of coding stimuli could, in my opinion, offer an interesting new tool to deal with this vertical hierarchy of explanatory levels.

Up to now, we have considered several theories of perception and the construction of self, but we have not said anything about memory, the faculty that seems to be primary and fundamental for the existence of every learning process and the very self.

The next chapter aims to address the complex relationship between memories and self-generated ones, by focusing on different notions of causation that are presupposed by different research methodologies and on the limits of their applicability. I will start from the interesting case of music-generator epilepsy. Subjects affected by this pathology help us to explore a paradigmatic type of memory, the musical memory. In these patients there is a clear phenomenological perception of songs, which belongs to their past lives but that they hear as completely real. These perceptions overlap the perception of the subject's environment, sometimes in a very dramatic way, preventing the subject to hear anything else. Sometimes these subjects are able to perceive both: the reminded song and, at the same time, what is happening in their environment, the voices of other people, and they are also able to interact normally with everything.

This leads to an extraordinary condition called 'doubling of consciousness' that I will explore in the last part of next chapter. This condition is also a characteristic feature of some epileptic episodes at their starting phase. In this case the subject vividly reminds something of her past before the episode. This condition is not a psychiatric disease. On the contrary, it is the coexistence in the subject of two dimensions: perception that comes from an exogenous stimulation and, at the same time, perception which comes from endogenous ones.

From a neurophysiological point of view, these cases are considered to be the effect of an enormous and aberrant activation of temporal areas. This could be the effect of a weak ictus, or a rare case of epilepsy, which involves temporal areas. Moreover, these conditions could be generated by a direct stimulation of different temporal areas during surgery without total anaesthesia. W. Penfield (1954) provides a taxonomy of these areas and the perceptions that they are able to generate. Similarly, many other neurophysiologists understand memory as an epiphenomenon of neural processes.

I will propose that a different interpretation of this kind of aberrant memories could help to shed new light on the process of normal perception and its relationship with time. I will argue that the notion of causality assumed by the neurophysiological investigation about memory processes involves a fragmented and modular understanding of its subject of investigation. This method proves to be fruitful for an analysis of certain specific motor faculties, but it fails in its application to the understanding of memory consciousness.

I intend to show the links between the assumptions of a modular realism, time directionality and closed memory, which are implicated by the neurophysiological investigations. In particular, I will discuss the notion of interaction between functional fields, as an explanation for memory processes. I will then highlight the implications that these approaches manifest with a typological and morphological categorization of stimuli and their limitations. Finally, I will thus outline the possibility of a different explanatory paradigm, based on timing coherence of neural activity, alternative to the one proposed by some interpretations of the predictive coding framework.

By arguing that memory is not caught by the conception of cause, time, and fragmentation as understood in the neurophysiological method of investigation, I will delineate a different paradigm of investigation, which is grounded on reflections that belong to the theoretical apparatus of the philosophy and phenomenology of time, and allows us to better understand these fundamental aspects of memory consciousness. In particular, I will explore the aspect of synchronicity to propose a re-reading of the waves phenomena that occur in the brain, the role that they can play in encoding information, and their correlation with conscious memory experience. This will be the subject of the next chapter.

Chapter 6

Self-generated memory:

When the boundaries between what happened and what is happening collapse

To investigate learning processes and related levels of consciousness it seems impossible not to talk about memory processes. Memory seems to be the faculty that sustains any possible learning and keeps together our sense of personal identity. Nevertheless, this is a very complex process to investigate, and nowadays it is still not clear how we store information. It is not even clear whether we really store something, or we reconstruct it from time to time and from new elements. It is also debated if the contents of memory occur randomly or in a principled way. These are the main aspects that I will explore in this chapter. I will investigate the connections between personal identity and time through a paradigmatic case: musical memories and self-generated ones.

Abnormal activity in the brain could be associated with some change in conscious life. There are many pathologies that show alteration in morphology and informativeness in neural components, but in most of these cases a comparison with the normal conscious life, or the one that the subject used to have before the occurrence of abnormal neural activity, seems to be impossible. Memory and its connection with actual perception are mental events that, in some conditions, could be compared. In the following section I aim to analyze these two dimensions under pathological conditions in order to better understand some of their main features and their different timing components.

6.1 Two reality dimensions at the same time

I start introducing the case of "music generator epilepsy" or "incontinent nostalgia" with the words that O. Sacks used for one of his patients:

> Mrs. OC, a little deaf, but otherwise in good health, lived in a hospital. One night in 1979 she had a very vivid dream and full of nostalgia: she was a child in Ireland and, above all, she felt the songs so often sung and danced down there. When she woke up, she noticed that the music continued very loud and distinct ... The conversation with her was not easy, partly because of her deafness but more because my voice was overwhelmed by the songs. After some days she was able to speak and to listen, between intervals. (Sacks 1985, Ch. 15)

I think that this case is very interesting because in self-generated memories the boundaries between what happened and what is happening collapse. Subjects affected by this pathology help us to explore a paradigmatic type of memory, musical memory. In these patients there is a clear phenomenological perception of songs, which belong to their past lives and they hear as completely real. These perceptions overlap the perception of the subject's environment, sometimes in a very dramatic way, preventing the subject to hear anything else. Sometimes these subjects are able to perceive both: the reminded song and, at the same time, what is happening in their environment, the voices of other people, and they are also able to interact normally with everything. This leads to an extraordinary condition called *doubling of consciousness* that is the most interesting aspect that I would like to discuss here. This condition is also a characteristic feature of some epileptic episodes at their starting phase. In this case, the subject vividly reminds something of her past, before the episode. This condition is not a psychiatric disease. On the contrary, it is the coexistence in the subject of two dimensions: perception that comes from an exogenous stimulation and, at the same time, perception which comes from endogenous ones.

6.1.1 Activation of temporal lobe: taxonomy of perception

How are these special memory episodes investigated and what are they, from a neural point of view? Mrs. OC was, unfortunately belatedly, subjected to EEG, when she still felt occasional fragments. The investigation was traced by requiring the patient to raise her right index finger when she was hearing the songs. In two hours of recording, she raised her finger three times, and the concomitant signal detected an intense activity in the temporal lobe of the brain when she heard the inner music, probably caused by a small stroke. Within a few months, the episode was resolved, leaving no trace in the memory of those songs.

Sacks call these episodes ‘incontinent nostalgia’ because he wants to highlight the strong emotional component that this kind of memory involves. For this reason, he does not exclude the possibility that there could be a connection between the content of a memory and a present event or a subject’s need.

From a neurophysiological point of view, these cases are considered to be the effect of an enormous and aberrant activation of temporal areas. This could be the effect of a weak ictus, or a rare case of epilepsy, which involves temporal areas.

Moreover, these conditions can be investigated by a direct stimulation in patients with severe epilepsy. Before excisions of abnormal area is carried out, it is possible to explore the exposed cortex, applying a gentle electrical current to it. Penfield (1954)[17] provides a taxonomy of these areas and of the perceptions that these direct simulations are able to generate. Like Penfield, many other neurophysiologists understand memory as an epiphenomenon of neural processes.

Graham, Brown, and Sherrington (1912) described local facilitation and intensification of motor responses by repeated stimulation at a single point in the anthropoid cortex, and Penfield found a similar effect with respect to motor and sensory areas of the cortex: a particular response can sometimes be repeated by interrupting the local stimulation and then reapplying it at the same or a nearby point, and the threshold of evocation of that particular response is lowered for a time by the first stimulus.

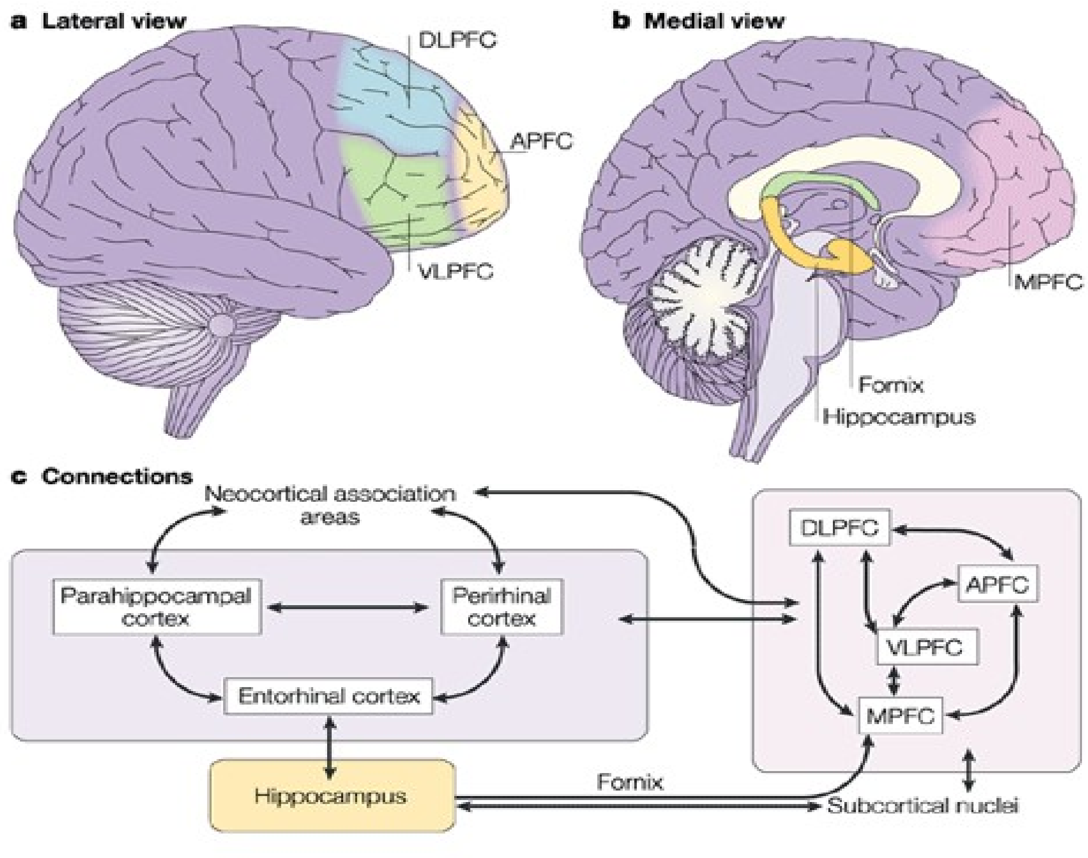


Figure 1 A neurophysiological model of relevant areas involved in memory processes.

[17] Penfield W. and Jasper H. 1954. This work refers to numerous patients that were treated for epilepsy. The precise indication of the neural points that, when stimulated, are able to generate perceptual contents, does not correspond to the same precise description of the very content. We need to take into account the condition of the patient and the short time available.

6.1.2 Psychological effects of neural stimulation

When we observe abnormal activations in different parts of the brain, we simultaneously observe phenomenological perceptions that occur in the subject. They are in general psychological experiences from the patient's past and, in Penfield's opinion, not particularly important ones (1954). All and only the details of those things to which someone had paid attention are still there. Throughout all these evoked experiences, patients are still aware of the fact that they actually are in the operating room.

What we have referred to as psychic responses include many different elements, made up of auditory, visual, somatic information, as well as interpretations, perceptions, comparisons, emotions. In fact, under the 'psychic' heading, there are two main kinds of responses.

1. Experiential: Phenomenological perception of past experiences, which includes past events and past interpretation.
2. Interpretative: Phenomenological perception of present experiences and the interpretation of what is going on.

6.1.3 Stream of consciousness and random memory

In Penfield's account (1954), either type of response, experiential or interpretative, attests the existence of a permanent ganglionic recording of the stream of consciousness. Timing is another argument that seems to reinforce this thesis. First, because patients who vocally reproduce their mental song seem to perfectly reproduce the real rhythm of the song, second, because no patient has reported that people who walked or spoke during the hallucination did so at an unusual or unexpected rate of speed. The record of that stream must be preserved in a specialized mechanism. In Penfield's words:

> The stream of consciousness flows inexorably onward, as described in the words of William James. But, unlike a river, it leaves behind it a permanent record that seems to be complete for the waking moments of a man's life, a record that runs, no doubt, like a thread along a pathway of ganglionic and

> synaptic facilitations in the brain. This pathway is located partly or wholly in the temporal lobes. (Penfield 1954)

This thought about the dynamical flow of consciousness and its reduction to bodily events, is having nowadays a great fortune. It goes back to a strong tradition in psychology and philosophy of mind, which sees William James as one of its most prominent advocates. James was influenced by Locke and by the Kantian concept of the 'I thought', or 'transcendental apperception' as the unifying principle of consciousness. He ended up analyzing deeply the role of emotions in the flow of consciousness, and he tried to assign a particular cognitive role to emotions, identifying them with feelings of bodily changes.

> My theory is that the bodily changes following directly the perception of an exciting fact, and that our feeling of the same changes as they occur IS the emotion. If we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its bodily symptoms, we find we have nothing left behind, no ''mind-stuff'' out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains. (James 1884)

This means that if a subject is exposed to an experience in a specific environment, his body causally reacts to it in a specific way, either instinctually, if it is a simple event that causes and instinctual response, or by producing a structured learned output, if the event is more complex and layered. So, for instance, if we see something dangerous for us, we can flee or respond in some instinctive way. This action involves some sequential changes in the body. We perceive the changes that occur in our body under these circumstances. These feelings are precisely our emotions, according to James's account.

The novelty and, at the same time, the aspect that makes this theory difficult to accept, is that it reverts the chronological relationship between an event, the reaction to it, and emotional perception. In this account, bodily changes, which are generated by circumstances and are aimed at protection according to an evolutionary process, occur in the subject and these bodily changes are perceived by the very subject, so emotions are feelings of bodily changes that occur because of some specific circumstances.

Many philosophers contested this approach, because of its strong implications on personal identity and free will. The role that environment plays in action construction seems to be too strong in this approach.

Antonio Damasio (2000)[18] is one of the neurologists who are sceptical about this approach, he is not a reductionist and he points out that the possible role of emotion in cognition and behavior seems to be variable, depending on the subjective state. So, if we accept James's account, it could be difficult to explain the variability of actions that we perform in the same context.

A second criticism of James's view maintains that his account ignores the role of subjective experience in emotion. Damasio notes that James seems to postulate a basic mechanism in which particular stimuli in the environment excite, by means of an innate and inflexible mechanism, a specific pattern of bodily reaction, but we experience flexibility non only in actions and response to an input, but also in the way we feel and perceive the same bodily changes. James does not seem to consider the process of mentally evaluating a situation and the role that it plays in our cognitive universe, and in relation to other kinds of mental states that cause emotions.

One of the most problematic aspects of James's account is that it may explain the emotion that occurs first, but it does not work for the subsequent emotions, which depend on a more complex structure of bodily changes and cognitive mechanisms. James's theory of emotions is often criticized for placing too much emphasis on bodily feelings and neglecting the cognitive and psychological role of emotion in subjective life.

But there are also many researchers who favour this theory and try to reinterpret James's claim in the light of his later works, as well as others' more recent works. One of the most famous is Matthew Ratcliff. He suggests that such criticisms are misplaced. He argues that James does not emphasize bodily feelings at the expense of cognition. Rather, bodily feelings are part of the structure of intentionality. In Ratcliff's interpretation James's

[18] In many other works, Damasio shows scepticism towards materialism, and he highlights the role of conscious perception and cognitive evaluation in driving individual action.

approach is an attempt to reconceptualize the relationship between cognition and emotions, by a methodological introspection that deconstructs our common sense claims about mental activity (Ratcliff 2005). Thus, James avoids taking at face value the existence of some mental states because we have confusing words to refer to them, like the words that we use for some emotions. He makes clear that the ontology of our mental states seems to be generated by the vocabulary that we use to describe a whole host of simultaneous perceptions.

If we accept the existence of conscious contents and emotions that occur in the individual and are generated by specific reasons, or even if we just accept the existence of feelings of bodily changes, it is clear that we have to approach the question of how this stream of consciousness can be encoded. How is information stored in order to make the occurrence of a memory, or its reproduction, possible? This is the question that I will discuss in the next part of this chapter.

6.2 Temporization as a code for memory

It may seem, from a point of view such as the one proposed by Penfield, that memory is a reproducible and immutable form that can be enucleated from the flow of consciousness. Memory, under this perspective, is implemented as an epiphenomenon of a localized and unusual activation in the temporal cortex. This recorded stream, the memory content, could be encoded in a permanent activation structure that recurs always identical and becomes available to conscious recruitment, under specific environmental and bodily conditions.

Nevertheless, it is possible that a specific memory is not represented by the location of a specific pattern of activation. A pattern is in fact specified by two main features: the first one is the spatial location of the active units, but the second one is the temporal distance between their activations. The focus of my perspective here is this second feature. It is not unlogic to think that a way of coding is the relative proportionality between the temporal distance of neural activation. A timing-based model is economical from the energetical point of view. This means that a memory content is a pattern of activation that could be spread on different units of time, as the proportionality between patterns of activation is preserved.

This alternative interpretation of this kind of aberrant memory could help to shed a new light on the process of normal perception and its relationship to time. The notion of causality assumed by the neurophysiological investigation on memory processes involves a fragmented and modular understanding of its subject matter. This method proves to be fruitful for an analysis of certain specific motor faculties, but it fails in its application to the understanding of memory consciousness. I intend to highlight here that there is a link between the assumptions of realism about discrete modules, time directionality and closed memory, which are implicated by some neurophysiological investigations that will be dealt with in the next section.

6.2.1 Memory: the other side of consciousness of the present

These approaches manifest a commitment to a typological and morphological categorization of stimuli. I would like to suggest that another explanatory paradigm, based on timing coherence of neural activity, is a possible alternative to the one proposed by current interpretations of the predictive coding framework, such as Heilbron and Chait's (2018).

I think that a research paradigm grounded in some reflections that belong to the theoretical apparatus of philosophy and phenomenology of time allows us to better understand these fundamental aspects of memory consciousness.

The word *memory* is a little misleading, for it refers to a content with defined boundaries, which is able to overlap the momentary consciousness of present reality and recurs identical, despite varying circumstances. However, this does not capture what is phenomenologically evident: a memory is not a form of identical reproduction, but a constitutive part of the consciousness of the present moment. Memory is in the present experience, or rather, it is a special part of the process of being presently conscious.

6.2.2 Multi-dimensionality, selfishness, and constructiveness of memory

What is recalled in memory and why? As we saw earlier, in Penfield-like approaches, memory is a portion of previous perception available for the stream of consciousness, which is activated by a stimulation and thus becomes momentarily conscious, but does not carry out a specific function. This interpretation, however, leaves a large part of what we call memory, untouched. It excludes the possibility of a reason on which the appearance of a specific content of memory depends. But, if we reconsider more deeply OC's memories of English songs, we realize that they brought with them a host of sensations, the warmth of the family, and the joy of childhood. She had never remembered these things before, a strong amnesia had tarnished all her childhood, but now a dreamlike vision of a living past was revealing a sense of great clarity, completeness and serenity never known. This experience seems to fulfil a specific need of personal unity.

That sonic experience conveyed the coincidence of a multiplicity of senses that came together in the form of a past experience. One could doubt the authenticity of these memories. As she had not recalled anything like that earlier, why not think that this experience was deceptively unreal, a pure construction of the brain, that arise from random and abnormal stress? This objection, which can be addressed to many mental processes, does not capture, I think, the problematic issue here. This criticism is sterile because it conflates the mnemonic mechanism with the imaginative one, by saying that abnormal neural activity could generate contents that have no correspondence with past reality. But, what is interesting here, is exactly the sense of property and the ego-centered perspective that characterizes and differentiates the experience of memory: the feeling of property that makes memory an undoubtedly private and real experience. It is an essentially private phenomenon and, for this reason, it is ineffable, but nonetheless completely realistic.

However, what could be saved of this objection is its focus on the constructive process of memory. Although memory does not build brand new content, it is an inherently dynamic and reconstructive process. This multi-layered structural dynamic has been recently analyzed from the standpoint of predictive coding.

6.2.3 Deductive construction of the past

What does predictive coding tell us about self and memory? The predictive coding perspective focuses on a specific aspect of our daily lives: the effectiveness of our actions requires to anticipate the conditions that will be generated by different possibilities of action.

The brain is constantly busy predicting the future and comparing it with the past, using old experiences to anticipate what is supposed to come next. After such comparison, the processing of correctly predicted information is suppressed, whereas prediction errors are transmitted to the next level of neural processing, in order to drive inference to achieve the overall goal of prediction-error minimization. In this dynamical process entorhinal-cortex, hippocampus and temporal lobe are strongly involved.

As we saw in the previous chapter (Sect. 4.1), Anil Seth provided a deductive account for ego construction and memory recall. This suggests that, between a set of incoming data, the ongoing activity selects the elements that satisfy the maximum coherence between them. So, ego and memory are not, under this perspective innate and immediate perceptions, they are sequential constructions that verify the questions: "What is more likely to be me?" in case of ego, and in case of memory: "what is more likely to be past?".

What is interesting to note here is the embodied nature of memory. Memory always manifests itself in different contexts, where it is invoked (perhaps through associations) by different details. So, depending on the context or time available, memory takes on variable colours, details and extension. This means that the aboutness of a memory, which appears in different contexts, could be similar, but the way in which it presents itself to the subject is always changing. We must consider the temporal extension of a memory, and also an equally fundamental aspect of it: the present conscious context in which it is inscribed, which is its condition of existence.

A memory is something necessarily embodied that superimposes on the conscious perception of the present moment, a conscious perception with a different degree of reality (because it seems that it has a different cause).

This may be possible through a temporal coherence between the temporal structure of the stimuli and the temporal pattern of activations in neural individual systems.

I will also propose (in Sect. 6.3) that a memory reproduces the same pattern of neural activations of a perception in the present moment, with a temporal distance between them that exhibits the same proportionality of the original one.

6.3 Memory: a slower perception

As anticipated in Sect. 6.2, I would like to suggest an interpretation of a memory content as a pattern of activation that could be spread on different units of time, when the proportionality between patterns of activation is preserved.

This is consistent with the possibility that the same content recurs in different conditions, perhaps because it is generated by a reproduction of the same temporal distribution. This could be generated either by aberrant activations or by coherent resonance with a temporal distribution of stimuli in the environment.

This is also consistent with an interpretation of time and perception that is not fragmented, as the one that I discussed at the end of Ch. 3 (Sect. 3.3). Let me briefly sum up those conclusions in order to explain the sense in which this timing structure may be a foundation of many different mental activities.

Contrary to the traditional interpretation of time as an entity that extends above and beyond objects, with a fixed and unchangeable ontological distinction between what happened in the past and what is happening now, some theories of the mind consider change or becoming as the basis of reality, in the case of memory content I suggest that this dynamic might be due to a proportional timing architecture of activations.

I intend here to propose that a conscious content in the present could be correlated to a pattern of activations that has a spatial component (location of active units) and a temporal one (the temporal distance between neural activations that compose the pattern). This content could be saved (maintained for future recall) as the reproduction of that pattern, but with a dilated temporal distance which preserves a proportionality with the original one.

From this point of view, recall of a past perception would imply a contraction of that dilation, which reconstructs the original speed and timing architecture of that specific pattern.

In addition, this way to store memories would be energetically convenient, because it would require distributing the energy necessary to activate neural units during a longer temporal span, as the proportionality is preserved.

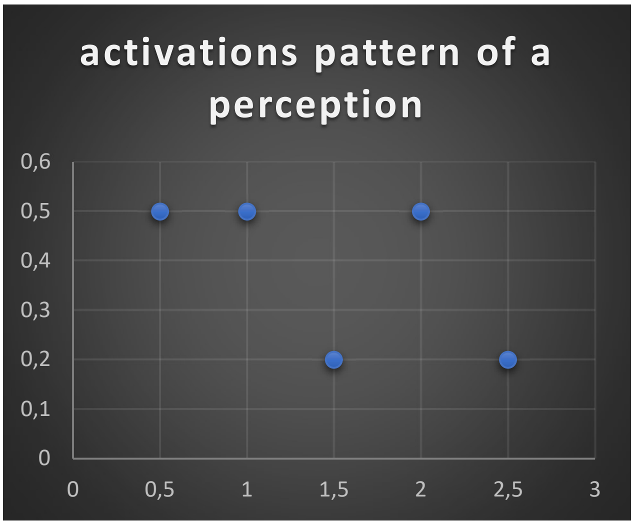
We could think about a metaphorical example to try to visualize what I mean with this kind of memory explanation. We could think of perception as a song, where every note represents a feature of the scene perceived (which in the brain is represented with a pattern of neural activations). These notes are distributed in the time flow as a specific configuration. There is a specific time distance between every note. This perception is stored in memory as a song played with a slower rhythm. When we recall this memory the rhythm of the song is brought back to the original one.

So, we could think about the activity of recalling a past perception as speeding up something that is stored as a slow pattern of activations. I try to represent this intuition of memory activity with two graphics (Figure 2 and Figure 3). The first one intuitively represents a simple pattern of activation composed by two neurons, which activate during a conscious perception. The second one represents the same pattern of activation as it is stored in memory.

We could also think about the existence of a speed threshold under which a pattern is not consciously perceivable. This explanation of how memory would store contents could also be consistent with what is going on during music generator epilepsy described by O. Sacks (see Sect 6.1). In these patients we observe very fast activations in temporal lobe, probably correspondent to patterns of activation that represent the content of original perception.

When a large amount of energy contributes to speed up those specific patterns, that specific memory comes up to perception and superimposes itself on momentary consciousness.

Figure 2 This graphic represents intuitively a simple pattern of activation composed by two neurons, which activate during a conscious perception. On y axis is represented one component of the neurons location and on x axis is represented the time flow in seconds. So, the neuron which occupies 0,2 position activates at 1,5s and 2,5s. he neuron which occupies 0.5 position activates at 0,5s, at 1s and at 2s.



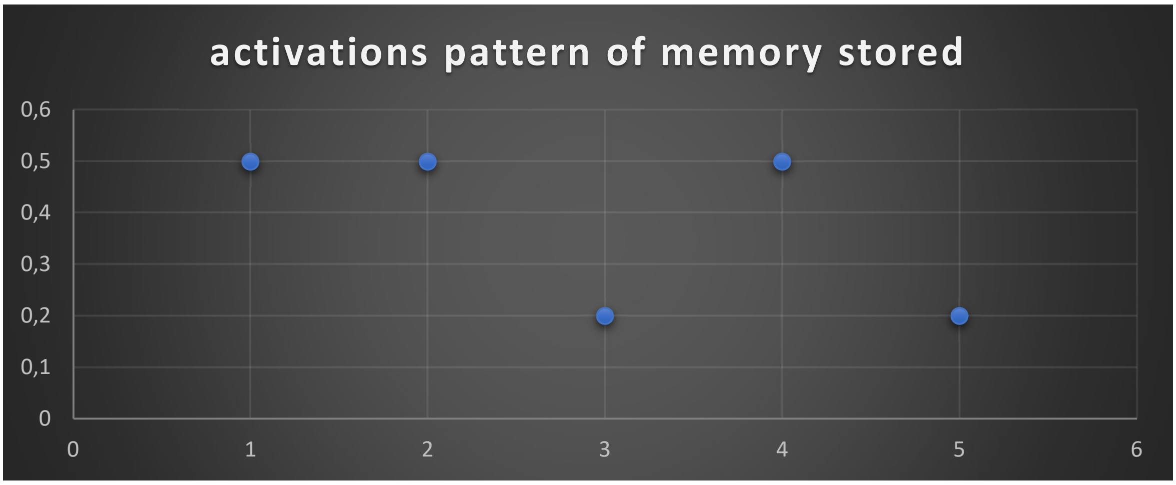


Figure 3 This graphic represents intuitively the same pattern of the perception represented the upper graphic, as it is stored. The neuron which occupies 0.2 location in this case activates at a doubled frequency of the perception condition, so it activates at 3s and 5s. And the neuron which occupies location 0,5 activates at 1c, 2s, and 4s. This is the same spatial located pattern, the proportionality between activations is preserved but it is slower.

Another way to visualize my view of what a memory might be, and how it might be stored in neural coding, is an analogy with frequency modulation in radio transmission. As we can modulate the frequency to transmit information, we could also think that this happens in neural system, through the modulation of the frequency of a pattern of activations. Figure 4 and Figure 5 represent this way of conceiving memory contents.

A wave of activations related to a conscious perception at the present moment has a specific frequency of activation. This pattern activates constantly in neural individual system but with a lower frequency. The information stored in this way is not consciously perceivable because it is below threshold, but it may nonetheless interfere with many other activations.

This interference could be the reason why memories are important components of our way to understand and see the world and are so relevant in learning processes. They coexist as a constant slow reproduction of patterns which constantly interfere with the whole neural network, contributing to change weights of specific activations, but they are consciously unperceivable when they are below a frequency threshold.

This last element is crucial in my approach as a possible defence against some criticisms that could emerge about slow activation of a neural pattern (same space portion at a different time) in which memory consists of. As an example, someone could be worried by the fact that stored memories are always in a sort of activation even though they are not conscious. This concern is motivated by the observation that considering this kind of memory process as an incessant process, seems too expensive, from the point of view of needed energy.

First, I refuse the strong opposition between what is conscious and what is unconscious in the mind. Under the perspective that I propose, the boundaries between these two notions are nuanced.

Moreover, even if we want to maintain this distinction, unconscious contents are very important in the constitution of what is conscious, so the energy used to keep them active is not wasted. In other words, we could say that the timing activation of slow patterns codetermines the timing of activations of faster ones, modifying the perception of what is real from the subjective point of view. This is a reciprocal constitutive process. Thus, slower activations are not less important than faster ones and deserve, for the organism survival, the necessary amount of energy.

Perceiving something as a memory requires the reproduction of something that connects that perception to a past perception. This is possible mainly in three different ways. First, conceiving memory as a pattern stored in a spatial location, but this also requires energy to keep those patterns alive.

Secord, conceiving memory as a part in the stream of consciousness as the one theorized by Penfield. But, as we saw before, this is not consistent with the plasticity and change of a memory and, also, it requires energy to keep this parallel stream alive so that the subject is able to bring to memory something.

For all these reasons, I believe the third way to conceive memory, the one that I suggested in the present chapter – as a slowdown of informational patterns - is not too expensive from an energetic point of view.

As suggested by R. Manzotti (2017) and other researchers, the present widens to incorporate everything we have experience of, and this is consistent with the fact that our experience extends into time to include past events.

As mentioned in Ch. 3 (Sect 3.2), this is true because every perception has a strong bodily component that requires an extended interval of time. This approach also explains also what is happening in Sack's patients affected by incontinent nostalgia.

No metaphysical gap keeps separate past, present and future, like no metaphysical gap separates a nearby object from a distant object. Both are parts of our present, not only because of the intrinsic dynamics of our neural activity, which I exposed in Ch. 1, but also because in this perspective the boundaries between some of our cognitive faculties, like perception of the present and memory of the past, seem to fade.

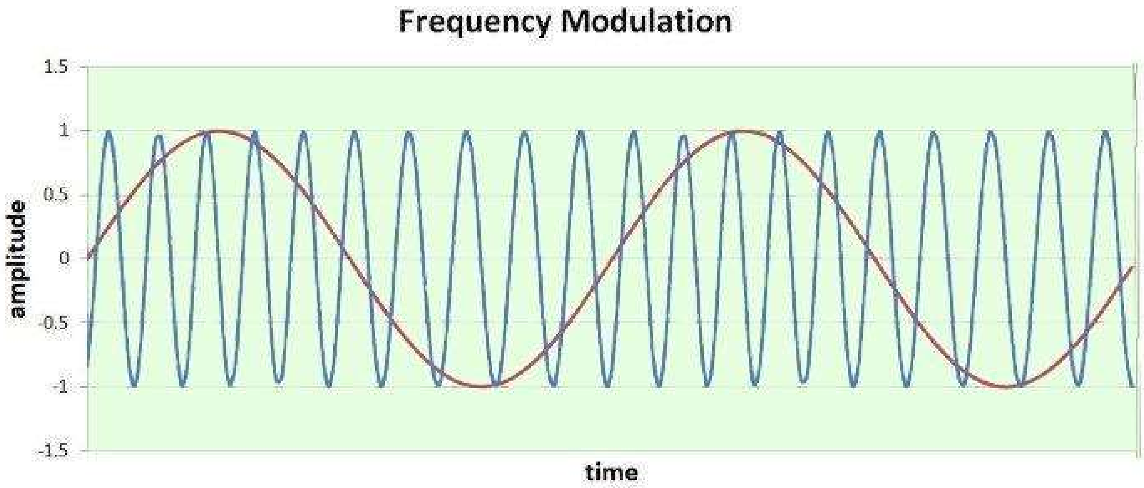


Figure 4 The blue line is a wave of activations that represent the conscious perception of a present moment, whether the red one represent the information stored but not perceivable because it is under speed threshold necessary.

Frequency of activations

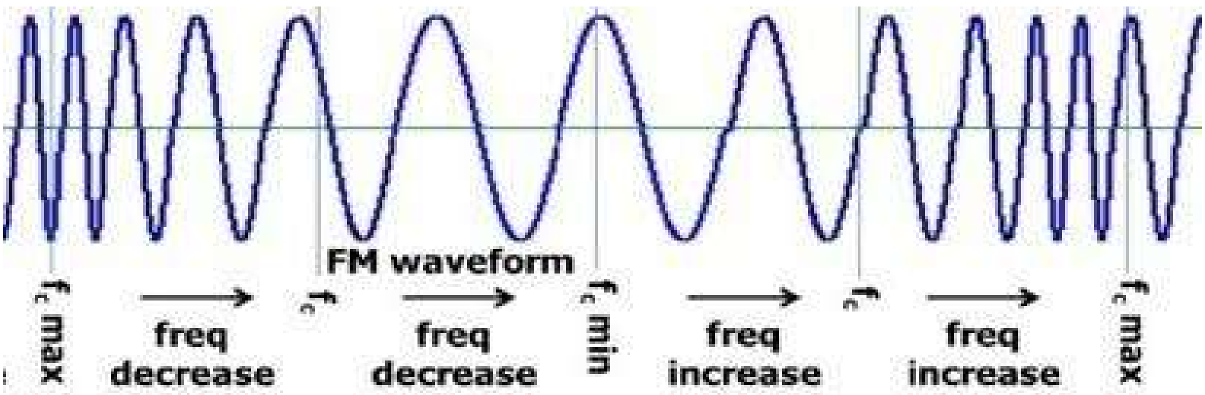
during conscious perception of

content C

Frequency of activation

during conscious memory

content C



Frequency of activations

as an unconscious storage of content C

Figure 5 The schema represents the frequency modulation of a pattern of activation as perception of C, memory storage of C, and conscious perception of that very remembrance of C after a period of time.

6.3.1 A dynamical model of aberrant neural activity

In this section, I would like to use this extended notion of temporality to reinterpret music generator aberrant neural activity, by employing a dynamical paradigm as the one suggested by Husserl's phenomenology (1893), to better understand the way of coding contents by means of a temporal architecture.

Husserl, as we saw in chapter 3, insists on the original dynamism of time, which requires us to conceive the present not as a static point, but as an act characterized by a constant consciousness flow between *impression, retention and protention*. By analysing retention, we explain the continuity of perception that Husserl contrasted to a fragmented vision of cognitive subjective life. I briefly recall here his famous example about music. Our perception (not memory) of a melody starts with a lively impression of the first note. However, it is soon displaced by the note (B) impression. But, while we perceive B, the note (A) resonates again. This mechanism, according to Husserl, ensures the presence of the past in the present, thus constituting the perception of an ongoing unity in experience. Here I would like to move a step further.

Retention is a real perception of the past and its role is to ensure a conception of an authentic past. At the same time, it guarantees the identity of the original impression. We trust retention because we have an intuitive impression of the consonance between perception and the real existence of its intentional content. For every retention is retention of everything that happened before. (According to Husserl, every retention inherits every previous retention.)

The causal power of the past, here, does not depend on a previously close event which causes some other event in the present. *All* the past is materially in the present. But if retention is very similar to impression, what can we say of memory, which we usually sharply distinguish from present impression? What is the relationship between retention and memory? This question has not a conclusive answer in Husserl. Here, using my account to the memories explanation as a slowdown of a present perception, I would like to suggest that the relationship between a present impression and a memory might be a proportionality between the time activations of their constitutive units.

Impression and retention **Perception and memory**

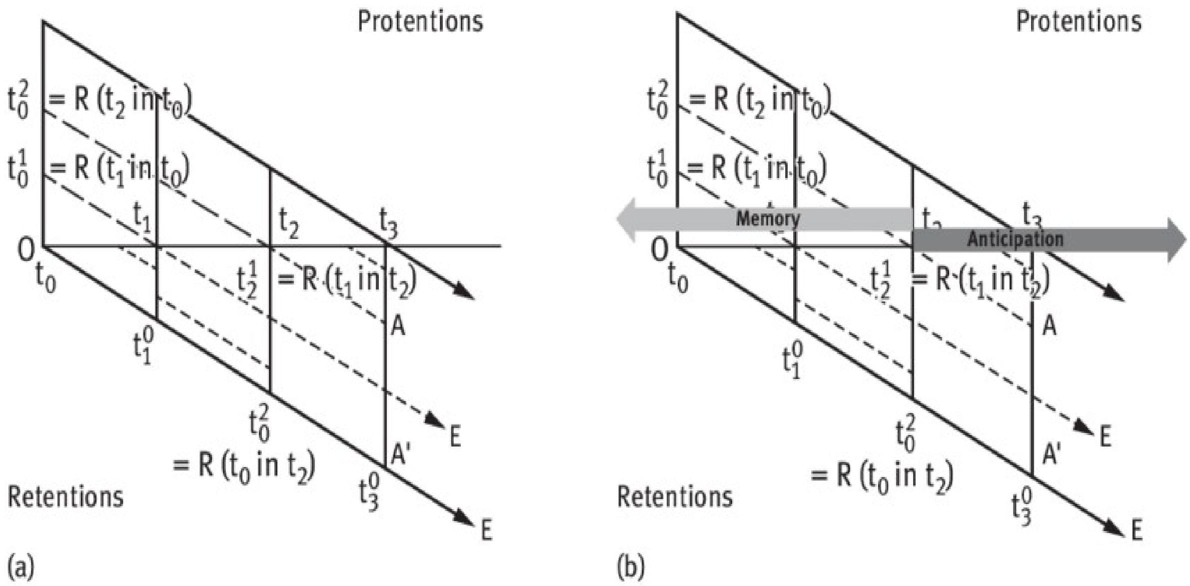


Figure 2 Husserl's (1928) diagram of time consciousness extended in both temporal directions: (a) the stream of consciousness of present as a continuity of cumulative retentions; (b) stream of consciousness between memory and anticipation.

6.3.2 Multiple consciousness and music

If we take seriously the coexistence of different dimensions of temporality and reconsider the case of music incontinent nostalgia, we can try to better understand what is going on. When stimulation produces an experiential response, patients usually recognize something of their past and, at the same time, they are acutely aware of the fact that they are in an hospital or lying on an operating table. Thus, they undergo a sort of awareness doubling and can even talk of it. But what does it mean? Is it possible that, in the same subject, an overlap of reasoning lines occurs, one about the present and one about the past?

This implies to explore whether reasoning, when divorced from awareness of sensory phenomena, has any place in the cortical record. Some patients sometimes speak of a multiplicity of unexpected thoughts coming into mind, which involve an ego multiplication in space and time. Another example is Penfield's Case A. D., who exhibits the production of two thoughts by temporal stimulation. This patient said that one thought had to do with the doctor in the room at the present moment and, at the same time, the other one had to do with the family in his kitchen, 20 years ago. The effect upon him was confusion and inability to explain. It might seem that two lines of reasoning or thinking could not coexist without interference and that, if thoughts were really reactivated, they would confuse the patient's present effort to rationalize. But the fact that the inner experience of multiple ego is difficult to explain does not mean that it is not real. These stories make us understand something that we inadvertently experience every moment: memory is a real consciousness doubling.

Some neurophysiological models speak of a unique individual consciousness emerging from biological processes. They talk of one consciousness that has, as attributes, certain mental abilities and, as content, either specific perceptions of reality, or well defined and always identical memories. But I think we should consider more seriously this duplicity or duplication of consciousness that takes place during the memory process.

The experience of memory is phenomenologically defined as a duplication of consciousness with nuanced boundaries. This intersection of dimensional planes, which could be reflected in a temporally coded architecture, is a complex structure that makes the boundaries between the present, the past and the future disturbingly labile.

I believe that musical memories in all their nuances, from humming inside to hallucinations, are so interesting because they entertain a special link with the construction of the ego. This aspect manifests itself in the amazing capacity of music to unlock the movements of parkinsonians in a fluid rhythmic sequence, or to bring back a temporary permanence of the past selves in patients suffering from amnesia and Alzheimer disease. Perhaps this is because of the essential link that music vibes weave with time. In the same unit of time music involves the individual system with a bombing of stimulations that have a specific, temporally coded, internal architecture.

The attentional focus is constantly involved in this close variability of impressions that contributes to develop and strengthen the perception of a continuous flow of consciousness. Not surprisingly, this link between organized distributions of stimuli which sustains memory processes and self-perception involves the same areas.

Chapter 7

Change in timing information. Learning and conscious agency

This chapter focuses on a timing-based level of explanation of cognitive processes that are involved in the learning process, in particular attention, memory and conscious perception of agency. Global workspace theory was examined in Ch. 3 where some of its problematic aspects were pointed out. In Ch. 4, the predictive coding framework and its main difficulties were discussed. In this chapter, I will explore a new framework that tries to mediate between these different perspectives, by overcoming their weaknesses but, at the same time, by keeping their positive aspects. This will be achieved by focusing on a timing perspective.

In the first part of this chapter, I will focus on informativeness, more specifically I will explore the notion of neural information, coding, and what we refer to when we talk of brain rhythms, neural spiking of a single cell or a group of cells, and the relevance of the timing component. Then I will explore the relationship between neural activations and their timing coherence, and some of its cognitive implications, especially with respect to memory and attention. I will finally argue that causality attribution and agency ascription in individual learning, as well as the very concept of identity, all depend on a timing relationship.

In chapter 1 we considered a neurophysiological perspective to learning that explains it by exploiting the concept of neural plasticity. We focused on morphological changes in the spatial architecture of different neurons through dendritic spines and the production of different synapses. As we saw, this approach is mainly centered on the production, trade and metabolism of neurotransmitters, receptors and the enormous number of molecules that modulate the process. So, this kind of explanation focuses on material changes in neural connections at the molecular or cellular level.

In this section, I delineate a different approach to neural explanation of learning processes. The focus will be on information, analyzed in its timing structure. Its basic axiom is that morphology is manipulated by information, which is codified in neural spiking. This means that learning is possible because of a previous change in the kind of coding and decoding in neural cells. What changes is the way in which neural cells codify information and send it to the cells that are connected to them.

What is relevant for the occurrence of a learned output is consonance between different areas, or groups of neurons. And even the consonance between few cells could be truly relevant in interpreting the information coding of a piece of information. In this case, the attention is not directed to spatial connections, but it is directed to pattern of activations and their timing coherence.

Under this perspective, learning is better explained as a change in neural pattern activity. Therefore, in order to better understand this process, we should explore the timing structure of the pulses of a neuron or a group of neurons. We can detect neural activations thanks to FMRI, MEG, or EEG, but we should analyze not only the part of the brain that is active under some conditions, but also the rhythm of neural dialog and its units of information or, so to speak, the alphabet of neural coding.

To understand learning processes, we need to pay attention to these changes in activation patterns. I will explore some discoveries in this research field, to better understand the relevance of timing, try to reinterpret learning processes under this perspective, and even give some hints on how to improve the outcome of these processes.

7.1 Coding of intensity and changing in timing information

Neural information is the root of neural plasticity, which determines learning processes. The neural communication code is generally described as the set of matching rules by which neurons represent the information that is received, processed, and transmitted. This will be briefly reviewed in Sect. 7.1.

Neural communication has a spatial component that defines which cells are activated in response to a given stimulus, so it specifies across the brain areas that show a simultaneous activity with a specific stimulus.

The neural communication code also has a temporal component (frequency codes and time codes), which describes the correspondence rules by which the electrical activity of a neuron encodes incoming signals and transmits the processed signal to subsequent cells. This second aspect of coding will be the focus of the present part. I will consider its philosophical implications for learning and agency perception in the second part of this chapter.

The study of neural coding began with Edgar D. Adrian's discovery (1926), that the train frequency of action potentials generated in response to a sensory stimulus grows with the intensity of the stimulus. This suggests that the frequency of action potential constitutes a code that contains information about the stimulus, and more specifically about its intensity. The cellular basis of neural coding lies in the fact that each neuron receives synaptic inputs, distributed on a morphologically complex dendritic arborization, which are integrated at the level of the axon emergence from the soma, where the outcome of this synaptic summation over time determines whether and how many action potentials will be generated, as we saw in detail in Ch. 2.

The study of neural coding aimed to describe the aspects of action potentials firing could be directly related to the various attributes of the stimulus: what is the information contained in that pulse train? What is the code used by neurons to transmit that information? The reverse problem, namely how it is possible to reconstruct a stimulus, is also studied by examination of neuronal firing (the problem of neuronal decoding). These questions obviously have not a unique answer and the theme of informativeness is very broad. My target here is just to try to correlate some aspects of timing to attention. I will briefly sum up some of the possible coding interpretations to conclude that, despite their differences, they share an important feature related to attention.

7.1.1 Information in frequency code

Adrian's experiments (1926) showed that the information of a stimulus is represented in the average firing of a neuron, which thus becomes the action parameter to measure the responses of individual nerve cells, depending on different attributes of a sensory stimulus. A frequency code is thus defined only with respect to parameters of the action potential train, relevant for the encoding and transmission of information: the average frequency of the pulses (number of action potentials per unit of time, measured in Hertz), regardless of their exact time distribution, in the time interval considered.

The average firing rate, that I mentioned in Ch. 3 can be defined by a time average over a sufficiently long interval (at least 0.5 sec., up to 1 sec.) to take into account the inherent variability of the firing, or by an average on a series of subsequent presentations of the stimulus (neuronal responses are described, in this case, by the so-called PSTH, Peri Stimulus Time Histogram). According to a frequency code, the exact time position of individual action potential in the process has no meaning and the variability of the intervals between action potentials is considered noise, while it could (and often has) great information content.

A frequency code, based on a temporal average of a single neuron firing, evaluated over a long-time window, or subsequent repetitions of a stimulus, is considered unrealistic today, because it seems too coarse to capture the variability and specificity of neural outcomes, Panzeri S. et al. (1999). First, reaction times for behavioral responses are too fast, compared to the significantly longer time required to measure the average firing frequency. This means that this model of information coding fails to explain or predict the human solicitous response to a fast stimulus, or a fast stimulus change like an incoming fast object that speedily changes direction. The time required to get a plausible measure of the average firing is not enough. Moreover, a frequency code is inherently ambiguous since the same average firing can be obtained with different combinations of attributes of a stimulus.

The limitations of the frequency code could be partially solved, however, considering that each neuron is part of a specific neuronal population that receives the same stimulus and that a population of neurons performs an average equivalent to a single, shorter, stimulus presentation. This thought is what motivated research about a different code: the population code, that we explore in the next section.

7.1.2 Information in population code.

Information is typically encoded not by individual neurons but by populations of neurons. The population code is the representation of a stimulus, through the joint activity of a large number of neurons, which determines an overall response. Population coding solves some of the problems of frequency codes by moving the time averages needed to give statistical significance to pulse trains with intrinsic variability into space domain (average over a population of cells). Population coding reduces statistical uncertainty due to neuronal variability and allows simultaneous, combinatorial representation of multiple different stimuli, through a population vector or other information theory decoding algorithms.

These kinds of studies are now possible with the methods of optogenetics, an emerging experimental approach that combines optical and genetic techniques, using light-activated ion channels. By interfering non-invasively and reversibly with the electrical activity of neural circuits, on the millisecond time scale, optogenetics today allows to directly address, and no longer just correlatively, the question of what code is used in a particular neuronal context.

7.1.3 Information in time code

A frequency code neglects any possible information contained in the temporal distribution of pulses. There is growing evidence that the temporal accuracy of individual action potentials in network firing (of the order of millisecond) is a significant element in neural encoding (time codes). In a time code, the time distribution of action potentials, or subgroups of action potentials, bears additional information content with respect to the stimulus, compared to the average firing rate. An example of time code is that the arrival time of the first action potential after a stimulus is a relevant variable of the neuronal response, especially when considering a post-synaptic cell that receives inputs from a population of neurons. In some models, particular intervals between action potentials, or groups of action potentials, organized within the spike train are relevant as time codes. A temporal code must be able to measure the position of an action potential, or groups of them, in relation to a precise temporal reference, which can be provided by the beginning of the stimulus, its periodicity, or by the relative temporal positions of action potentials in a population of neurons.

There are also many interesting studies about the connection between some timing architectures, which seem to encode some aspects of metabolism of neurotransmitters, and some enzymes and hormones, which modulate neural plasticity and trophism. The phenomenon of synaptic plasticity dependent on time distribution, known as *Spike-Time Dependent Plasticity* (see Sec. 7.4), shows how specific temporal relationships between action potentials can encode different synaptic responses.

7.2 From intensity to spike timing and attention

Despite all the debate about an exact definition of information coding in the neuronal system, it seems clear that the intensity of stimuli can be encoded in some way through the number of spikes in a time unit.

The same typology of stimulus hit the individual system with different levels of intensity, and this can be translated into the modulation of timing in the same neural areas.

The proximity of height peaks in neural spiking is often observed to be coherent and synchronic with high intensity of the input, to which a given circuitry is subjected, whether low oscillatory frequencies seem to indicate low intensity of the stimulus. The timing of the neural activations conveys information about the stimulus and its processing can indicate something relevant also about the conscious content connected to that stimulation. We must also observe that intense stimuli involve our attention in a stronger way. Intense contents are more likely to become aware because they manifest more urgency, as they require a response.

This claim, about a necessary connection between intensity of stimuli and attention, could be challenged by some empirical evidence that shows that, under certain conditions, we could be completely blind to some very evident features of the environment. This is the case of many illusions that we already discussed in Ch. 2.

As a paradigmatic example, we could consider the *invisible gorilla* (Chabris and Simons 2011), which shows everyday illusions governed by beliefs about the world in a top-down process. In their famous experiment a group of people are playing soccer. We are required to pay attention to some aspect of the scene, like the number of ball touches of a specific player. In doing this, the majority of watchers do not notice that a big gorilla crossed slowly the group of players. The big and strange stimulus was there, but our attention was blind to it. We constantly trust in the fact that our mind reproduces an intersubjective state of things in a perfect way, and laws and processes that rule our society are built on this claim, but simply this is not true. They write in a previous work:

> Our experience of a rich, stable visual world often leads to the intuitive belief that our representations of that world are correspondingly detailed and precise. But increasing evidence for "change blindness" the inability to detect large changes to scenes from one glance to the next, has inspired claims that little to no information about the world is preserved in visual short-term memory.
>
> (Chabris C., Simons D. Schur.T - 2002, p.78)

We have approached a similar argument when we talked about visual and auditory illusion. Here there is a strong claim about the fundamental role played by attention and the beliefs about world laws that seem to guide our perception.

The claim that our perception is illusory is quite old in the history of human thought. But my point, here, is to move further, beyond perception illusion and attention inaccuracy. I would like to question the very nature of the attention that is supposed to guide our conscious perceptions. From a phenomenological standpoint, whether our perception is false or illusory is not relevant. As I previously remarked in Chs. 3, 5, and 6, one of the goals of the present work is to import some relevant insights from phenomenology and apply them to the field of neural research. In order to achieve this goal, this part is intended to be more constructive. I do not want to investigate in depth the supposedly illusory nature of subjective perceptions of the intersubjective real world.

Our subjective perception, of either time or the world, is not illusory for ourselves. I think that, in order to understand conscious perception in learning, we should not forget or avoid considering this important and distinctive aspect, even though this contrasts with the positions of many contemporary authors, like Dennett (2005).

In particular, I wish to focus on the nature of attention, by exploring those morphological and timing aspects that can be relevant for its existence. I focus on attention and memory timing features because they are fundamental components of learning processes. In Sect. 7.2.1, I will approach some morphological aspects of the neural system that are relevant for attentional processes. I will consider the hypothesis that they are a coherent neural implementation of the workspace theory and the predictive coding theories of attention and perception.

Then, in the last part of this chapter, I will analyze more deeply the timing properties of information in attention processes, to try and define a timing-based paradigm that could help to differentiate between voluntary learning processes and involuntary ones, as well as among the different agency perceptions connected to them. The paradigm that I propose will incorporate some relevant aspects of epiphenomenal, predictive, and enactivist theories, by considering timing as the connection between subjective internal processes and external stimuli.

7.2.1 Neural morphology of attention

In Ch. 1, learning processes were considered from a neurophysiological point of view, by focusing on neural plasticity features of the single neural cell or, more broadly, of cell networks. Here, the focus will be on the morphology that networks can assume depending on learning circumstances, because I think that especially the third kind of organization (fan disposition) could be associated with some epiphenomenal approaches.

Before exploring timing aspects, we need to describe some morphological structures that seem to be strongly involved in the emergence of consciousness and learning processes. At the neuro-anatomical level, three different fundamental morpho-functional structures, that connect the different parts of the brain in circuits of relevance, overlap (LeDoux - 2002): the first is a large three-dimensional weft, that consists of a deep-level localized part, the thalamus, (Fig. 1) which receives sensory and other signals. From this structure, some foils get out that connect to the cortical part (Fig. 2). These circuits seem strongly active in thought and conscious action.

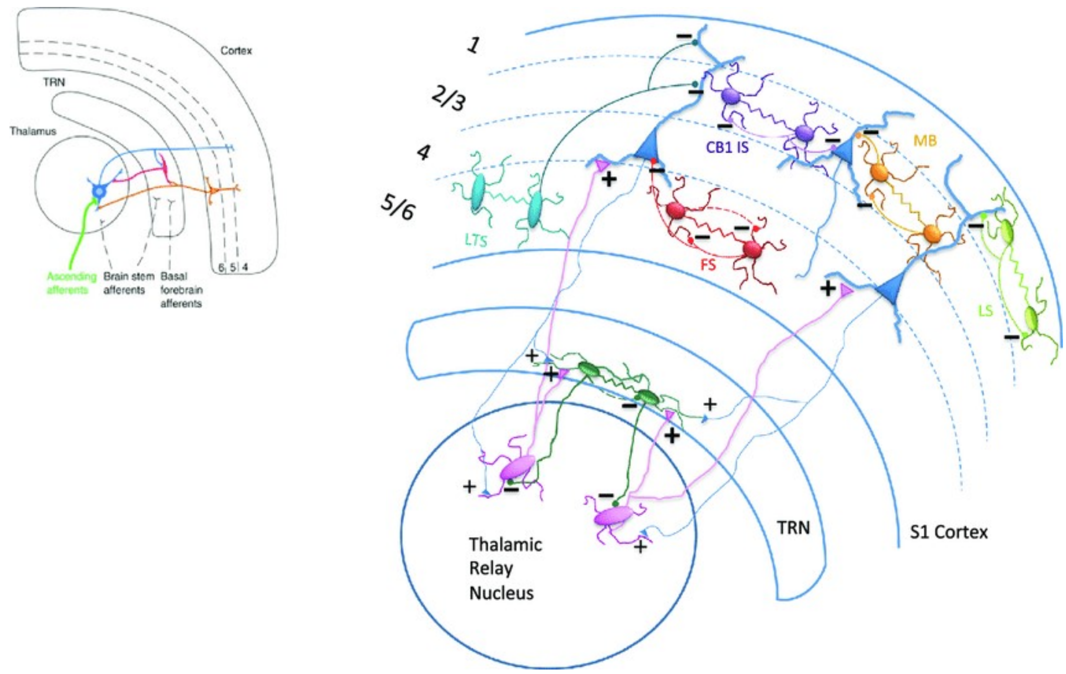


Figure 1. Scheme of thalamic cortical connections through different cortical layers. Neurons with the same colors have the same shape and functional role.

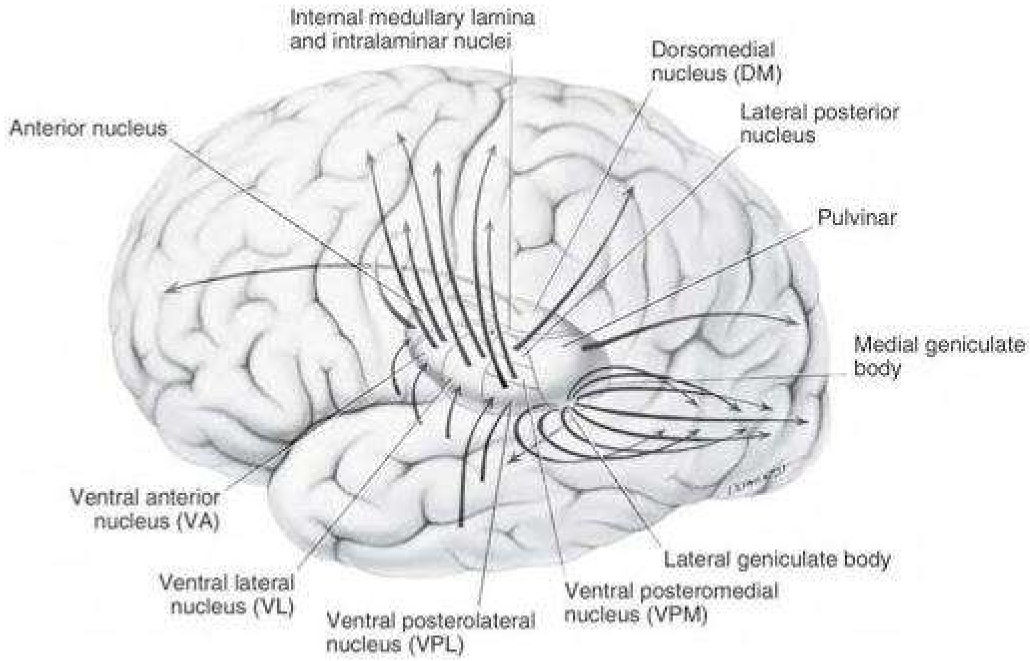


Figure 2 Scheme of the main directions of information flow that involves the thalamus.

The second topological organization is structured according to a complex of one-directional chains that connect in parallel the cortex to a set of its appendages. These mechanisms are involved in automatic action, or in those actions that would become automatic after a repeated reproduction of the gesture, no longer requiring a strongly conscious and careful participation of the agent (Fig.3)

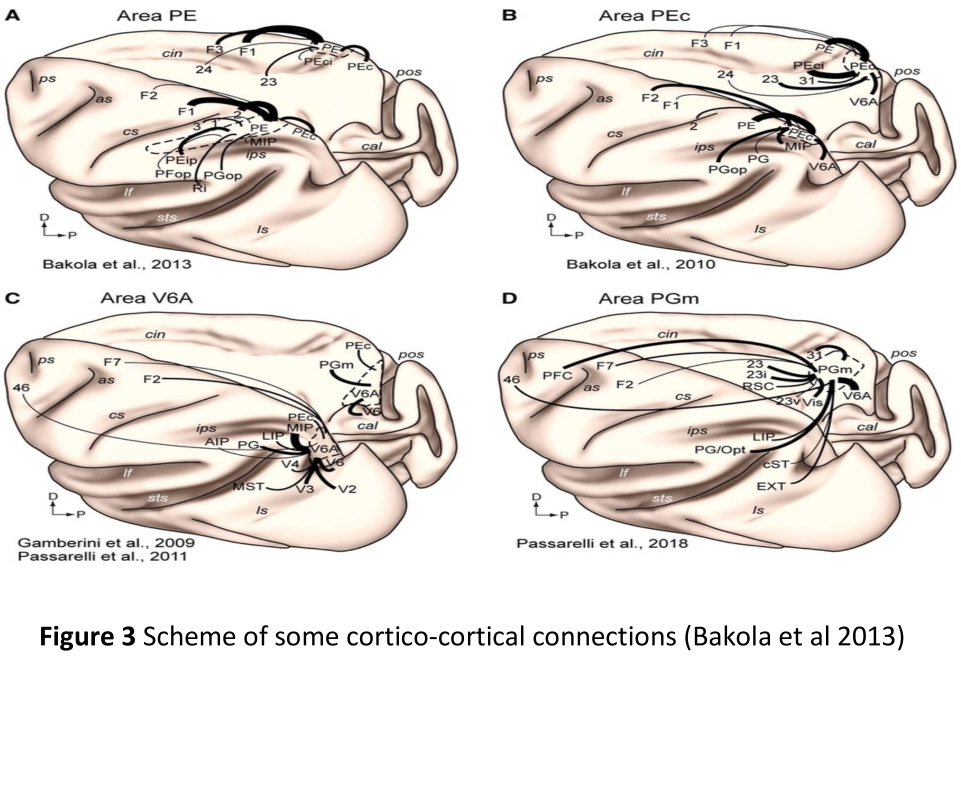


Figure 3 Scheme of some cortico-cortical connections (Bakola et al 2013)

Finally, there is a third type of organization that is different from the previous ones and consists of a widespread set of connections disposed of on a fan structure. The neurons included in these architectures seem to activate whenever something relevant in a small-time window happens and determines the flash activation of all surrounding structures.

This means that perception and learning processes not only depend on neuronal plasticity and the huge number of neurons present in each of the three morphological structures, but it also depends on the overlap of these structures that constantly interact, thus making even more complex to isolate the action circuits that would be connected to certain cognitive processes.

Attention—a process preliminary to consciously perceiving something and storing contents in memory, so that learning becomes possible—seems to be primarily associated to the functioning of the third morphological architecture. Attention selects, among a constant bombardment of stimuli, some stimuli that have a greater weight. Intersubjectively there is a world full of things and events, but a subject only perceives certain aspects of them, as we saw in Sect. 7.2. Subjects select certain pieces of information that depend on their mental states and the state of their synaptic connections.

The process of attention is the base of memory and learning. In fact, in order to be aware of something, it is necessary that attention be focused on it for a sufficiently long time. Some neurophysiologists have proposed that the ability to pay attention is strongly connected to the morphological and functional side. So-called *value systems* (LeDoux 2003) are complexes of neurons arranged according to a morphology of the third type (*fan*), which branch from the central structures to all cortical areas, and beyond to the entire brain. These very long branches are responsible for simultaneously informing the entire structure of sudden and significant changes that occur in the whole body of the subject.

This morphological aspect may be consistent with the functional explanation modelled by the theory of global workspace as described in Ch. 3. This morphological fan-structure could be also relevant to endorse another paradigm of consciousness explanation proposed by Dennett D.

But even if it is consistent with it, form a morphological point of view, I do not think that Dennett account captures an important aspect, because it is focused on hierarchy of contents, but it does not consider in appropriate way the timing relevance between them and the timing structure of the stimuli in the environment.

Dennett proposed his multiple drafts model of consciousness in 1991, which suggested that there is no special location in the brain that can be seen as the qualia-containing consciousness module. Instead, he suggested that consciousness is distributed throughout the brain. He extends the model by creating a similar figure that he calls metaphorically as *Fame in the Brain* (Dennett 2005). This is a functional minimalist theory that is very close to the workspace theory proposed by Baars exposed in Sect 3.2. Dennett suggests that the mind acts as an echo producer. Under this perspective, in order for something to be considered conscious, there must be enough correlating neural events that go with it.[19]

The difficult question, the *hard problem*, as Chalmers (1999) called it, exemplified by the famous bat of Nigel (1976) makes no sense to be asked. In Dennet's opinion there is no Cartesian Theatre in our brain in which all our conscious experiences are represented, but, more radically, there is not neither the sense of self, so our sense of agency is something that does not deserve to be explained. On the contrary, the brain is composed of contents that anarchically assume the right to be cerebrally popular, acquiring a certain fame.

Personally, even if I agree that consciousness has not a spatial location in the individual brain, I do not think that the method to avoid considering phenomenal properties is a good one to understand consciousness. It forgets its most relevant feature, and also I do not think that we can abstract timing components of the brain and environments in approaching consciousness. I think that our sense of identity and agency is something real that we need to take in account to approach consciousness, also I think that this is strongly connected to timing processes and could help to discriminate different learning processes. This is the topic of the next section, where I will argue for this.

[19] Dennett does not provide a conclusive function that connects a neural event and a conscious one, so it is difficult under this perspective to understand what he really means as *neural correlates*.

7.3 Timing in learning

Now that I have outlined some neural morphological aspects of the attentional process, it is time to investigate its timing aspects. Before approaching attention and conscious perception from the timing perspective, I want to talk about some crucial timing features of learning processes.

I talked a lot about learning from a material, morphological and informational point of view but this chapter wants to be mainly about time and timing and its role in learning processes. So, here I want to highlight one very relevant aspect of neural plasticity that here could be relevant in this timing context: Spike Timing Dependent Plasticity (STDP) in learning.

I talked a lot about Hebbian reinforcement that means the strengthening of synaptic junctions in Ch. 1, but I said nothing about the time window in which it operates. Spike timing dependent plasticity is precisely what guarantees the causal interpretation because it refers to the sequence of activations in a synaptic bilateral conjunction (Recurrent Synapse).

If we think about two neurons connected to each other and one activates for a stimulus that occurs before, then there will be a strengthening of the synapse in one direction (exactly the direction which starts with the neuron that activates with the first stimulus) and not in the opposite one (Fig. 4). This is demonstrated in several studies and could offer a possible argument for the way in which the brain interprets causality, just like a sequence of temporal activations. This is also coherent with a different level of explanation of causality, a psychological one, proposed by Hume (1746). In his opinion in fact, causality experience is only the experience of repeated connections in the time.

There is another important aspect that we should highlight here: the time window necessary to activate this connection has a grain of milliseconds, in fact an interval of a second is not efficient anymore. (Levy and Steward 1983)

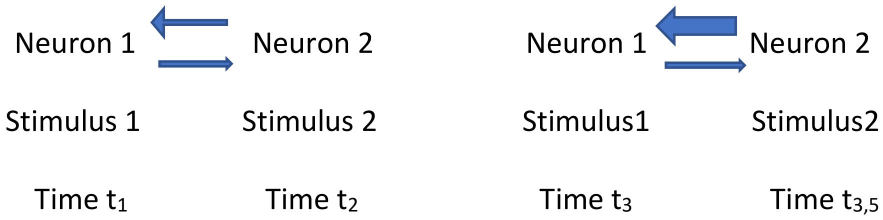


Figure 4 The scheme shows STDP, there is a strengthening of neural junction from neuron 1 to neuron 2 and also the time sensitivity of the second activation could increase generating a shorter activation.

7.3.1 Timing contiguity and ascription of real identity

This observation leads us to approach another problem that is crucial for the attribution of identity and, probably, for the sense of subjective action. This is the problem of the so-called size-distance invariance. Starting at a very young age, humans are able to recognize the identity of the same object despite its changing appearances in shape and size due to the distance and angle of perspective. More or less, all the features and aspects of a thing always change depending on distance, circumstances, light condition, movements, or perspective. But still, we are always able to recognize the identity of a thing despite all these different appearances and it seems that we possess this ability we are infants.

So, even if this is not meant to be a conclusive answer, I would propose that timing contiguity is fundamental for the ascription of identity. The way in which the mind is able to recognize features that must be considered the same despite their different appearance, in contrast to features that should be distinguished, depends on timing.

Timing generates the ability to discriminate between different things in the environment. If some activation configurations always occur with a very close temporal contiguity, they are associated in a meaningful unity by the subject. I approached individual body involvement in perception when I discussed Noè’s work in Ch. 3. My present proposal can be seen as an extension of the argument which considers bodily timing involvement.

7.3.2 How does the brain count time?

As mentioned before (Sect. 7.1.1), intense stimuli are represented at the timing level (in the sense that a stronger stimulus is represented by a higher frequency than the one that represents a weaker stimulus), and they have a stronger impact. It seems that attention (and conscious perception) is the result of a process of comparison between different stimuli through a proper timing neural structure, where the speediest activations seem to be the most implicated in attentional contents. Is there in the neural system a structure that acts as a metronome? May there be an internal clock that gives a unit of time to the system? A metronome independent of other structures that acts as a unit of temporal measurement and makes it possible to select high speed signals, which are more urgent with respect to others and so have the dignity to become conscious?

There are two different possible answers to this question that imply a completely different understanding of the workings of attention, and so of individual consciousness. First, we could say that there is a specific area that always pulses at the same frequency, which works as a timer (Gibbon et al. 1988). In this sense we have to suppose that it starts to pulse at some point of the development of the neural system. The features which define different contents of perception would then depend on those internal and subjective rhythms that are related to this original timing unit. Under this perspective, attentional content is a step-by-step sum of speedier activations compared to this timer. But, as I will argue in the next section, this is only one possible answer, for we should also consider the existence of multiple clocks and an alternative way of measuring time might depend on their integration or overlap.

There are several reasons why the mind should be able to detect and measure time. We need this to explain our understanding of synchronic perceptions that are not synchronic at all in the real intersubjective world. The ability of the brain to measure time is also needed for a simpler job like understanding the provenance of a sound[20] or for interpreting a sentence thanks to the capacity to correctly separate syllables and words.

[20] The brain must calculate, in a very short time, the distance that separates our ears, to be able to understand exactly where sound waves hit the body, and so where the sound comes from.

So, how can the brain measure time? Is there a specific area that is structured for this, or are there multiple interconnected areas that do this complex and fundamental job? And what is the relevance of these observations for distinguishing different learning processes and different senses of individual agency? The next section aims to answer these questions.

7.3.3 Multiple clocks and consonance

This section aims to make clear that it is exactly the timing capacity that is one of the key elements not only of learning, but also of our sense of identity and agency. This capacity is the most basic feature of the brain working, it connects an individual mind with its environment, it is crucial for the sharing function of the workspace and for predictive and comparing activities.

One of the most relevant rhythms in our life is the circadian one. When we talk about this, we refer to the regulation of the individual that determines a period of sleep and a period of activity. Surprisingly enough, many experiments demonstrate that it does not only depend on the environment light conditions. In fact, even if the individual is in a condition of artificially regulated light, or even in complete darkness, she is able to preserve a circadian rhythm (Lavie 2001). This makes us think of a specific area that is specialized in measuring time.

The suprachiasmatic nucleus was considered one of the best candidates, because when it was transplanted from an OGM rat (that was genetically modified to have a circadian period shorter than the normal one) to another normal rat, produced the same circadian rhythm in the new brain (Ralph 1998). But in humans this is not enough. It is not true that the suprachiasmatic nucleus is able to regulate all the other rhythms with a finer temporal grain. In fact, we need to measure different time periods and durations, depending on cognitive activities that process different stimuli with different timings.

If there are so-called *place cells* that activate when we occupy different regions of space, it seems that there are not analogue cells for time. It is true that there are so called *clock cells,* especially the cells of the suprachiasmatic nucleus, that exhibit a proper inner rhythm, in the sense that they cyclically realize different proteins and neurotransmitters (Reddy 2006).

This mechanism that regulates our circadian rhythm is known as Transcription Translation Autoregolatory Feedback Loop and it is organized through to the modulation levels of the protein named Period (Kanopka and Benzer 1971). But it is not true that they respond to stimuli in a cyclical way.

Moreover, we need to highlight a very important aspect of time measuring. Circadian rhythms organize events in the body (like the production of hormones and proteins) that are shorter than the circadian period. There is an infra-periodic measure and an over-periodic measure. The first measure is used to organize events that are shorter than the period. The circadian clock, in fact, has a period of 24 hours and, with respect to it, hormones and molecules are produced in accordance with a shorter period (many times in 24 hours). The second measure is used to count and organize events that are longer than the period.

Drug abuse studies and neuropharmacology are also useful to investigate clockwise functioning in the brain and also time perception. These substances are responsible for changes in neural activity because they have a mimic activity. This means that they act like endogenous neurotransmitters altering the transmission of information in a specific network for example, sometimes preventing the reuptake of some molecules, sometimes stimulating the abnormous production of some molecules. They also often have an impact on time perception, but this may be a misleading observation.

On the one hand, saying that time perception is altered implies the existence of an internal normal perception; this in turn presupposes an internal normal clock with respect to which perception of a duration turns out to be altered. But this is not a good argument, first because alteration of time perception strongly depends on the subject and the circumstances in which the drug assumption occurs, and second because there could be the same temporal alteration (for instance the perception that time passes very fast) with drugs that act on the individual system in opposite ways, like cocaine (stimulant) or marijuana (relaxing).

Moreover, some experiments testify that we are able to exercise, and so to improve, the ability to evaluate a given time duration without improving the ability to guess longer durations. This means that through the repetition of the same task, we become, step by step, more precise in estimating the duration of and experience that has a specific time window of duration, but we fail when we are supposed to estimate a thinner or bigger window also of that very experience (Wright et al. 2010). We need some systems that are sensible to very fine time scales and others that are sensible to coarser time scales.

Figure 5 We could try to figure out the work of different clocks in the brain through different clepsydrae that count different time portions. We could have different clocks that are sensible to different time grain that the brain uses to organize different activities, but in some specific moments they are coherent between them. Like the moment represented in the picture.

All these findings induce us to think about an alternative explanation of timing activities, that takes into account the multiplicity of relevant time windows. We should thus consider the existence of multiple clocks in the brain, that are sensible to different features and granularities of time, can perform different activities, and might be connected to the nature of our cognitive perception of time and agency. We could try to figure out this timing structure thinking about a variety of clepsydras with different dimensions, that count different portions of time, as shown in Fig. 5.

We can think, in fact, of many periodic oscillators with different frequencies that overlap each other only on some of their peaks, so that their partial consonance forms a periodic structure. In this case we do not have a pulse that functions as a measuring unit, but the measurement is due to the consonance between different clocks. This means that time is processed through a population code in which every neuron plays a distinctive role in the global dynamics of the system.

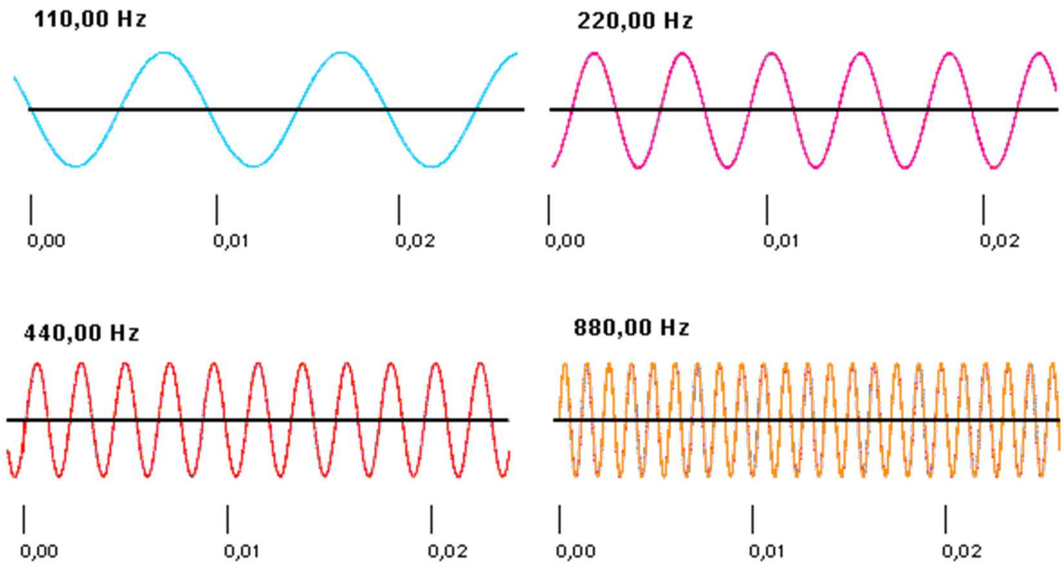


Figure 6 This example is taken by music harmony, but it is useful to visualize different clocks in the brain that show a coherent peak at some points (in this figure such points correspond to 0,006 and some of its multiples).

7.4 Timing and conscious agency perception in learning

Timing sequences could also be the root of different agency perceptions in different motor learning procedures. When I started this work the question that I moved from was about the difference that we can detect in voluntary (active) or afferent (passive) learning, despite in either case neural changes are more or less the same.

If we consider the learning process of a motor action, we see in either case neural changes due to neural plasticity that can be explained by Hebbian learning and STDP.

When we learn something new through a voluntary process, we have a strong perception of individual agency. At the same time a modification in our brain occurs, due to neural plasticity, that could represent the effect of the reproduction of the action. But we can obtain a strong modification in neural pathways also in the case of passive reproduction of a motor action through the use of protheses, orthoses or other devices.

As it was mentioned in the Introduction, some of these tools limit the degrees of freedom of action movements, and so are able to strengthen the networks that correspond to specific and effective movements (Pittaccio et al. 2017).

Other kinds of tools aim to improve learning through a passive stimulation of a peripheral configuration of body parts. An example is a robotic glove that touches sequentially and repeatedly a sequence of fingers, helping the subject to perform a song which corresponds to that sequence of fingers on the piano keyboard. It is important to stress that, in both these cases, the subjective sense of agency is not vivid as in the active learning cases.

Obviously, it is not possible to experimentally compare exactly the same areas and the changes in neural plasticity that are the effects of voluntary learning on the one side, or afferent learning on the other side. First, because of the enormous number of connections. Second, because experimental conditions to compare these two different ways of learning require two different subjects and, consequently, two different brains, or the same subject, but with the involvement of her two different hemispheres.

Nevertheless, if we consider all the studies on neurophysiological, morphological and temporal observations that we analyzed in the present work, we could try to delineate a timing-based explanation for the different senses of agency related to these two different kinds of learning.

In the case of voluntary learning, the causal process seems to reproduce the pathway of STDP, as the speed of a neural configuration related to a specific motor task becomes faster and faster with repetitions. In the case of conditioned learning, even if we want to reach the target of a learned output, the sense of agency is compromised by a different timing configuration. The difference is neither in the motor part that is active in the neural individual system, nor in spatially different connections. The reason that justifies different senses of agency might instead be primarily linked to different timing, or more precisely, to the way in which multiple clocks integrate each other.

This is not a reductionist theory of consciousness, because even if neural components are fundamental, what is more fundamental is timing coherence between internal multiple clocks and this is also modulated by the temporal architecture of environmental stimuli. But neither is this an extended theory because the relevance of the environment depends on timing proportionality and coherence, and not on its material components.

Moreover, my view is not very close to the global workspace framework. What is the difference? Differently from workspace theory - which we can briefly summarize as the sharing of contents processed by different functional modules - in my view this sharing is distributed in space and also in time. This means that the global workspace framework is based on a fragmentation of informational contents and spatial and functional properties of the patterns involved, while my view is mostly focused on timing architectures. Potentially, under my perspective, the same information could be processed in completely different brain locations if the temporal structure of the activations is preserved.

Under my perspective, consciousness is a matter of timing organization and it depends on the way in which different external stimuli, with proper temporal structure, interfere and are processed by different clocks in the neural system.

These multiple clocks are able to change information, and this determines further changes in the whole individual body, which is the interface or the medium of interconnection between external stimuli and the individual system.

The consonance in internal structures in the subject generates a change in the organization of multiple clockwise structures, and this interconnection may be seen as the onset of the conscious perception of agency in performing an action.

7.5 Methodological observations and practical applications of the timing-based framework

The present work aims to further develop some observations of neuro-phenomenology. This is obtained by applying them to timing processes to investigate a special feature of consciousness, namely, the sense of property of an action in a learned gesture. I moved from a phenomenal element: we perceive differently something that we voluntarily learn and something that we are conditioned to learn. Starting from this I investigated neural processes through the lenses of timing processes to incorporate environment in the productive process of the self.

With neurophenomenology the process of naturalization of phenomenology has begun. In the words of its founder Francisco Varela, neurophenomenology is "a way to marry modern cognitive science with a rigorous approach to human experience" (Varela, 1996, p. 1). Basically, it is a science that incorporates phenomenological investigations of experience into neuroscientific research on consciousness.

As we saw before, especially in Ch. 4 and Ch. 6, this step was started by Husserl, who was unsatisfied by reductionist explanations. Against the supporters of the empirical approach, interested only in the evidence of psychic facts and completely indifferent to the knowledge of the current structures of knowledge, Husserl objects that evidence exists only in his actual experience, which is renewed at every moment. A truth in fact can only be grasped in its becoming, which cannot be a computable set of psychic acts, but lived temporality. In this sense, the temporal process does not represent a real part of the flow of the lived, but an ideal immanence through which we can always return to acquired evidence. Consciousness of the temporality and temporality of the consciousness are, under this perspective, thought as a cycle.

From a theoretical point of view, neurophenomenology integrates three elements: phenomenology, the theory of nonlinear dynamical systems, and experimental brain science. This means that the explanatory hypotheses advanced by neurophenomenology always take into account the structural invariants of consciousness identified by phenomenologists and the data of brain activity measurements made with modern neuroscientific tools. The connection between these two types of data is provided by the theory of dynamical systems, which allows large-scale brain phenomena and conscious experience to be considered as the two poles of a recursive circle.

From a methodological point of view, however, the novelty of neurophenomenology lies in the method of reciprocal bonds. This method consists in guiding neuroscientific investigation on the basis of phenomenological analyses, and conversely. In this way, neuroscience broadens its perspective to include the world of experience as it appears to consciousness, while phenomenology is enriched by data from the biological domain.

The implicit feature of this mutual constraint approach is the attention to the dynamics of the correlation between the biological level and the level of consciousness. It is not a question of establishing a mere isomorphism between the nervous system and experience, nor merely of building correspondence between the two levels, but of establishing binding emergency relationships enshrined in principles preferably describable in quantitative terms (Petitot et al., 1999).

My target in the present work was to apply this method to the enquiry about the sense of agency in learning processes through a focus on timing, and my faith that this idea could be effective is supported by some promising experimental results in a similar application of these methods to different problems. A clear example is offered by neurophenomenological accounts to epileptic premonitory mental states, that I explored in Ch 6. Here I will briefly consider an analogous example by Le Van Quyen that nicely shows the potential of this methodology and presents some encouraging preliminary data (Le Van Quyen, Soss et al., 2005).

Central to neurophenomenology is the combination of quantitative measures of large-scale neural activity with detailed first-person descriptions of the categorical features of experience. Accordingly, as a guide for their neurodynamic analysis, the authors focus on the role of integrative neural mechanisms such as neural phase synchrony in epileptic seizures, and the collection of refined first-person descriptions of preictal states.

Without downplaying the critical role of this circuitry, it is worth noting that the authors' analysis of prodromic symptoms not only revealed bodily feelings, such as "the feeling of heat inside my body", but also a sense of decrease of energy, a lack of concentration, words or physical balance. The latter of these symptoms are more related to the pre-reflexive form of self-consciousness and are likely to involve a broad range of brain circuitries critical in speech, movement, and voluntary attention.

The authors conclude that prodromic symptoms seem to have "no localization value" in contrast to auras, which have well localized epileptogenic zones. Thus, they choose to study these preictal states as changes in the dynamic and systemic mode of interaction of the whole brain, rather than attempting to localize first-person features. Their analysis made possible the discovery of a "preictal state" characterized by a desynchronization of the neuronal assemblies related to the epileptogenic focus up to several hours before the seizure onset. This dynamic model not only provides a representation of a dynamic mode, or signature, compatible with the phenomenology of the state, but it also provides explicit predictions of the probability of a crisis occurrence.

This example shows that the neurophenomenological method starts by investigating phenomenal events and connects them not only to neural events that contemporarily happen, but also to a timing distributed structure which incorporates past events. This leads us to search for events which are always related to the emergence of a specific event, like an epileptic attack, in order to treat it in the appropriate way, even before its very emergence.

Most recent cognitive theories of consciousness hypothesize a minimum time needed for the emergence of neural events that connect to a cognitive event (Dennett, Kinsbourne, 1992). This non-compressible temporal structure can be analyzed, as well as the manifestation of long-range neuronal integration in the brain linked to a diffuse synchrony (a similar idea was discussed in Ch. 3). This link sheds light on the nature of phenomenological invariants through a dynamic reconstruction that is at their base, thus providing the synchronization process with a tangible experiential content (Varela 1997, p. 32).

Since this neurophysiological approach proved to be fruitful, I maintain that a neurophenomenological approach to timing could be equally fruitful to better investigate the sense of agency. Some practical impact could be found in the field of neural manipulation, for instance we could try to manipulate the frequency of a pattern of activations reproducing the same location but modifying its speed, preserving the temporal architecture through magnetic stimulation.

At the same time we could investigate the phenomenological perception of the subject to see if something changes in the sense of presence and in the temporal location that he can attribute to that specific event.

We could also use this model to approach some psychological diseases. As an example, we can try to manipulate traumatic memory by decreasing activations speed in those areas which are activated when the subject brings back to memory a traumatic experience.

In my opinion, if the two opposing pairs—processing of information vs. self-creation and representation of the environment vs. the construction of a significant world on a perturbing background—are analyzed through the lens of timing, they are brought together in a dynamic circularity that can even be fruitful from a practical point of view.

By the famous expression "consciousness is not in the head" (Varela, 1997) the neurophenomenological approach stressed that the affective mind is not limited to neural events but is distributed through the whole body. The affective-temporal states emerge from a mutual co-determination and eco-implication between mind and body. Time is, under this approach, perceived as a phenomenon: a phenomenon—at the same time semantic and pathic—that is a lived experience and not a measure of change. Time, therefore, is not an arithmetical meter, but incarnate temporality, incarnate mind, lived corporeity (Merleau-Ponty 1969).

In my work I wanted to push further these considerations up to saying that time and time proportion is the determining factor in information coding, in emergence of memory and in the possibility of learning. However, my view differs from standard neurophenomenology in that it also considers the temporal structure of the environment, and not only the one of the subject’s body. In my hypothesis, the coding is due to a decrease or, more generally, to a modulation in the speed of the activations that make up a perceptual pattern, since perception has been shown to be temporally structured.

Timing is, at the meantime, crucial for consciousness of the ownership and intentionality of a motor gesture, since it has been shown to be attributable to a different temporal architecture of the areas involved in the gesture learned. This motivates the idea to approach learning through an embodied, situated and timing-based conception.

This situated and timing-based approach to consciousness in learning could even have some interesting implications in the practice of electrical stimulation through magnetic tools such as TMS, rTMS, tDCS, tACS, tRNS, tES, or other devices used for the enhancement of learning performance, to better direct the activity and effect that these these tools have on individual brain.

Moreover, this kind of perspective could help us to approach some ethical issues that the use of these tools often encounters. In fact, it may be useful to overcome some misleading difficulties about the subjective impact that the clinical application of these tools sometimes presents, shedding a new light on the meaning of voluntary action and its possible representation in this context.

Some other practical applications might be in the design of a productive environment to improve learned output, structuring settings that take into account the aspects exposed: the relevance of timing components and of surroundings features in individual performance. These possible applications will be the matter of subsequent dedicated research.

Conclusion

In this work have delineated a situated and timing-based model for consciousness in two different learning processes. At first, in Ch. 1, I approached learning from a neurophysiological point of view, as a process of remapping of neural configurations that occurs thanks to neural plasticity. I discussed the relevance of different aspects of environmental influence on neural plasticity during the development of the brain, from the dimension of the cellular environment, those structures that surround neurons, up to a much higher dimension, the role of parents' behavior in pregnancy. I considered the relevance of the environment in neural change throughout life and I traced a situated approach to learning from a morphological point of view, by highlighting the active role of the environment in both these processes.

I then approached the composition of the environment in the individual neural system. *Multimodal Integration*, the process that stands between stimulus and response, refers specifically to the temporal architecture of subjective perception. Stimuli have a temporal structure that differentiates them, a temporal distribution, and they impact on the spatial construction of the body that processes them depending on this aspect. This, through multiple evidence, has been delineated as a highly integrated and semantically oriented reconstructive process. The intersubjective dimension becomes very important, because the neural resonance generated by mirror mechanisms strongly interferes with the subject's elaboration and distribution of weights within his/her system, but also because the external observer contributes to the normalization of discordant or contradictory perceptions. This assumption of reciprocal integration between systems was justified in the light of the neurophysiological analyses of timing, and also by appealing to theories that underlie the aspects of continuity and the dynamics of the mind.

In Ch. 3 I discussed the notions of *integration* and *interaction between fields,* associated with consciousness. I highlighted the implications of these concepts with respect to a typological and morphological categorization of stimuli, and their limitations. I argued that consciousness cannot be understood within a fragmented conception of causality. I explored global workspace theory and I traced the possibility of a different paradigm that, in order to better understand this fundamental feature of consciousness and its links with time and unity, is based on some reflections that belong to the theoretical apparatus of phenomenology. With these tools, I argued that perceiving is an embodied relationship to the world. The body and its movements are strongly involved in the generation of perception. I used a central aspect of the phenomenological reflection on temporal consciousness that is rooted in an equally dynamic and extended interpretation of perception. Perception is a complex structure that hosts in itself more or less extensive and intricate remembrances and also expectations. Our perception of time is structurally extended and composed of different degrees of existence and presence.

One of the most extraordinary abilities of the human brain, useful for learning, is that it is constantly projected to the future. We figure out possible conditions, expected because we have an available set of data that could be coherent with a set of outputs. In addition, on a finer scale, during the performance of an action, it seems that we constantly form a model of the world, and our perception tries to anticipate the perception that will be connected to a specific change in the conditions of our activity.

In Ch. 4, I explored this aspect, which is crucial in the predictive coding approach. The effectiveness of our actions seems to require anticipation of those world conditions that different possible actions will produce. Under this perspective, in order to act and reach a desired target, it seems that something that drives our movements is anticipated by predictions of our brains. Despite some limitations that I highlighted of this predictive perspective, an interesting point is that it focuses on the relevance of what is unexpected. It is exactly what differs from what is expected that passes through hierarchical levels, arrives at being conscious, and has the strongest impact on future model-based predictions.

In Ch. 5, I focused more specifically on the ego perception, by discussing a predictive approach to the notion of *self* and starting from an analysis of interoceptive interference. Then, I explored the possibility that this predictive coding framework might have some conceptual anticipations in German idealism. I considered some aspects that Fichtian idealism shares with the constructive likelihood principle. I then explored another conceptual basis, which idealism, in its third step, may offer to the development of enactivism and 4E cognition theories. With these tools, I explored the theme of reflexivity, or the role of action in self-perception and identity ascription.

These considerations support the view that the essence of personal identity should not be searched merely inside the subject, but that instead it has a situated and timing-based character. Phenomenal knowledge (which is typically meant to be indubitable because it is intrinsically perspectival and subjective) arises from reflection, in the sense of being both sensory extrinsic and dependent on an inseparable interconnection with the other and the environment. In a more narrative sense, personal identity contains a multiplicity of aspects that also concern the outside of the very subject. Identity is something that refers to how others see us and what we see of ourselves. We carve out our personal identity within possibilities of action that we learn by reflection and mental representation of our limited corporal extent and the actions of others, on which we model our actions.

Memory seems to be the faculty that sustains any possible learning and keeps together our sense of personal identity. This was the main theme of Ch. 6, where I investigated the connections between personal identity and memory through a paradigmatic case: musical memories and self-generated ones, when the boundaries between what happened and what is happening collapse. I argued that memory is not a reproducible and immutable form that can be enucleated from the flow of consciousness, implemented as an epiphenomenon of a localized and unusual activation in the temporal cortex. I considered the possibility that a specific memory is not represented by the location of a specific pattern of activation identically repeated under some specific conditions. A pattern is in fact specified by two main features: the first one is the spatial location of the active units, but the second one is the temporal distance between their activations. The focus of my perspective is on this second feature.

I argued that a way of coding memories might be the relative proportionality between the temporal distance of neural activations. This means that a memory content is a pattern of activation that could be spread on different units of time, as the proportionality between patterns of activation is preserved. Furthermore, a timing-based model seems to be advantageous from the energetical point of view.

Musical memories in all their nuances, from humming inside to hallucinations, are very interesting because they entertain a special link with the construction of the ego. Perhaps this depends on the essential link that music vibes weave with time. In the same unit of time music involves the individual system with a bombing of stimulations that have a specific, temporally coded, internal architecture.

In Ch. 7, I recollected some timing aspects explored in the previous chapters in order to propose a timing-based perspective for understanding the neural changes that occur in learning processes, and I tried to propose a paradigm of explanation that, by taking into account the active role of the notion of synchronicity and coherence, could help to shed light on the differences between conditioned and voluntary learning and the different conscious perception associated to them. Namely, I focused on a timing-based level of explanation of cognitive processes that are involved in the learning process, in particular attention, memory and conscious perception of agency to argue that causality attribution and agency ascription in individual learning, as well as the very concept of identity, all depend on a timing relationship.

To do this I started focusing on informativeness and the relevance of the timing component, according to the basic assumption that morphology is manipulated by information, which is codified in neural spiking. This means that learning is possible because of a previous change in the kind of coding and decoding in neural cells. What changes is the way in which neural cells codify information and send it to the cells that are connected to them. What is relevant for the occurrence of a learned output is consonance between different areas, or groups of neurons.

The neural communication code has a temporal component, which describes the correspondence rules by which the electrical activity of a neuron encodes incoming signals and transmits the processed signal to subsequent cells. I briefly summed up some of the possible coding interpretations to conclude that, despite their differences, they share an important feature related to attention. It seems clear that the intensity of stimuli can be encoded in some way through the number of spikes in a time unit.

The timing of the neural activations conveys information about the stimulus and its processing can indicate something relevant also about the conscious content connected to that stimulation. We must also observe that intense stimuli involve our attention in a stronger way. Intense contents are more likely to become conscious because they manifest more urgency, as they require a response. I considered the hypothesis of a coherent neural implementation of the workspace theory and the predictive coding theories of attention and perception.

Then, I analyzed more deeply the timing properties of information in attention processes, to try and define a timing-based paradigm that could help to differentiate between voluntary learning processes and involuntary ones, as well as among the different agency perceptions connected to them.

The paradigm that I proposed incorporates some relevant aspects of epiphenomenal, predictive, and enactivist theories, by considering timing as the connection between subjective internal processes and external stimuli.

Timing contiguity is fundamental for the ascription of identity. The way in which the mind is able to recognize features that must be taken to be the same despite their different appearance, in contrast to features that should be distinguished, depends on timing. Timing generates the ability to discriminate between different things in the environment.

I argued for the need of multiple clocks in the brain to organize information and coherence, and I surmised that the timing capacity is one of the key elements not only of learning, but also of our sense of identity and agency. This capacity is the most basic feature of the brain working, and it connects an individual mind with its environment. These multiple clocks in the brain are sensible to different features and granularities of time, can perform different activities, and might be connected to the nature of our cognitive perception of time and agency. Many periodic oscillators with different frequencies overlap each other only on some of their peaks, so that their partial consonance forms a periodic structure. In this case we do not have a pulse that functions as a measuring unit, but the measurement is due to the consonance between different clocks.

I used all these tools to shed some light on the different senses of agency that occur in two opposite learning processes. In the case of voluntary learning, the causal process seems to reproduce the pathway of STDP, as the speed of a neural configuration related to a specific motor task becomes faster and faster with repetitions. In the case of conditioned learning, even if we want to reach the target of a learned output, the sense of agency is compromised by a different timing configuration. I proposed that the difference is neither in the motor part that is active in the neural individual system, nor in spatially different connections. The reason that justifies different senses of agency might instead be primarily linked to different timing, or more precisely, to the way in which multiple clocks integrate each other.

Under my perspective, consciousness is a matter of timing organization and it depends on the way in which different external stimuli, with proper temporal structure, interfere and are processed by different clocks in the neural system. These multiple clocks are able to change information, and this determines further changes in the whole individual body, which is the interface or the medium of interconnection between external stimuli and the individual system. The consonance in internal structures in the subject generates a change in the organization of multiple clockwise structures, and this interconnection may be seen as the onset of the conscious perception of agency in performing an action.

Bibliography

Adrian E.D. (1926), The impulses produced by sensory nerve endings. *The Journal of physiology, Mar 18, 61* pp. 49-72

Agostino (398), *Confessiones*
trad. it. di Vitali C., *Le confessioni*. Collana BUR, Milano, Rizzoli Editore, 1958

Baars J.B. (1983) Conscious Contents Provide the Nervous System with Coherent, Global Information. *Consciousness and Self-Regulation* pp. 41-79

Baars J.B (1997) In the theatre of consciousness. Global Workspace Theory, a rigorous scientific theory of consciousness*. Journal of Consciousness Studies*, Vol 4, N. 4, 1 April, pp. 292-309

Baars J.B., Newman J., Taylor J.B. (1998) Neuronal mechanisms of consciousness: a Relational Global Workspace framework*, Toward a Science of Consciousness II: The second Tucson discussions and debates. Cambridge,* MA: MIT Press, pp. 269-278

Barret L. (2017), The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *January,* pp. 1-23.

Bergson H. (1896), *Matière et mémoire*, Philosophie Demi-Toil
trad. it. Pessina A., *Materia e memoria*. Lateza (2009).

Beurle R.L. (1956), Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society of London*. Series B, Biological Sciences, Vol. 240, No. 669

Broad C.D. (1947), *The Mind and its Place in Nature*. Kegan Paul, Trench, Trubner & Co

Cajal R. (1984), *The neuron and the Glia Cell.* Charles C Thomas Pub Ltd

Cao L., Thut G., Gross J. (2017), The role of brain oscillation in predicting self-generated sounds. *NeuroImage 147*, Feb 15, pp.895-903.

Caporale N., Dan Y. (2008), Spike Timing–Dependent Plasticity: A Hebbian learning rule. *Annual Review of Neuroscience 31* February, pp.25-46

Carey A.A., Tanaka Y., Van der Meer MA. (2019), Reward revaluation biases hippocampal sequence content away from the preferred outcome. *Nature Neuroscience 22*, pp.1450-1459

Chabris C., Simons D. (2011), *The Invisible Gorilla. How our intuitions deceive us.* Harmony Books

Chabris C., Simons D., Schnur.T (2002), Evidence for Preserved Representations in Change Blindness. *Consciousness and Cognition 11*, pp. 78-97

Chalmers D. (1996), *The conscious mind. In search of a fundamental theory.* Oxford University Press

Clark A. (2016), *Surfing uncertainty: Prediction, action, and the embodied mind*. New York, Oxford University Press.

Clore G. L., Ortony, A. (2008), Appraisal theories: How cognition shapes affect into emotion. *Handbook of emotions*. The Guilford Press. pp.628-542

Colombetti G. (2013), *The feeling body. Affective science meets the enactive mind.* MIT Press 2013

Damasio A. (2000), *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*, Mariner Books London

Damasio A. (2010), *Self Comes to Mind: Constructing the Conscious Brain,* Pantheon

Dehaene S., Naddachè L. (2001), Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition 79*, April, pp. 1-37

Dehaene S., Kerszberg M., and Changeux J.P., (1998) A neuronal model of a global workspace in effortful cognitive tasks. PNAS November 24, pp. 14529-14534

Dennett D. (1991), *Consciusness explained.* Penguin

Dennett D. (2005), *Sweet dreams. Philosophical obstacles to a science of consciousness.* MIT Press

Dennett D., Kinsbourne M. (2011) *Time and the observer: The where and when of consciousness in the brain.* Cambridge University Press

Desmaisons D., Vincent JD., Lledo PM. (1999), Control of Action Potential Timing by Intrinsic Subthreshold Oscillations in Olfactory Bulb Output Neurons. *The Journal of Neurosciences 19*, pp.10727-10737

Eccles C. J. (1994), *How the Self Controls Its Brain.* Springer

Eccles C. J., Popper K. (1977), *The Self and Its Brain* Springer, Berlin
trad. it. Mininni G., Barbara Continenza - L'io e il suo cervello. Strutture e funzioni cerebrali. Armando Editore, 1981 Vol2
trad. it. Mininni G., Barbara Continenza - L'io e il suo cervello. Dialoghi aperti tra Popper ed Eccles, Armando Editore, 1981 Vol 3
trad. it. Mininni G., Barbara Continenza - L'io e il suo cervello. Materia coscienza e cultura. Armando Editore, 1981 Vol 1

Edelman G. (1993), *Bright Air, Brilliant Fire: On Matter of The Mind*. Basic Books

Edelman G. (1988), *Neurobiology. An introduction to Molecular Embriology*. Basic Books, New York

Edelman G., Mountcastle V.B. (1978), *The Mindful Brain. Cortical Organization and the Group-Selective Theory of Higher Brain Function.* MIT Press

Edelman, G. (1987), *Neural Darwinism. The Theory of Neuronal Group Selection*. Basic Books, New York.

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

Fichte J.G. (1794), *Grundlage der gesammten Wissenschaftslehre*
trad it. Boffi G., *Fondamenti dell'intera dottrina della scienza.* Bompiani 2003

Friston K. (2005), A theory of cortical responses. *Philosophical Transactions of Royal Society*, *Biological sciences 360*, Royal society publications, pp. 815-836.

Gibbon J. et al (1988) Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, Vol 95, Jan pp. 102-114

Gibson J. (1979), *The ecological approach to visual perception*. Houghton Mifflin, Boston

Graham Brown T., Sherrington CS. (1912), On the Instability of a Cortical Point. *Proceedings of Royal Society* Vol. 85, No. 579 pp. 250-277 Royal Society

Heidegger M. (1927), *Sein und Zeit*, Tübingen, Niemeyer
trad. it. di Chiodi P., *Essere e tempo* rivista da Franco Volpi, Milano, Longanesi, 2005

Heilbron M., Chait M. (2018), Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex?. *Neuroscience Oct 1 389,* pp. 54-73

Hoekzema E. et al. (2017) Pregnancy leads to long-lasting changes in human brain structure. *Nature Neuroscience 20*, p. 287–296

Hume D. (1748), *An Enquiry Concerning Human Understanding*
trad it di *Dal Prà M., Ricerca sull'Intelletto umano,* Laterza 2009

Husserl E. (1913), *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie.* Erstes Buch: Allgemeine Einführung in die reine Phänomenologie" in Jahrbuch für Philosophie und phänomenologische Forschung Halle a. S.: Max Niemeyer 1913
trad. it. di Filippini E., *Idee per una fenomenologia pura e per una filosofia fenomenologica. Vol. 1: Introduzione generale alla fenomenologia pura.* Torino, Einaudi, 1976
trad. it. di Filippini E., *Idee per una fenomenologia pura e per una filosofia fenomenologica. Vol. 2: Ricerche fenomenologiche sopra la costituzione.* Torino, Einaudi, 1976
trad. it. di Filippini E., *Idee per una fenomenologia pura e per una filosofia fenomenologica. Vol. 3: La fenomenologia e i fondamenti delle scienze.* Torino, Einaudi, 1976

Huxley T. (1923), *Method and results.* Nabu Press 2012

James W. (1890), *The Principles of Psychology.* Holt and Co, New York.

Konopka R.J., Benzer S. (1971), Clock mutants of Drosophila melanogaster. *Proceedings of the National Academy of Sciences, September 68,* pp. 2112-2116

Konorski J. (1948), *Conditioned reflexes and neuron organization*. Cambridge University Press

Larrabe M.G., Bronk D.J. (1947), Prolonged facilitation of synaptic excitation in sympathetic ganglia. *Neurophysiology 10*, March, pp.139-54.

Lavie. P (2001) Sleep-wake as a biological rhythm. *Annual review of psychology Vol. 52,* pp. 277-303

Le Doux J. (2002), *Synaptic Self*. *How Our Brains Become Who We Are.* Penguin Putnam.

LeDoux J. (2003), The Emotional Brain, Fear, and the Amygdala. *Cellular and Molecular Neurobiology November* 23, pp. 727-38

Le Van Quyen, Soss et al. (2005) Preictal state identification by synchronization changes in long-term intracranial EEG recordings. *Clinical Neurophysiology*, *vol. 116, March 2005*, pp. 559-568

Levy W.B., Steward O. (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience Vol 8*, Issue 4, Aril, pp. 791-797

Livingston R.B. (1966), Brain mechanisms in conditioning and learning. *Neurosciences Research Program Bulletin* pp. 349-354

Lloyd D.P. (1949), Post-tetanic potentiation of responses of monosynaptic reflex pathways of the spinal cord. *Physiol*ogy 33, pp.147-170.

Marshour G.A. (2013), Consciousness, Anesthesia, and Neural Synchrony. *Anesthesiology July*, Vol. 119, pp 7-19

Martin S.J. Morris R.G.M., (2002) New Life in an Old Idea: the synaptic plasticity and memory hypothesis revisited. *Hippocampus 12* pp. 609-636

Maturana H.R., Varela F.J. (1980), *Autopoiesis and Cognition. The Realization of the Living* D. Reidel Publishing Company U.S.A

trad it. di Stragapede A., *Autopoiesi e cognizione. La realizzazione del vivente,* Venezia, Marsilio 1985

Manzotti R. (2017) *The spread mind. Why Consciousness and the World Are One.* Books LLG, New

trad it. di Panini A., *La mente allargata. Perché la coscienza e il mondo sono la stessa cosa*, Il Saggiatore 2019

McGurk H., MacDonald J. (1976), Hearing lips and seeing voices. *Nature 264*, pp.746-748

Merleau-Ponty M. (1945), *Phénoménologie de la perception.* Paris Gallimard
trad. en. Smith C., *Phenomenology of Perception* New York: Humanities Press, 1962

Mikhail I. Rabinovich J. Friston K., et al. (2012), *Principles of Brain Dynamics: Global State Interactions*. Mitpress

Morgan C.L. (1923), *Emergent Evolution.* Henry Holt and Co

Nagel T. (1974), What Is It Like to Be a Bat? *The Philosophical Review. 83*, pp. 435–450. *Nature Neuroscience 22*, pp. 1450-1459

Noë A. (2004), *Action in perception*. Cambridge, Mass: MIT Press.

Noë A. (2012), *Varieties of Presence*. Harvard University Press.

Nunez P.L. (1974), The brain wave equation: A model for the EEG. *Mathematical Biosciences* Vol 21, December, pp 279-297.

Ondobaka, S., Kilner, J., & Friston, K. (2017), The role of interoceptive inference in theory of mind. *Brain and Cognition 112*, pp. 64-68.

Panzeri S. et al. (1999). Correlations and the encoding of information in the nervous system". Proc Biol Sci. 266 pp. 1001-12.

Penfield W., Jasper H. (1954), *Epilepsy and the functional anatomy of the human brain*. Little Brown, Boston

Petitot J., et al. (1999), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science.* Stanford University Press

Pittaccio S. & al. (2013), Passive ankle dorsiflexion by an automated device and the reactivity of the motor cortical network. *5th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*

Pockett S. (2012), The electromagnetic field theory of consciousness: a testable hypothesis about the characteristics of conscious as opposed to non-conscious fields. *Journal of Consciousness Studies 19*, pp.81-90

Ralph M.R., Foster R.G., Davis F.C. (1990), Transplanted suprachiasmatic nucleus determines circadian period. *Science 23, Feb Vol. 247, Issue 4945,* pp. 975-978

Ratcliffe M. (2005), William James on Emotion and Intentionality. *International Journal of Philosophical Studies* Vol. 13, pp. 179-202

Reddy. A.B *& al.* (2006), Synchronization and Maintenance of Timekeeping in Suprachiasmatic Circadian Clock Cells by Neuropeptidergic Signaling Current.
Biology, Vol 16, pp. 599-605

Rose S.A. Handsaker R.E., Daly M.J., Carroll M.C. (2016), Schizophrenia risk from complex variation of complement component 4. *Nature 530* pp. 177-183

Rossi S. (2020), *Il cervello elettrico*. Raffaello Cortina Editore

Sacks O. (1985), *The Man Who Mistook His Wife for a Hat.*
trad. it. di Morena C., *L'uomo che scambiò sua moglie per un cappello.* Collana Biblioteca n.174, Milano, Adelphi, 1986

Seth A. (2013), Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive science* Vol. 17 November, pp. 565-73

Seth A., Keisuke Suzuki K., Critchley H. (2012), An interoceptive predictive coding model of conscious presence. *Frontiers of Psychology 2*, *January,* 395

Sumby W.H., & Pollack I. (1954), Visual contribution to speech intelligibility in noise.
Journal of the Acoustical Society of America 26. pp. 212-215

Teufel C., Flecher PC. (2020), Forms of predictions in nervous system. *Nature Reviews Neuroscience* Vol 21, pp 231-242

Thompson R. F. Spencer W.A. (1966), Habituation: A model phenomenon for the study of neuronal substrates of behaviour. *Psychological Review 73*, pp.16-43.

Tommaso (1264), Liber de Veritate Catholicae Fidei contra errores infidelium
trad it. di Centi T. *Somma contro i Gentili,* Torino, UTET, 1975

Tononi G., Massimini M. (2013), *Nulla di più grande.* Baldini & Castoldi

Tononi G., Oizumi M., Albantakis L. (2014), From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory *3.0. PLoS Computational Biology, May 8*

Wilson H.R. (1973), Cowan J.D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetic 13*. October, pp. 55-80

Woodbrige P. (1957), Changing concepts concerning depth of anesthesia. *Anesthesiology 18* pp. 536-50

Varela F.J. (1996), Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, pp. 330-349

Acknowledgements

This work has greatly profited from the collaboration and support by many people. It is especially important to mention them here because they have been the very condition of possibility for this research. As I have argued in this dissertation, every kind of learning is possible and conceivable only by virtue of a profitable environment that allows for its evolution. Many ideas in this work grew up from a continuous and fertile exchange with extremely competent people during the whole course of my doctoral program. Unfortunately, the following list can only be incomplete. My special thanks goes to my family for their constant, unshakable and indispensable support. My supervisor, Prof. Marco Giunti, accompanied me along the entire path of my project, he spurred me on, motivated, corrected, and always provided me with precious suggestions, useful and interesting thematic insights. He taught me a lot, both from the strictly technical point of view and from the one of research life and methods. I also thank my colleagues in the Department of Pedagogy, Psychology and Philosophy of the University of Cagliari, who were always open to dialogue and welcomed me into a dynamic, motivated and cohesive group. My visiting experiences at foreign universities were periods of enormous growth in many respects. I wish to especially thank Prof. Giovanna Colombetti, my supervisor during my visit at Exeter University. She made the time I spent there plenty of stimulating experiences and useful advice, she helped me to deepen a fertile line of inquiry with respect to perception and affectivity, and gave me many opportunities to present and discuss my ideas. Prof. David Yates, my visiting supervisor at the University of Lisbon, involved me in many useful activities, discussions, and extremely fruitful dialogues, especially with regard to several aspects of predictive coding. Professor Robert Clowes oversaw my activities at Nova University of Lisbon and allowed me to present and discuss my research in highly stimulating contexts. Prof. Ana Sebastião accepted me in the *Mind and Brain Program* and gave me the opportunity to develop and broaden my skills in the field of neurophysiology. Finally, I would like to thank all my friends and colleagues in the foreign departments I visited. I worked side by side with them and we really shared a very important part of this thrilling and beautiful journey.