POLITECNICO DI TORINO

MASTER THESIS

Disorders of Consciousness: using the Perturbational Complexity Index to distinguish between voluntary and involuntary movements

Author:
Federica Sgrò

Supervisors:
Gabriella Olmo
Vito De Feo

ICT for Smart Societies

16 April, 2019

POLITECNICO DI TORINO

Abstract

Collegio di Ingegneria Elettronica, delle Telecomunicazioni e Fisica (ETF)

ICT for Smart Societies

Disorders of Consciousness: using the Perturbational Complexity Index to distinguish between voluntary and involuntary movements

by Federica Sgrò

The debate behind the word "*consciousness*" has been vivid for centuries among ethicists, philosophers and neurologists. Distinguishing between different **Disorders of Consciousness (DOCs)** in after-coma patients, in particular, is still difficult, and the diagnosis is based on scales whose assignement may be subjective and thus mistaken. Finding an accurate mathematical way to interpret a patient's movements and understand if they are completely voluntary or just involuntary reflexes may be a more objective way to classify their level of consciousness.

This thesis work proposes to find a way to distinguish among voluntary and involuntary movements in healthy subjects by analysing the **Event Related Potentials (ERPs)** behaviour obtained from EEG recordings. For each subject and experiment (voluntary, semi-voluntary and involuntary actions), after recording the EEG signals while performing the movements, the ERPs were extracted and a measure of the Integrated Information Theory firstly proposed by Giulio Tononi, called **Perturbational Complexity Index (PCI)**, was calculated on the basis of the state transitions before and after the on-set of the movement.

The values obtained from the algorithm were finally used to validate two different classifiers, one based on the **Minimum Distance Criterion** and the other using **Support Vector Machines (SVM)**. In both cases, significant results were obtained in the case of the binary classifier (considering only voluntary and involuntary movements), and in particular 88.9% specificity and 77.8% sensitivity were achieved with the minimum distance and 66.7% specificity and 83.3% sensitivity with SVM. The multi-class classifiers, however, led to many errors due to the ambiguous nature of the semi-voluntary experiment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The debate behind consciousness has been vivid for centuries, and it still is largely discussed in many disciplines, such as philosophy, ethics, and neurology. This introductory chapter aims at clarifying the difference between conscious and unconscious state and at coupling the most abstract aspects of the subject with a more scientific approach from which we can start our work.

1.1 Consciousness

1.1.1 What is consciousness?

The first problem that arises in writing this work is understanding what consciousness is, and especially finding a uniform way to define it and calculate it. Yet, whoever tries to look up for the term on a dictionary or on the internet will probably be overwhelmed by the vastity of different definitions and meanings given for the same word.

Here are just a few examples taken from some of the most influential english dictionaries:

1. *"The state of being awake and being able to think and notice things"*, Cambridge Dictionary;
2. *"The state of being aware of and responsive to one's surroundings. A person's awareness or perception of something; the fact of awareness by the mind of itself and the world."*, Oxford Dictionary;
3. *"[...] The upper level of mental life of which the person is aware as contrasted with unconscious processes."*, Merriam-Webster Dictionary;
4. *"The infinite awareness of the present"*, Urban Dictionary.

If we try on our own to give a meaning to the word *"consciousness"*, all the above definitions seem to suit, despite the differences in one another. And if we try to extend to other areas of knowledge, this scattering of interpretations gets only worse. *"Consciousness"* may be the awareness of mental events for psychology, the ability to distinguish between good and evil for ethics, the competence to separate the self from the external world for psychiatry, or it may even be referred to Karl Marx's class consciousness. In philosophy, each author gave its own interpretation, varying widely across the history and

even differentiating among a myriad of types and categories. This can give an idea of how the topic is still lively and deep within the different disciplines, and how many books will still be written discussing the vasitity of meanings the word *"consciousness"* may contain.

1.1.2 Wakefulness and Awareness

The previous subsection probably caught the reader into a vortex of expressions and definitions from which it seems practically impossible to get out. However, two words seem to be recurring: *awake* and *aware*. From this, the *consciousness* definition offered by neurology and neuroscience seems to be the most suitable and most appropriate to start to work on. [9]

In neurology terms, consciousness is represented by two components:

- **Wakefulness**: Characterized by a vigil state, fundamental for the wake-sleep cycle;
- **Awareness**: Perception of the self and the environment.

The conscious state exists only if both elements are present, but it's possible to observe wakefulness in absence of awareness. However, the opposite is never achievable.

Everyday we experience the loss of consciousness as soon as we fall asleep, or maybe someone sensed it while fainting or under the effects of anesthesia. This, however, are all temporary states which last only for hours, very different from the states of uncosciousness that may happen after a trauma or an accident.

1.1.3 Disorders of Consciousness (DOCs)

At this point, the reader may observe that the easiest way to understand if someone is conscious or not is by directly asking him/herself. This is also what each one of us instinctively does when a person passes out: asking *"Are you there? Can you hear me?"* and wait for the reply or at least a movement of approval in case the person is unable to speak. Some may argue, however, that this may also be misleading. Just think of all the artificial intelligences out there that can "talk" to us and help us in everyday life like they are perfectly conscious of themselves and the world. When you ask *"Alexa, are you there?"*, Alexa replies, tells you that she is existing and hearing you and ready for your orders. And yet, how can we define as "conscious" a non-living human being? This is also the point behind the famous **Turing Test** to identify if a machine is able to "think", one of the biggest theories in computer technology, and the **Zombie Theory** developed by the philosopher and cognitive scientist David John Chalmers. As interesting as both these concepts are for the debate behind the conception of "consciousness" and "free will", they will not be discussed further in this work, which will focus purely on the state of consciousness in human beings.

Even in human beings, yet, it is rather easy to say when a person is perfectly conscious and able to move and communicate with the external

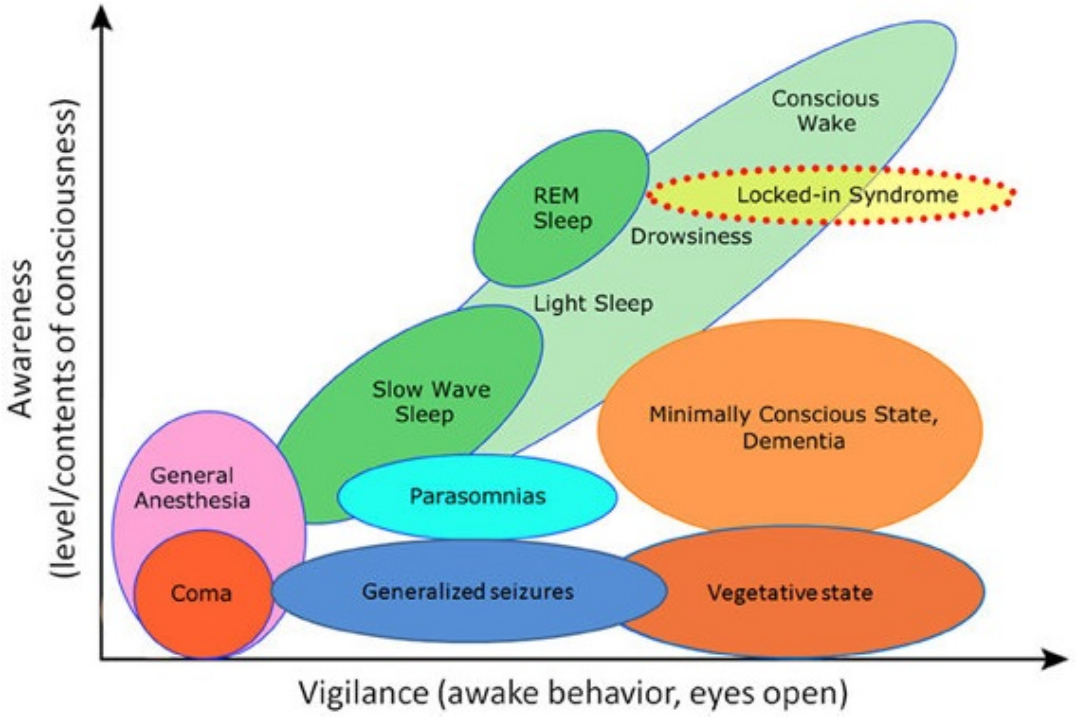

FIGURE 1.1: Awareness and Wakefulness in different states of consciousness.
Notice that it is possible to experience wakefulness in absence of awareness, but the opposite never happens. (Adapted from LAUREYS)

world, but a wide range of conditions and undefined states may arise in case a person is laying in a bed unable to speak, for example after an accident, and his movements cannot clearly be identified as perfectly voluntary. To go on in our analysis, it is necessary to try and make some order in the extensive spctrum of **Disorders of Consciousness (DOCs)**, comprehending all those states in which a patient is not able to communicate his or her conscious state:

- **Coma**: It corresponds to a state of profound sleep and unresposiveness in which the patient does not wake, and it is due to a suffering in the brainstem, regulating basic functions. A patient usually exits the coma state by waking and regaining the normal wake-sleep cycle, but this does not necessary mean a recover of cosciousness.

- **Vegetative State (VS)**: It happens after coma, when a patient correctly regains the wake-sleep cycle and the thalamocortical functions. The Vegetative State was brightly described by Bryan Jennet and Fred Plum in 1972 as an *"unresponsive wakefulness"*: there is no reply to painful stimulus and no reply to simple commands, but some non-intentional movements and automatic reflexes are observed.

- **Minimally Conscious States (MCSs)**: It is a wide range of not-perfectly defined states in which the patient is not able to communicate with the external world, but some signs of non-automatic movements can

be observed. Differently from VS patients, in the case of minimally conscious states it is possible to regain full consciousness also after many years.

- **Locked-in State (LIS)**: It happens when a patient after coma regains not only the correct wake-sleep cycle, but also full consciousness of him/herself and the surrounding world. Due to a neural damage, however, he/she is permanently paralyzed and is unable to communicate his/her condition, at least at the starting stage. Most of the patients in Locked-in State, however, eventually recover the ability to move the eyes, since these movements are not dependent on the brainstem, and they can communicate again with the external world.
- **Total Locked-in State**: It is a very rare condition of Locked-in State in which unfortunately the patient is also unable to move the eyes. Because of the total incommunicability of this state, it is practically impossible to correctly estimate how many people are under this condition.

This last definition of total LIS probably left the reader in a state of distress. Just imagine to be blocked into a body that you cannot move, perfectly capable of understanding the outside world but in no way being able to communicate this to anyone, without anyone noticing that you are there... Or again, think about wrongly considering a patient in a permanent vegetative state, completely unable to recover, while in truth being in a MCS with the possibility of regain consciousness under intensive rehabilitation. The whole point behind the constant study behind the concept of consciousness suddenly makes sense: understanding when a person is there and move more decisevly across the blurred lines of Minimally Conscious States.

1.1.4 Between Mind and Matter

Now that we have a better idea about what consciousness is, the problem that arises is to understand how an abstract concept like this can be linked to the phisical world. The separation between *"mind"* and *"matter"* has been at the center of the philosophical debate since Cartesio's *"res cogitans"* and *"res extensa"*, but a way to tie the two things together must be present. Without this connection, it would be impossible to translate a phisical state into a mental image (for example, the act of touching an object becomes the sensation of touching that object) or the opposite (thinking about touching the object and actually moving the body to touch it). In other words, there must be some aspects in the brain activity which explain the presence of consciousness and can be used as a starting point in our study.

1.2 The Human Brain

To find which aspects of the brain activity play a key role in the definition of consciousness, we need first to find our way into the human **Central Nervous System (CNS)**, and in particular the human brain. The CNS is an important

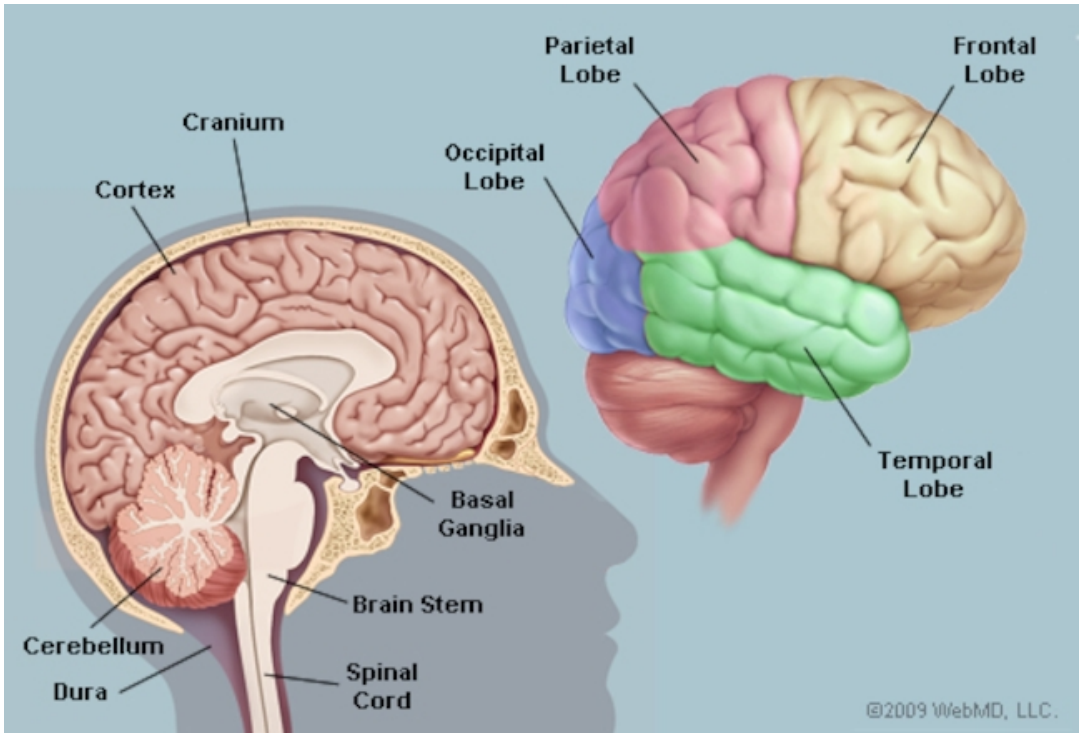

FIGURE 1.2: Anatomy of the human brain.

system taking care of a person's perception and personality, coordinating also the activity within the organs. It is structured into the **spinal cord**, spinal nerves allowing for the transmission of sensory and motor stimuli to the rest of the body, and the **brain**.

1.2.1 Anatomy of the Human Brain

Inside the brain we can identify three different entities. The **brainstem** plays an important role in the regulation of basic functions, such as the sleep cycle and cardiac and respiratory activities. The **cerebellum**, instead, controls the posture, coordinates movements and takes care of sensory stimuli and motor information. The largest portion of the human brain is the **cerebrum**, in which the **cerebral cortex** is the outer layer.

The cerebral cortex plays a key role in memory, attention, perception. thought and language. It can be divided in two hemispheres, left and right, and four regions or lobes:

- **Frontal Lobe**: The primary motor cortex, responsible of speech and personality;
- **Parietal Lobe**: In charge of the elaboration of stimuli, in particular for pain, tact and temperature, and language processing;
- **Temporal Lobe**: Responsible of hearing/smelling senses, visual memory, language comprehension and emotion association;
- **Occipital Lobe**: The visual cortex, in charge of the elaboration of visual stimuli, colour differentiation and motion perception.

1.2.2 The role of neurons in consciousness

The **thalamocortical system**, which corresponds to the thalamus and the cerebral cortex, and the cerebellum are both accomodating **neurons**, cells that receive, process and transmit information through electrical and chemical signals.

While the presence of neurons is a sign of high transmission of information, both for sensations and actions, it is not enough to assess the presence of consciousness. Indeed, the cerebellum has a very high number of neurons (about 80 billions), and yet from direct experience we can see that life and consciousness can go on even in its total absence. [23] Those patients for which it was necessary to completely remove the cerebellum, for example because of a tumor, struggle to speak fluently and coordinate movements, but they in no way change their conscious experience and they keep feeling sensations, perceive colours and shapes, hear sounds and savour tastes. The thalamocortical system, on the other hand, hosts a minor number of neurons (only 20 billions) but is the relevant part in the conscious experience. An important damage to this system can cancel from the subject a whole dimension of the conscious experience (for example, the visual perception of colours) or, in worst cases, leave an empty body without anyone inside.

What is the practical difference between the two entities is a central point in the defiition of "consciousness", and will be further discussed later in the text.

1.2.3 The Electroencephalogram (EEG)

One of the most essential techniques to study the brain activity is the **Electroencephalogram (EEG)**, invented by Hans Berger in 1929. The EEG permits to identify the electric potentials associated to those currents flowing perpendicularly to the scalp, thanks to some electrodes placed directly on the scalp and a conductive gel between each electrode and the skin in order to provide a stable electrical connection. [13] The electrodes, in particular, are usually placed according to a standard called the 10-20 International System of Electrodes Placement (IS 10-20), which is based on the area of the cerebral cortex underlying a certain electrode, shown in figure 1.3. Higher resolution systems based on 10-20 are possible, and in this case other electrodes are placed on intermediate areas.

Since the signals recorded are usually between only 10-100 μV, the amplitudes are very near to the electrical noise generated by the device. Moreover, most of the recordings suffer from some artifacts, which interfere with the useful signal, such as the eye and muscle movement or an electrode temporary detachment. For this reason, attention must be payed not only during the recording, to avoid any interference with other instruments, but also in filtering, artifact-removal and deleting any trial that is corrupted.

The EEG can identify a spontaneous activity of the brain, which is always present. The frequency spectrum is characterized by bands, called **rythms**, such as delta rythm (< 4 Hz) , found during dreamless sleep, theta rythm (4-7 Hz), found in sleep, meditation and hypnotic state, alpha (8-15 Hz), found

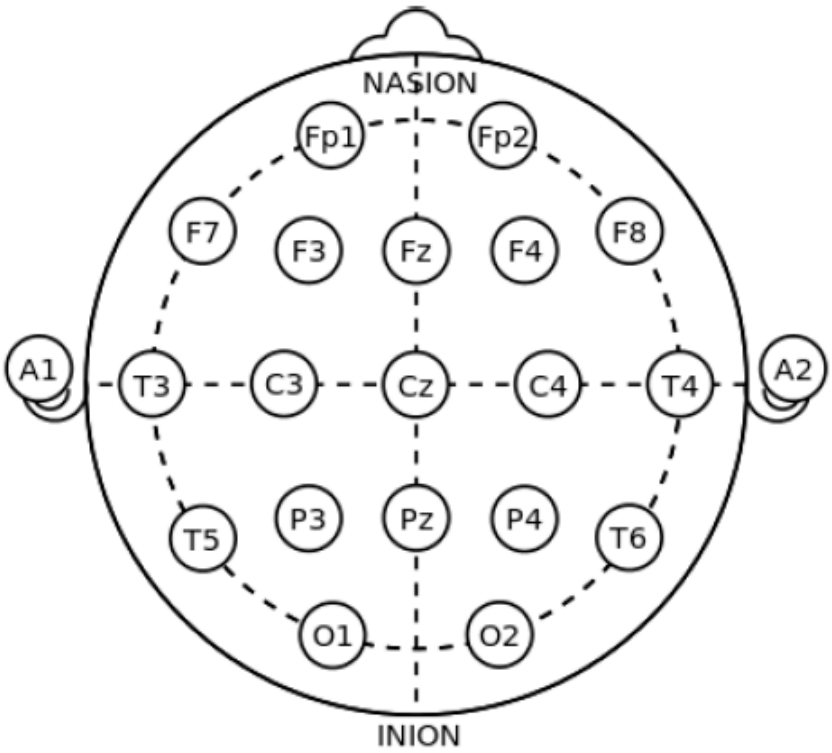

FIGURE 1.3: The 10-20 Electrode Location System. Letters correspond to the lobe or brain area: Pre-frontal (Fp), Frontal (F), Temporal (T), Occipital (O), Parietal (P), Central (C). Even numbers correspond to those electrodes placed on the right side, odd numbers to those on the left. "Z" electrodes are those placed along the midline.

in a very relaxed state of wake or in the moments immediately before falling asleep, Beta (16-30 Hz), in states of normal waking consciousness.

Chapter 2

The Road so far

Thanks to chapter 1, the importance of assigning the correct DOC to a patient is very clear. However, the only way that nowdays neurologists have to give a diagnosis is by using scales, such as the **JFK Coma Recovery Scale-Revised (CRS-R)**, developed by Giacino, Kalmar and Whyte in 2004 and consisting in 6 subscales for a total of 23 items. [6] In this cases, a standardized scoring is given on the basis of the patient's responses and a DOC state is assigned.

The problems arising from this type of assessment are mainly two:

- The scale assignment is based on the subjective perspective of the examiners, who can misinterpret or misread signals from the patient;
- The only aspects taken into consideration are the responses to stimuli.

But what happens if the analysts distort the patient's responses? Or again, if the patient is not able to communicate with the external world?

The absence of response is wrongly read by the examiner as absence of consciousness, and this is something that according to studies happens to more than 40% of the diagnosed vegetative state patients. This number needs to be reduced to zero, and what is needed to achieve this is a methodology based on objective parameters. This chapter is a brief excursus into some of the most interesting reaserches and studies conducted in the past years in this direction.

2.1 Readiness Potentials (Kornhuber and Deecke, 1964)

One of the most relevant studies in the field of brain activity was conducted in 1964 by Hans Helmut Kornhuber and Lüder Deecke, who discovered the socalled **Bereitschaftspotential (BP)**. [8]

2.1.1 Event Related Potentials (ERPs)

As we already mentioned, EEG is usually employed to analyze the spontaneous brain activity. However, it can also be used to identify the brain response to a specific sensory, cognitive or motor event, the **Event Related Potentials (ERPs)**. ERPs provide a direct and instantaneous measure of the

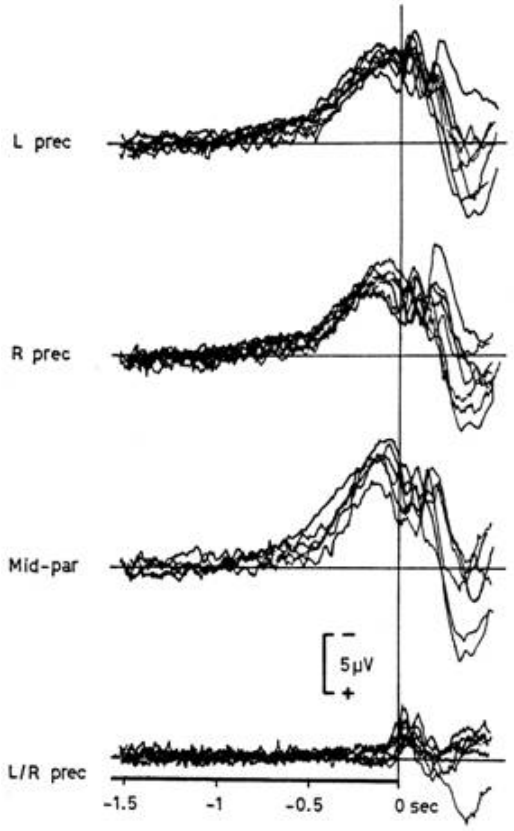

FIGURE 2.1: Readiness Potential (RP).
It is possible to identify the early component (Early BP) from -1 to -0.5 s and the late component (Late BP) from -0.5 to 0s. (KORNHUBER and DEECKE)

neural activity and they reflect the synchronous activity of a very high number of neurons, activating due to a stimulus. [12]

The best way to calculate ERPs is by averaging together recordings from many trials. This way, random brain activity is averaged out and the brain activity time-locked to the stimulus can be extracted.

2.1.2 Movements and Readiness Potentials

In 1964 Hans Helmut Kornhuber and Lüder Deecke by using the ERP technique discovered an activation process preceding of about a second a voluntary movement. This electrical activity, registered in the supplementary motor area, is called Bereitschaftspotential (BP) or **Readiness Potential (RP)** and it reflects the intention of the subject to perform a movement. It has an early component, called **Early BP** lasting from about 1 to 0.5 second before the movement, and a late component called **Late BP**, approximately from -0.5 to 0s. Before limb movements, the RP is initially equally distruted across the two hemispheres, but it starts to lateralize before the onset of the motion (**Lateralized Readiness-Potential (LRP)**).

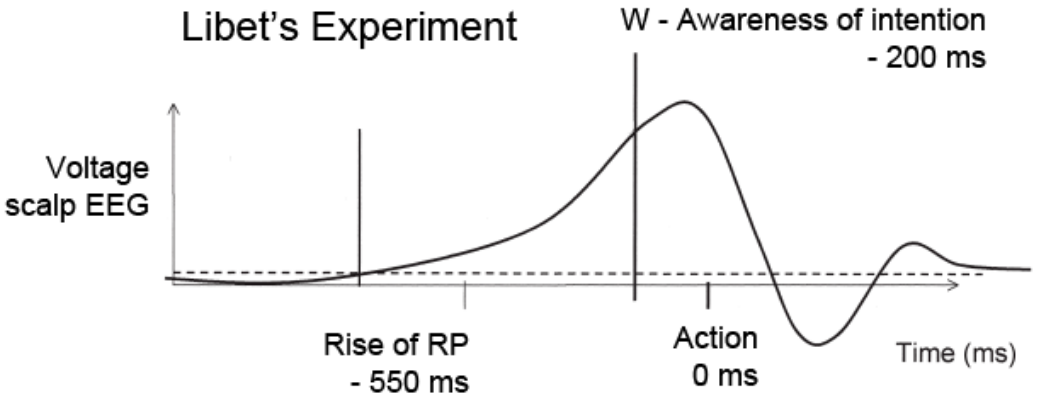

FIGURE 2.2: Libet's Experiment.
The Readiness Potential starts 350 ms before the subject is aware of his/her intention to move.

2.1.3 Libet's experiment

In 1980s, Benjamin Libet conducted a series of experiments in order to find the timing occurring between executing an action and being conscious of doing it. [10] [11] In his experiments, he was asking the subjects to perform a movement, such as pressing a button or flexing a finger: an electromyograph (EMG) was used to register the instant the action started, while an oscilloscope and an electroencephalograph were recording the neuronal activity. By averaging together 40 trials for each subject according to the ERP technique, it was shown that the RP starts about 350 ms before the subject is fully conscious of his/her decision to fulfill the motion, as shown in figure 2.2.

2.1.4 Contingent Negative Variance

The RP must not be mistaken for the **Contingent Negative Variation (CNV)**, a different event-related potential component shown in figure 2.3.

Firstly discovered by W. Grey Walter in 1964, the CNV appears in cases of warning-go tasks. [24] Two different stimuli are given: the first one is called *warning stimulus* and it simply "alerts" the subject that the second stimulus will arrive soon, the so called *imperative stimulus*, asking the subject to perform an action.

The CNV is a sustained negative component appearing between the warning stimulus and the imperative stimulus. The shape of the potential is similar to that of the Readiness Potential, but it shows a negative trend.

2.1.5 Readiness Potential and "Free will"

Libet's experiment set fire to the already lit debate about the concept of *"free will"*: if the brain prepares the action before the subject is even aware of the intention to move, can we really say that we "freely" decided to perform that action? This is, again, a discussion that will probably remain vivid for

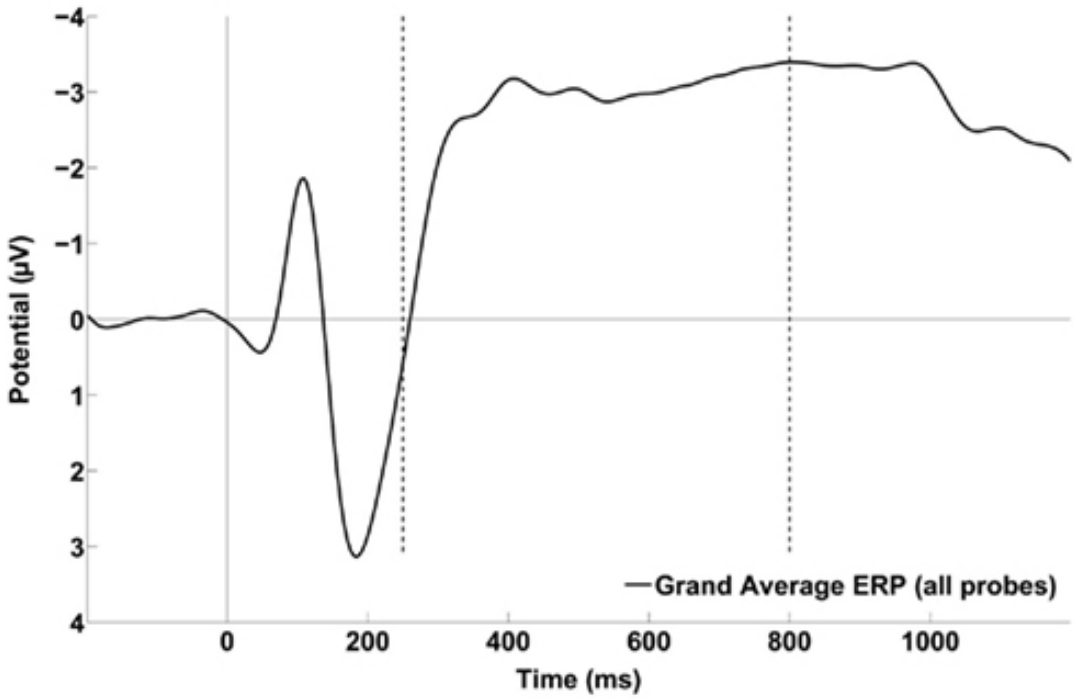

FIGURE 2.3: Contingent Negative Variation (CNV).
The ERP in the figure represents the Contingent Negative Variation timelocked to the stimulus onset. The dotted lines represent the warning stimulus and the imperative stimulus, among which the CNV is defined. Before the CNV, a short negative component and a short positive component also appear.
(NG, TOBIN and PENNEY [14])

centuries among philosophers and psychologist, and will probably never see a unanimous answer.

As lively as the debate may be about the philosophical implications of the RP, what is essential for us engineers are actually the practical implications: the RP is recorded only in case of a voluntary movement, and not is case of an involuntary one. Being able to distinguish between voluntary/involuntary actions may be a way to identify the presence of consciousness.

2.2 Information Integration Theory (Tononi, 2004)

We already mentioned before in this text that high information is not enough to assess the presence of consciousness. According to Giulio Tononi, this is due to the fact that information should also be integrated. [22]

2.2.1 Integrated Information

As we said in subsection 1.2.1, neurons are the main center of conscious experience, since they transmit information. However, despite neurons are present in a much higher number inside the cerebellum, this organ does not have any role in the state of consciousness of a subject, who could still feel and

perceive him/herself and the external world even after a cerebellectomy. The thalamocortical system, instead, has a minor number of neurons, but a lesion or injury to it leads the patient to "lose" a part of the conscious experience.

The difference among the two structures is that the cerebellum neurons are high in terms of information quantity, but basically are independent one another. In other words, the cerebellum acts as a "*zombie*": it receives and elaborates information, but it doesn't "feel" or "perceive" this information, and the reason behind this is still unknow.

The neurons inside the thalamocortical system, instead, provide for **integrated information**: this means that the repertoire of available states cannot be sub-divided into a repertoire of states available to independent components.

2.2.2 The Φ measure

Tononi proposes the $\Phi(S)$ function as a measure of the information integration of a certain subset S.

First of all, it is important to state some basic definitions from Claude Shannon's **Information Theory**. One of the key measures of information is **Entropy**, defined as the average uncertainty of a random variable A:

$$H(A) = -\sum_{a} P(a) \log_2 P(a) \tag{2.1}$$

where $P(a)$ is the probability mass function of the source A. Entropy is measured in bits and it corresponds to the number of bits required on average to describe the random variable. It is zero when A takes only one value with probability 1, or a maximum of $\log_2 1/K$ if $P(a)$ is equiprobable, with K equal to the total number of possible outputs.

The **Joint Entropy** of two random variables A and B with joint distribution $P(a, b)$ is defined as:

$$H(A, B) = -\sum_{a,b} P(a, b) \log_2 P(a, b) \tag{2.2}$$

The **Mutual Information** is the reduction of uncertainty due to another random variable, and it can be written as:

$$MI(A; B) = H(A) + H(B) - H(A, B) \tag{2.3}$$

Now that some basic notions about information theory have been defined, suppose to have a subset S of elements from a system, and suppose to divide it into two complementary parts A and B. Give maximum entropy to A and determine the entropy of the responses B which can be induced by A. The **Effective Information** between A and B is:

$$EI(A \rightarrow B) = MI(A^{Hmax}; B) \tag{2.4}$$

Notice that $EI(A \rightarrow B)$ is high if different outputs from A induce different firing patterns from B, while it will be low or even zero if the connections

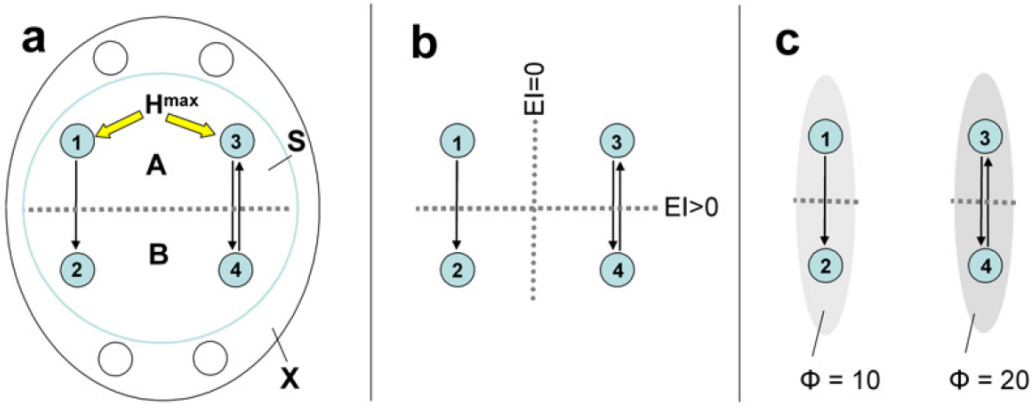

FIGURE 2.4: Effective Information, Minimum Information Bipartition, Complexes.
a) The subset S is made of four elements and is part of a larger system X. The subset S is bisected into two comlement A and B by a bipartition (the dotted line). Arrows indicate causal effect of A on B and B on A. Effective Information (EI) is calculated by injecting maximum entropy. b) The horizontal bipartition (black dotted line) yields positive effective information, while the vertical bipartition (grey dotted line) leads to EI = 0, and thus corresponds to the Minimum Information Bipartition (MIB). Notice that also other bipartitions are possible but not represented here, and they all lead to positive EI. c) The two subsets (1, 2) and (3,4) represented in figure with grey ovals are not part of any other subset with larger Φ, and thus constitute complexes. (TONONI)

between A and B produce scarce effects or if these effects are always the same. The effective information for both direction measures the repertoire of possible causal effects of A on B and of B on A:

$$EI(A \rightleftharpoons B) = EI(A \rightarrow B) + EI(B \rightarrow A) \tag{2.5}$$

The **Minimum Information Bipartition** of a certain subset S corresponds to the bipartition for which the normalized effective information reaches a minimum, in other words the *weakest link*:

$$^{MIB}A \rightleftharpoons B = \min \frac{EI(A \rightleftharpoons B)}{H^{max}(A \rightleftharpoons B)} \tag{2.6}$$

Finally, the **Information Integration** for the subset S is simply defined by the $\Phi(S)$ function, the non-normalized effective information for the minimum information bipartition:

$$\Phi(S) = EI(^{MIB}A \rightleftharpoons B) \tag{2.7}$$

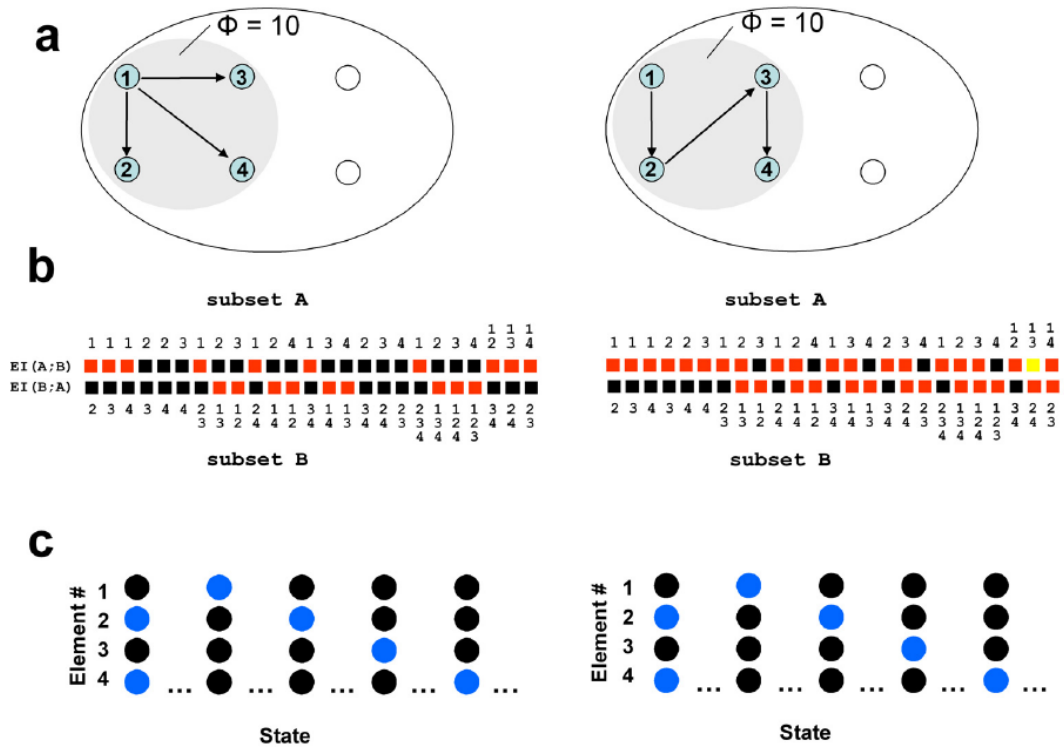

FIGURE 2.5: Two complexes with the same Φ.
a) The represented complexes (in the grey ovals) have the same value of Φ, but the connections among elements are very different: the left one has a "divergent" architecture, while the other is a "chain". b) The effective information matrices for both complexes are represented (black: zero, red: intermediate value, yellow: high). Notice that despite the same Φ value, the two matrices look different. c) The state diagrams of both complexes are represented (blue: active elements, black: inactive elements) for the four elements. The state diagrams look the same, but they have a different meaning, because the effective information matrices are different. (TONONI)

2.2.3 Complexes

A **complex** is a subset S having $\Phi(S) > 0$ and not included in any other larger subset having larger Φ. In particular, the **main complex** is the complex of a given system having the maximum $\Phi(S)$.

The relationship between effective information, minimum information bipartion and complexes is better explained in figure 2.4.

The **Effective Information Matrix** defines the relationship between the structural elements of a system: it contains the values of the effective information between each subset and any other subset, or in other words the informational relationships among the elements. Two complexes may have the same value of Φ (corresponding to the quantity of information), but very different effective information matrix (quality of information). The values of the variables mediating these interactions among elements of a complex specify the particular conscious experience at any time and are represented

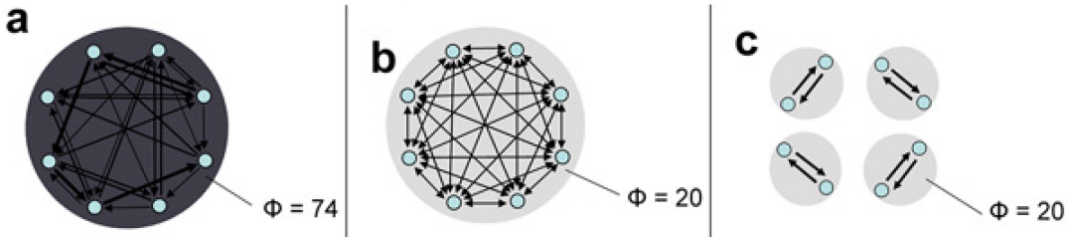

FIGURE 2.6: Information integration for a thalamocortical-like system.
a) The picture represents a system that is both specialized and functionally integrated, resembling the anatomical organization of the thalamocortical system. Each element is connected to a different subset of elements with different weights, in an heterogeneous arrangement. b) The same amount of connectivity but homogeneously distributed leads to a loss of specialization and thus a reduction of information integration. c) The same amount of connectivity but distributed in a way to form independent modules eliminates functional integration, and yields to separate complexes with lower Φ. (TONONI)

by **Activity States**.

In figure 2.5 there is an example of two complexes with the same Φ measure but different architectures, and their respective effective information matrix and state diagram.

2.2.4 Integrated information theory and thalamocortical system

As we already know, a lesion or ablasion of the thalamo-cortical system impairs the conscious experience of a subject. This would lead to thinking that consciousness lies in a distributed thalamocortical network, and not a single cortical area.

The thalamocortical system is made of elements which are functionally specialized. It is divided into systems, each of them dealing with a different function, and subdived into specialized areas. These specialized elements are then linked by a rich network of connections within an area and between areas. The organization of the thalamocortical system emphasizes at the same time specialization and integration, as shown in figure 2.6.

In terms of graph theory, a well-functioning system is a network providing different connection patterns for different elements (in other words, heterogeneous links) and all elements can be reached from any other element. There are two ways in which a thalamocortical-like system can lose information integration and thus reduce its Φ value:

- by replacing the heterogeneous link connections with homogeneous ones (reduction of specialization);
- by rearranging the system into smaller modules (reduction of integration).

The reason why the cerebellum is not a central part of the consciousness experience is that its organization is such that the cerebellar patches are activated independently from one another. This means that its connections rather generate many small complexes of low Φ (small integration).

Moreover, depending on certain conditions, the same thalamo-cortical system organization may lead to more or less consciousness. This is the case, for example, of what happens when a subject is under anesthesia or simply sleeping. Indeed, some key parameters may change the readiness of some neurons to respond to stimuli and thus reduce the Φ of the system.

2.2.5 For the future

The theory of integrated information seems to be an optimal way to measure the level of consciousness of a subject in almost every condition, also when unable to move or communicate with the external world, for example in locked-in state. However, at the time of publication of this Tononi's work, the validity of this theory still had to be experimentally proven. Further studies in this direction may be interesting to better understand which parameters take part in the information integration of the thalamocortical system and which conditions determine a certain quantity and quality of consciousness.

2.3 Decoding Algorithms and Information Theory (Quiroga and Panzeri, 2009)

For years, studies were conducted on the basis of a single-neuron multiple-trial approach. However, the brain typically makes decisions by evaluating the activity of a large neuronal population, and for this reason it would be best to shift to a **multiple-neuron single-trial framework**. As Rodrigo Quian Quiroga and Stefano Panzeri observed in 2009, this would be the optimal approach to obtain information about how neurons decode stimuli. [15]

2.3.1 The Multiple-neuron Single-trial approach

After recording the neuronal activity through implanted microwires, the first problem that arises in this approach is to distiguish the activity of single neurons from extracellular registrations. This can be obtained by using **Spike-sorting algorithms**, which cluster the spikes on the basis of their shape and through which it is possible to understand which spike belongs to which neuron, as shown in Figure 2.7.

Then, the information obtained from the neuronal population needs to be extracted by using objective measurements. An interpretation of the resulting patterns may be given by using two separate approaches, either by using **Decoding Algorithms** or by applying the **Information Theory**.

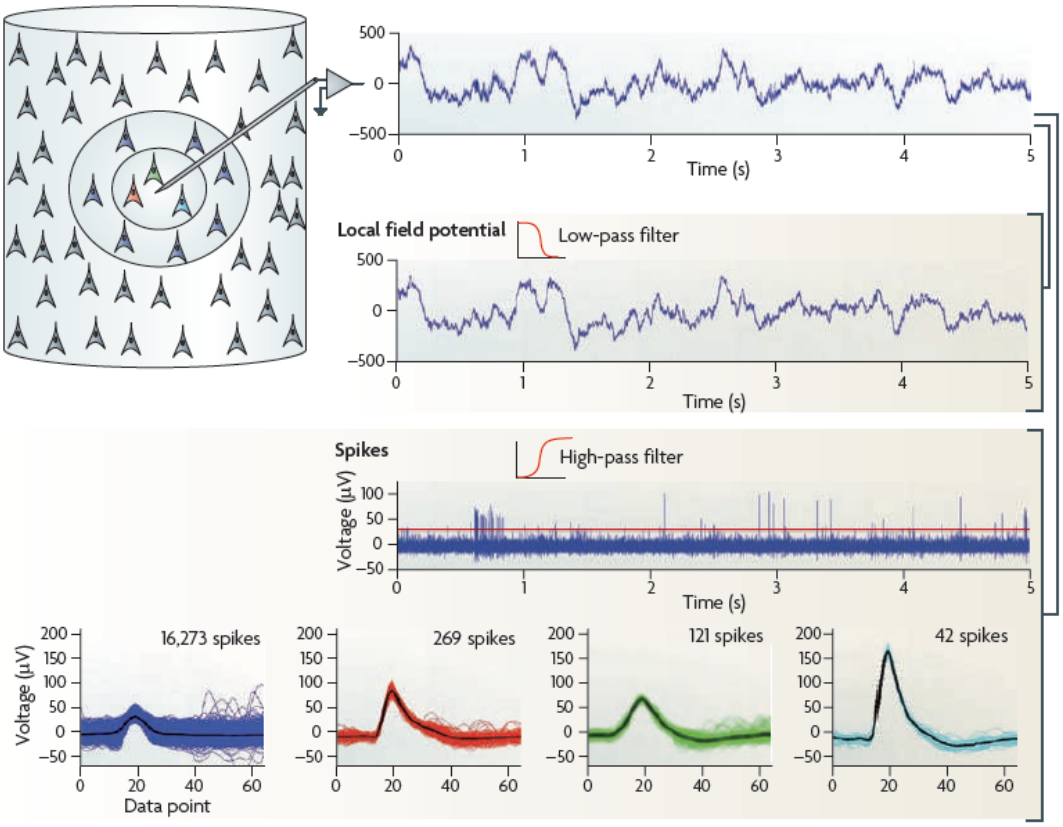

FIGURE 2.7: Spike-sorting algorithm.
A low-pass filter is used to identify the mean potential generated by the neurons near the electrode, called Local Field Potential (LFP). An high-pass filter, instead, is used detect the spikes through an amplitude threshold, in order to distinguish them from background activity. The spikes are then sorted according to the extracted features. (QUIROGA and PANZERI)

2.3.2 Decoding Algorithms

Decoding algorithms are able to predict which stimulus or behaviour elicits a particular neuronal response, as shown in Figure 2.8. They work through machine-learning: some trials are taken as training while the remaining are used as testing samples to validate the algorithm. Trials to be decoded are assigned to the nearest class, which is, in other words, the most "similar" to the given trial.

The results are then presented in the form of a **confusion matrix**, in which each element c_{ij} represents the normalized number of times the stimulus i was predicted as j. Of course, the wish is to have a diagonal matrix, corresponding to a perfect decoding. In case of equiprobable stimuli, instead, each element of the matrix will be equal to $1/K$, where K is the total number of stimuli.

2.3.3 Information theory

Information theory is a discipline studying the quantification, compression and transmission rate of the information, originally developed in 1948 by

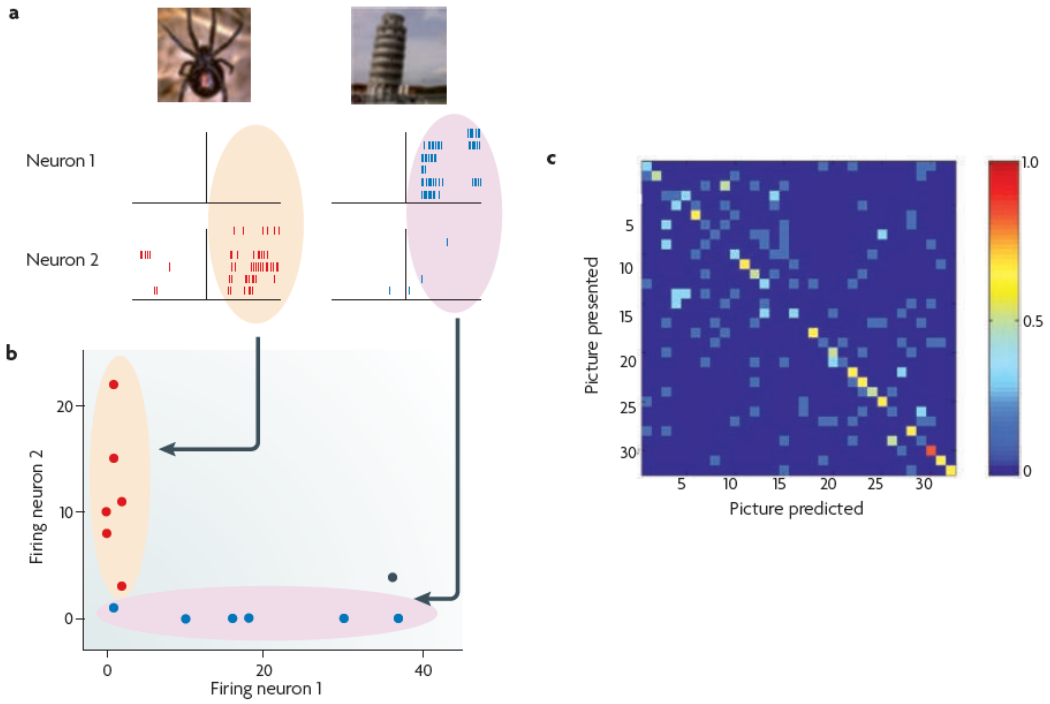

FIGURE 2.8: Decoding algorithm.
Suppose we want to predict which image elicited which single-cells response. a) For simplicity, only two pictures are here represented. b) Each point of the N-dimensional space represents a trial: trials in which the first picture was presented are red points, trials in which the second one was presented are blue points. The black point is a new trial, which will be assigned to the nearest neighbour (the second class). c) The outcome of the decoding algorithm is shown in the form of a confusion matrix: a perfect decoding means a diagonal matrix. (QUIROGA and PANZERI)

Claude Shannon. With information theory, it is possible to quantify the total knowledge of a stimulus gained through the known neuronal responses.

The **Shannon Entropy** of a stimulus with probability distribution $P(s)$ is defined as:

$$H(S) = -\sum_{s} P(s) \log_2 P(s) \tag{2.8}$$

It is measured in bits and it gives a measure of the average uncertainty about which stimulus is presented, or also the average amount of information gained for the stimulus presentation. Entropy is zero when the same stimulus is presented each time, or a maximum of $\log_2 1/K$ if the K stimulus are all equiprobable.

The residual uncertainty about a stimulus after observing the neuronal response is then given by the weighted average entropy of the posterior distribution, where $P(s|r)$ is the conditional probability:

$$H(S|R) = -\sum_{s,r} P(r) \cdot P(s|r) \log_2 P(s|r) \tag{2.9}$$

Notice that $H(S|R) \neq H(R|S)$ and $H(S) - H(S|R) = H(R) - H(R|S)$.

At this point, we can define the **mutual information** as the gained information or reduction of uncertainty about the obtained stimuli given the neuronal responses, where $P(s,r)$ is the joint probability:

$$MI(S;R) = \sum_{s,r} P(r)P(s|r) \log_2 \frac{P(s|r)}{P(s)} = \sum_{s,r} P(s,r) \log_2 \frac{P(s,r)}{P(s)P(r)} \tag{2.10}$$

Mutual information is zero if the stimuli and the responses are independent, while a perfect knowledge about the stimulus given the neuronal response gives a maximum of $MI(S;R) = H(S)$.

2.3.4 Merging the two approaches

As we got from the previous subsections, decoding algorithms and information theory are two very different approaches to the problem and they quantify different aspects of knowledge. Decoding algorithms are able to predict the stimulus that caused a certain response, but the amount of extracted information, even with an optimal decoder, can still be less than the information available in the neuronal responses. Information theory, on the other hand, is used to quantify the total knowledge of the presented stimulus.

What Quiroga and Panzeri suggest is to link the two approaches by computing the mutual information between the actual stimuli and the ones predicted by the decoder, $MI(S;S^P)$, or in other words between the rows and columns of the confusion matrix obtained with the decoding algorithm. In this way, we could be able to have a measure about how the neurons can report the predicted stimuli and the uncertainty of these predictions.

This idea has not been applied yet purely in the field of consciousness, but it may be interesting for further studies.

2.4 Approximate Entropy (Sarà, Pistoia et al., 2011)

According to Marco Sarà and Francesca Pistoia, consciousness depends on the interaction of wide neural networks, with a degree of unpredictability. For this reason, in 2011 with their research group they proposed to use a non-linear parameter such as **Approximate Entropy (ApEn)** to experimentally quantify the level of consciousness. [18]

2.4.1 Approximate Entropy calculation

Approximate Entropy is a way to quantify the unpredictability and the irregularity of a time-series. In this case, we will consider the EEG data, and the ApEn is defined as the logarithmic likelihood that patterns of data

close to each other will remain close for the next comparison within longer pattern. First of all, for each EEG dataset m-dimensional vectors sequences $p_m(i)$ are costructed, where i ranges from 1 to $N - m + 1$. At this point, the mean of the fraction of patterns of length m that resemble the pattern of the same length that begins at i is defined as:

$$C_i^m(d) = \frac{1}{N-m+1} \cdot D \tag{2.11}$$

where D is equal to the number of vectors for which the distance $|p_m(j) - p_m(i)| < d$, m specifies the pattern length and d the criterion of similarity. For their study, Sarà and Pistoia set $m = 2$ and $d = 20\%$ of the standard deviation of 400 EEG data.

The Approximate Entropy is then calculated as:

$$ApEn(N, m, d) = \frac{\sum_{i=1}^{N-m+1} \ln C_i^m(d)}{N-m+1} - \frac{\sum_{i=1}^{N-m} \ln C_i^{m+1}(d)}{N-m} \tag{2.12}$$

In other words, what Sarà and Pistoia did was to take 400 adjacent EEG datapoints and they divided them into 200 subsets of length $m = 2$. Then, they determined the number D of subsets with the criterion of similarity $d = 20\%$ of the standard deviation of the 400 points and calculated ApEn according to equation 2.12. This calculation was made for each EEG electrode: what could be taken into consideration for quantifying the level of consciousness is both the average ApEn among electrodes or also the minimum and maximum values.

What was expected in the study was that a reduction of ApEn was a sign of functional isolation of the source from the other sources, and so a sign of low consciousness.

2.4.2 Results

In their study, Sarà and Pistoia evaluated the ApEn for a set of healty subjects and for vegetative state patients for which they were assigned two different scales (the Extended Glasgow Outcomes Coma Scale and the Coma Recovery Scale-Revised).

As it is shown in figure 2.9, the mean ApEn can be used to discriminate healthy subjects from patients, but it is not enough to differentiate among the patients. However, after 6 months from the measurements, patients with the lowest mean ApEn either died or persisted in vegetative state, while those with the highest values became minimally conscious or showed partial/full recovery. The relationship between ApEn and both coma scales follows an exponential model and it is shown more in detail in figure 2.10.

The results obtained by Sarà and Pistoia in their reaserch prove that Approximate Entropy is a valuable way to quantify the level of consciousness and that non-linear parameters should be further studied to understand their role in the DOC classification.

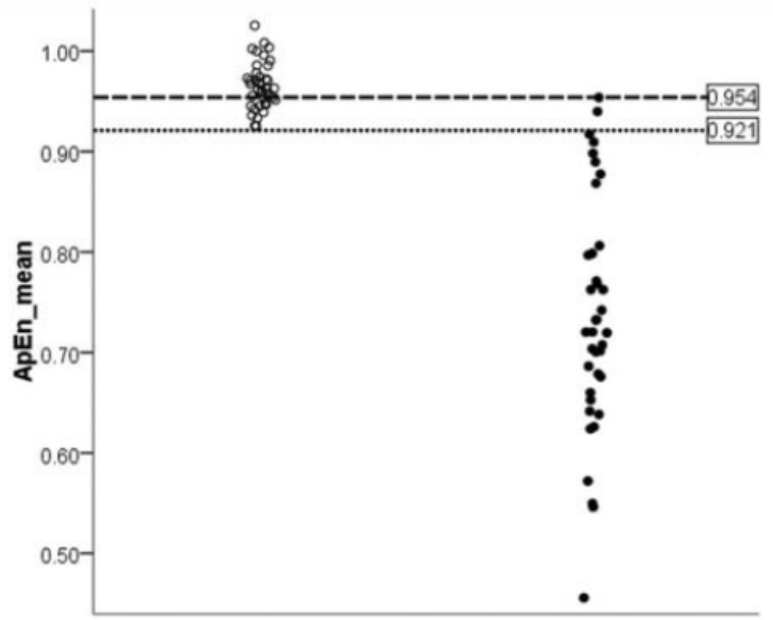

FIGURE 2.9: Mean Approximate Entropy for healty subjects and patients.
The open circles represent the mean ApEn for healthy subjects, while closed circles are the mean ApEn for patients. The dotted and slashed line respectively represent the two cutoffs respectively at 70% specificity 100% sensitivity and 100% specificity 94.7% sensitivity. (SARÀ, PISTOIA et al.)

2.5 Connectivity Biomarkers (Höller et al., 2014)

Consciousness can be examined both by asking the subject to follow commands or by using resting parameters.

Despite examining consciousness by using active paradigms would be the easiest to identify conscious patients, as Yvonne Höller observed in 2014 this can be efficient only for those subjects who are willing to cooperate. Indeed, a depression phase is a common experience among disabled people, and it is especially frequent among locked-in patients, who may voluntarily decide to not collaborate. For this reason, Höller and other reaserchers in their studies decided to focus on **connectivity biomarkers** extracted from resting EEG and use them to try to distinguish among different consciousness levels. [7]

2.5.1 Classification

In Höller's experiments, EEG were recorded in quiet conditions on both awake patients in MCS or VS, for which a stable CRS-R score was assigned, and on healthy subjects. The data was pre-processed and segmented into 2 seconds-long segments. For each of them, a list of features was evaluated and these features were averaged on the segments. In particular, the classification was performed as a two class-problem: healthy vs MCS, healthy vs VS, MCS

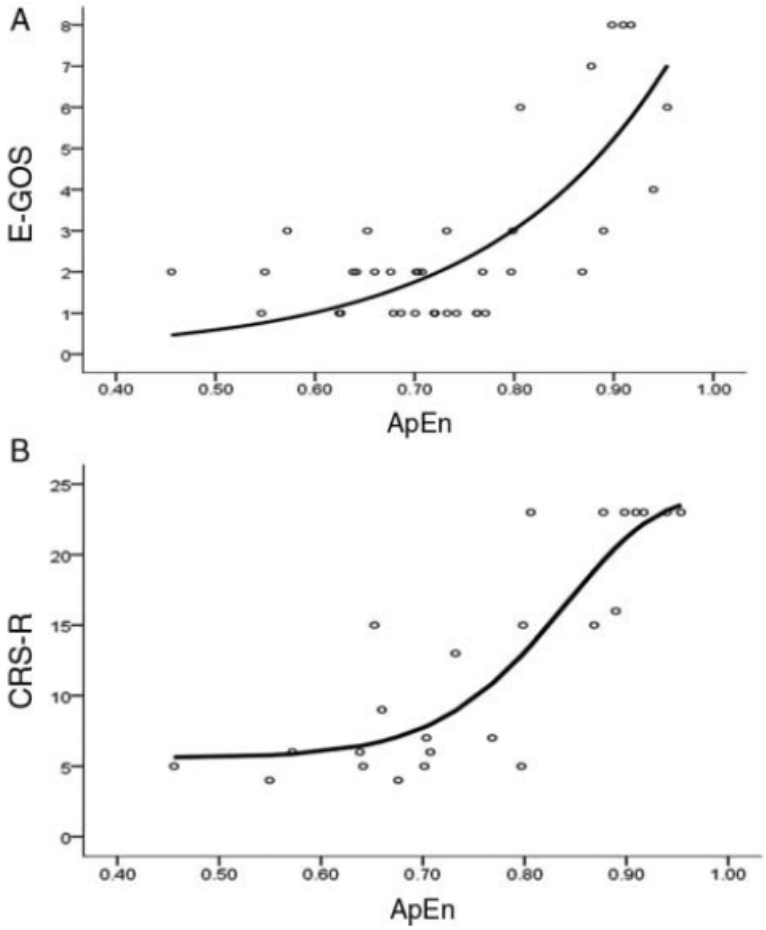

FIGURE 2.10: Relationship between Approximate Entropy and Coma Scales.
The coma scales are referred to the 6-months follow-up. a) Extended Glasgow Outcomes Coma scale (E-GOS) also includes the death cases and was evaluate for the overall patients group. b) Coma Recovery Scale-Revised (CRS-R) is better focused on consciousness transitions and was calculated only for those patients who remained alive. (SARÀ, PISTOIA et al.)

vs VS. **Support Vector Machines (SVM)** with a linear kernel function were used.

2.5.2 Results

In this thesis, we won't list and explain all the 44 features used for the reaserch, since not all of them were relevant and it is beyond the scope of our work, but the most important will be mentioned since they led to interesting results.

In fact, almost all the biomarkers used in the reaserch were able to differentiate between patients and healty subjects, but only a few could distinguish among patient groups. In particular, only three features related to the multivariate frequency response accurately classified among the three different levels of consciousness: **Partial Coherence**, **Directed Transfer**

Function (DTF) and **Generalized Partial Directed Coherence (GPDC)**. More precisely, the most accurate results were obtained for partial coherence.

These results proved that connectivity patterns on resting EEG can be used to accurately differentiate not only among healthy subjects and patients, but in some cases also among MCS and VS cases.

Chapter 3

An Integrated Information measure: the Perturbational Complexity Index

As we saw in section 2.2, the idea of Integrated Information Theory described by Tononi is to perturb the system and estimate the amount of integrated information by using the Φ function. However, what was missing in the study was a way to have this measure in an easy and objective way, for example by using EEG recordings.

For this reason, in 2013 Adenauer G. Casali and a group of researchers, among which Tononi himself, proposed to measure the integrated information from the thalamocortical system of a subject by using the **Perturbational Complexity Index (PCI)**.

3.1 PCI as Lempel-Ziv compression algorithm complexity

Casali defined the PCI as the algorithm complexity of the compression of the spatio-temporal pattern of cortical activations after a **Transcranic Magnetic Stimulation (TMS)**. In other words, from EEG recordings of TMS-evoked potentials, a matrix of significant sources is extracted and compressed by using the Lempel-Ziv algorithm: the PCI corresponds to the normalized algorithm complexity. [1]

The procedure for the PCI calculation is represented in figure 3.1 and its steps are better described in the following subsections.

3.1.1 Recordings of the brain response

As we said, in order to measure the amount of integrated information, the first step is to perturb the system. In their study, Casali decided to obtain this perturbation by direct activation thanks to Transcranic Magnetic Stimulation.

TMS is a non-invasive neurostimulation procedure, consisting in using electromagnetic induction to cause the flow of electrical currents into certain areas of the brain. In particular, for their study Casali oriented the maximum induced electric field perpendicularly to the cortical gyrus and the intensity was adjusted to obtain a significant response. The neural response to TMS

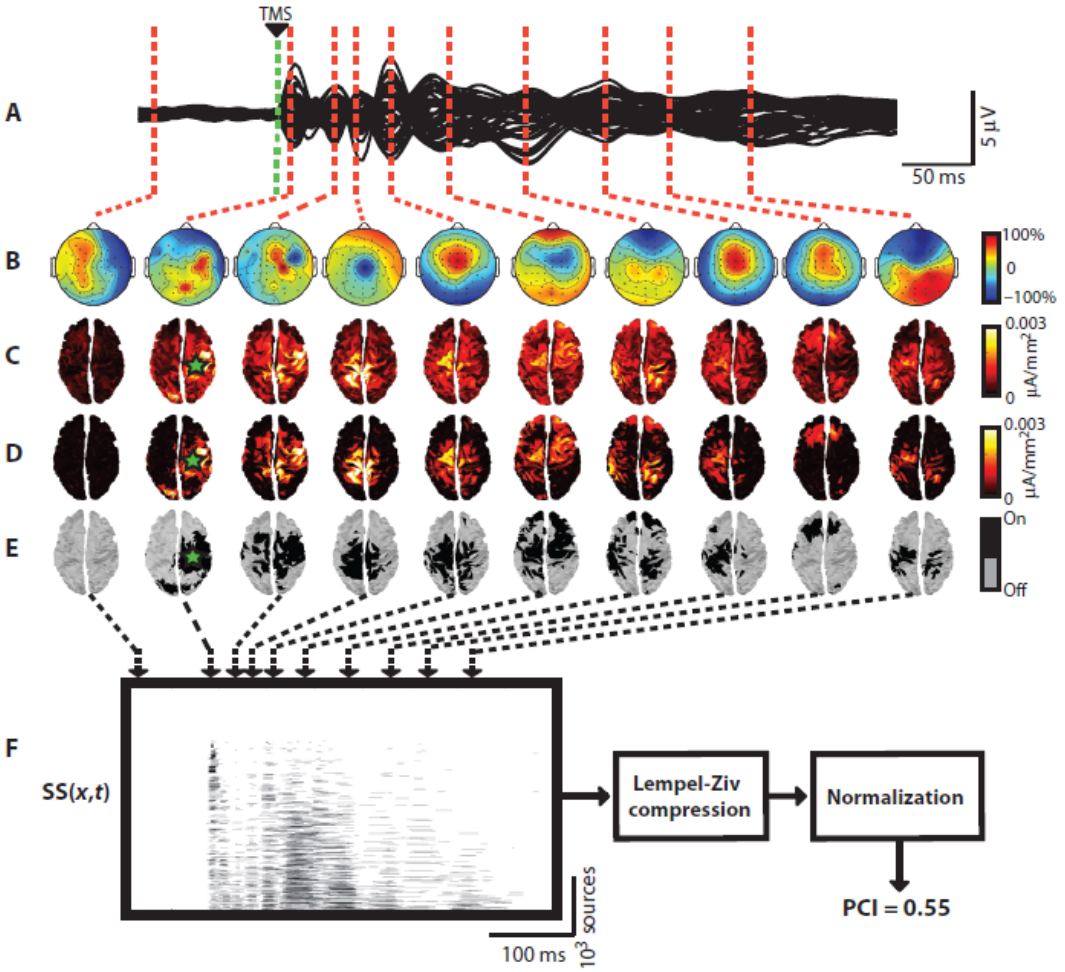

FIGURE 3.1: PCI calculation based on Lempel-Ziv algorithm complexity.
a) Averaged TMS-evoked potentials recorded from all EEG channels in one representative wakeful subject. b) Color-coded maps of the instanteneous voltages at selected latencies. c) The corresponding distributions of cortical currents. The green star is the location of the TMS stimulation. d) Significant TMS-evoked cortical currents. e) Binary distribution of significant sources: $SS(x,t) = 1$ if source x is significant at time t, 0 otherwise. The sources in $SS(x,t)$ are sorted on the basis of their total activity during post-stimulus period. f) The integrated information content is evaluated by measuring the Lempel-Ziv algorithm complexity. (CASALI et al.)

was recorded with an **High Definition EEG (hd-EEG)**, an improvement of the traditional EEG using more electrodes and sensors and providing higher accuracy.

More than 200 stimuli were acquired for each subject, and in offline data processing all trials containing artifacts and channels with bad signal to noise ratio were rejected.

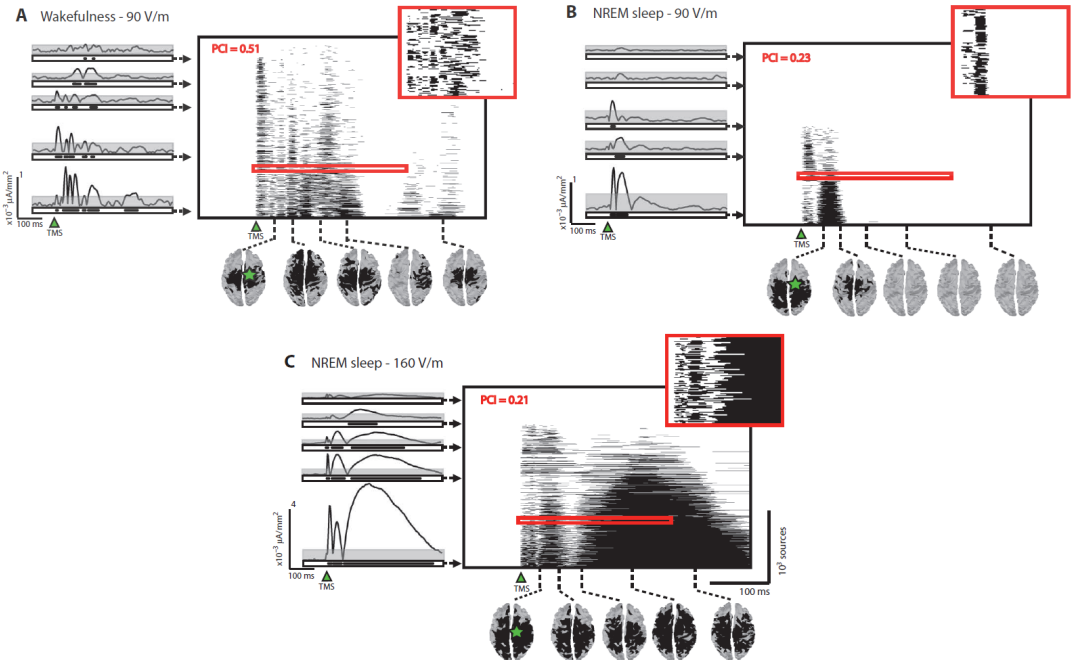

FIGURE 3.2: PCI and loss of integration/differentiation. On the left side, TMS-evoked currents for some representative sources, where the grey area is the statistical threshold applied to identify significant sources. On the right side, spatio-temporal matrix of significant sources $SS(x,t)$ for representative healthy subjects. The red frames show an extended set of the matrix. The green star identifies the location where TMS was applied. a) $SS(x,t)$ for alert wakefulness and 90 V/m TMS intensity, resulting in $PCI = 0.51$. b) $SS(x,t)$ for NREM sleep and 90 V/m TMS intensity, resulting in $PCI = 0.23$. Notice that the low value of PCI is due to the low number of active sources (low integration). c) $SS(x,t)$ for NREM sleep and 160 V/m TMS intensity, resulting in $PCI = 0.21$. Notice that the low value of PCI is due to the high redundancy in the $SS(x,t)$ matrix (low differentiation). Also, changing the intensity of the TMS does not define a drastic change in the PCI value. (CASALI et al.)

3.1.2 Source Modelling

Source modelling of the EEG activity was performed by modelling conductive head volume according to the 3-spheres BERG method, which allows to approximate the scalp potential in a 4-shell spherical head model. Then, each source's activity was centralized on the mean and normalized on the standard deviation of the baseline level. Surrogate average baseline was obtained by shuffling pre-stimulus samples at the single-trial level and the maximum absolute value of the surrogates across all sources was calculated. This was repeated 500 times to obtain a distribution of 500 bootstraps and according to this a threshold T was extracted.

On the basis of the threshold T, the **Significant Sources matrix** $SS(x,t)$ was calculated by setting $SS(x,t) = 1$ if source x was significant at time t,

$SS(x,t) = 0$ otherwise. The matrix was computed for the immediate cortical response, which is 300 ms after the stimulus.

3.1.3 Lempel-Ziv algorithm

After the Significant Sources matrix is obtained, $SS(x,t)$ can be compressed, for example by using the **Lempel-Ziv data compression algorithm**. This algorithm works as a dictionary of symbols from an aphabet: the symbols are grouped into strings and each string is converted into a smaller code. The **Lempel-Ziv complexity** c_L is the number of times a new subsequence is present while scanning the sequence. In particular, the sequence is given by the subsequent columns of the $SS(x,t)$ matrix.

The **normalized Lempel-Ziv complexity** is defined as:

$$\overline{c_L} = c_L \frac{\log_2 L}{LH(L)} \tag{3.1}$$

where $L = L_1 \cdot L_2$ is the total number of spatiotemporal samples of the matrix (L_1 sources and L_2 timesamples) and $H(L) = -p_1 \log_2 p1 - (1 - p_1) \log_2(1 - p_1)$ is the source entropy, being p_1 the fraction of ones contained in the matrix.

The **Lempel-Ziv Perturbational Complexity Index (PCI^{LZ})** is finally defined as the normalized Lempel-Ziv complexity of the spatio-temporal patterns of the TMS-evoked cortical activations $SS(x,t)$. The result, as shown in figure 3.2, is that PCI^{LZ} results lower both if less sources are significant on the overall post-stimulus interval or if their behaviour is standardized (in other words, if the matris $SS(x,t)$ is very redundant).

3.1.4 Results

The algorithm defined above to calculate PCI^{LZ} was firstly used on healthy subjects in both wakefulness and in a state of unconsciousness due to sleep or anesthesia. As it is shown in figure 3.3, the **Significant Current Density (SCD)** or strenght of the response in the post-stimulus interval (300 ms) is similar in both cases of alert wakefulness and in unconsciousness, while the number of significant samples in the spatio-temporal response $SS(x,t)$ only defines the number of active sources over the time interval and does not give an idea about the level of differentiation of information. The values of PCI, instead, result significantly higher in alert wakefulness, when the cortical responses are both integrated and differentiated. In other words, PCI can be used to distinguish between states of consciousness and unconsciousness.

Interesting results were also obtained by applying the algorithm to calculate PCI^{LZ} in graded states of unconsciousness. As shown in figure 3.4, in an intermediate level of propofol sedation the PCI values result still lower than the ones obtained in wakefulness, but also higher than the ones in deep sedation. The same can be said about different sleep stages: as we could expect, the lowest PCI values were obtained during **Non-Rapid Eye Movement (NREM)** sleep, during which there is little eye-movement and

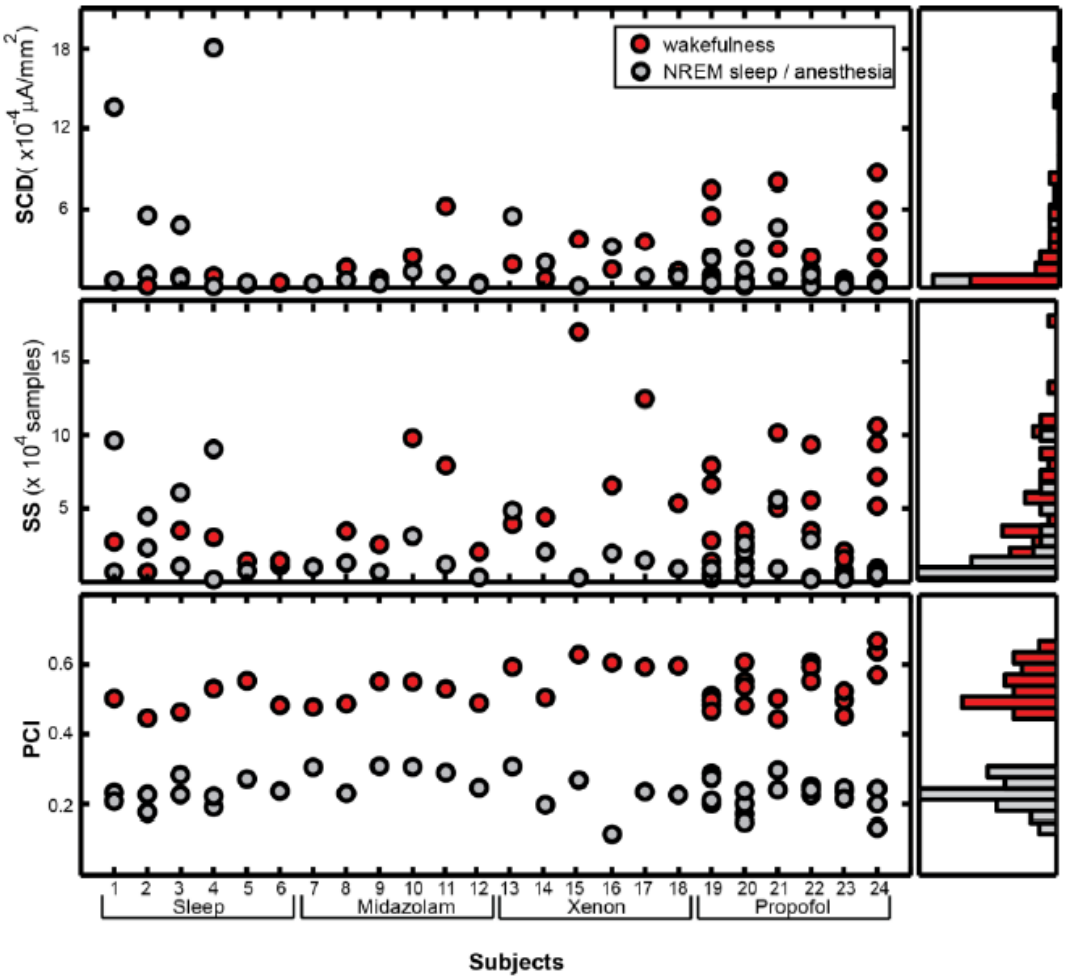

FIGURE 3.3: Parameters of cortical responses to TMS during wakefulness and sleep/anesthesia.
On the left side, the parameters for each subject in wakefulness and sleep/anesthesia, on the right the distributions of each parameter. a) Significant Current Density (SCD), or strength of the response, cumulative over the post-stimulus time. b) Number of total significant samples cumulative over the post-stimulus time. c) Perturbational Complexity Index (PCI). Notice that only PCI is a significant parameter to discriminate among wakefulness and sleep/anesthesia. (CASALI et al.)

usually no dreaming, while during the **Rapid Eye Movement (REM)** phase, in which we experience vivid dreams, PCI values are considerably similar to those in wakefulness. Consequently, PCI is also sensible to graded changes of consciousness, and this is a good premise to distinguish among different states of DOC.

This is proven by the results obtained when PCI^{LZ} was applied to severely brain-injured patients. As shown in figure 3.5, those patients with a stable diagnosis of vegetative state or unresponsive wakefulness syndrome (UWS) follow the behaviour of unconscious healthy patients, while those patients diagnosed with Locked-in Syndrome result in a PCI comparable with the ones in wakefulness. Those subjects with an intermediate level of consciousness,

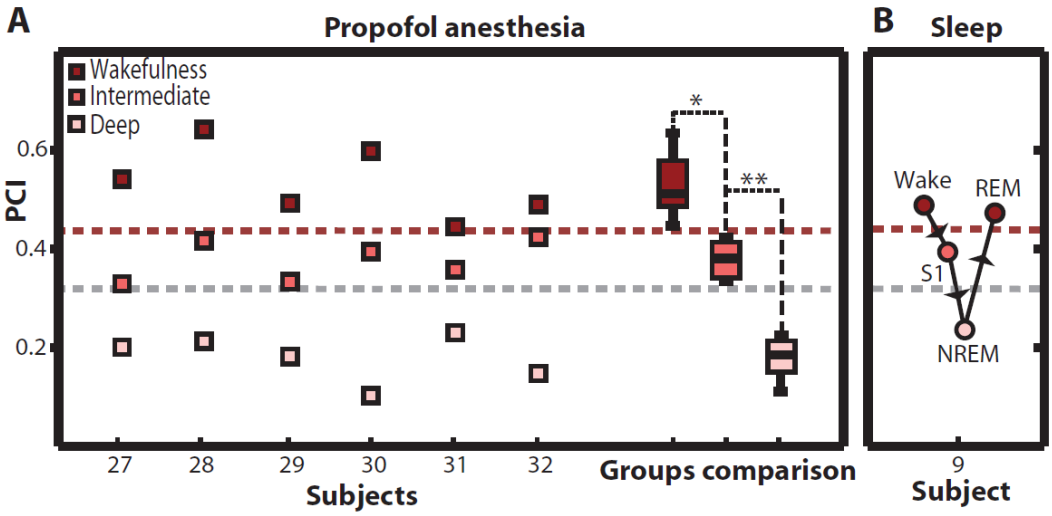

FIGURE 3.4: PCI and graded changes in consciousness. The grey dashed line is the maximum observed complexity during unconsciousness ($PCI = 0.31$), the red dashed line is the observed minimum complexity in alert wakefulness ($PCI = 0.44$). a) PCI in six subjects in wakefulness, under intermediate and in deep levels of propofol ansthesia. b) PCI in one subject in wakefulness, sleep stage 1 (S1), NREM and REM sleep. (CASALI et al.)

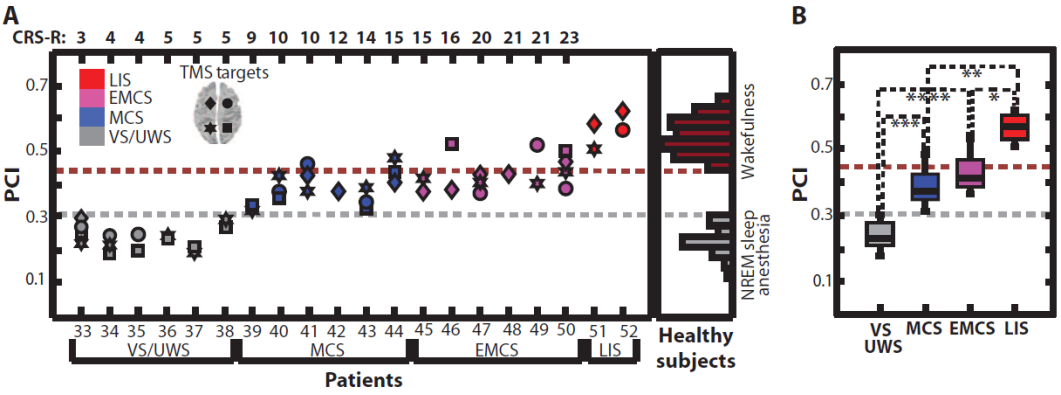

FIGURE 3.5: PCI in brain injured patients. The grey dashed line is the maximum observed complexity during unconsciousness ($PCI = 0.31$), the red dashed line is the observed minimum complexity in alert wakefulness ($PCI = 0.44$). a) On the left side, PCI values for 20 severely brain-injured patients for diffrent TMS targeting. On the right side, PCI distribution for healthy subjects in wakefulness and sleep or anesthesia. b) Box plots for PCI in brain-injured patients (CASALI et al.)

such as Minimally Conscious State (MCS) and Emerged Minimally Conscious State (EMCS), result in an intermediate PCI value. This proved that PCI evaluation can be used to identify the state of consciousness of a brain-damaged patient.

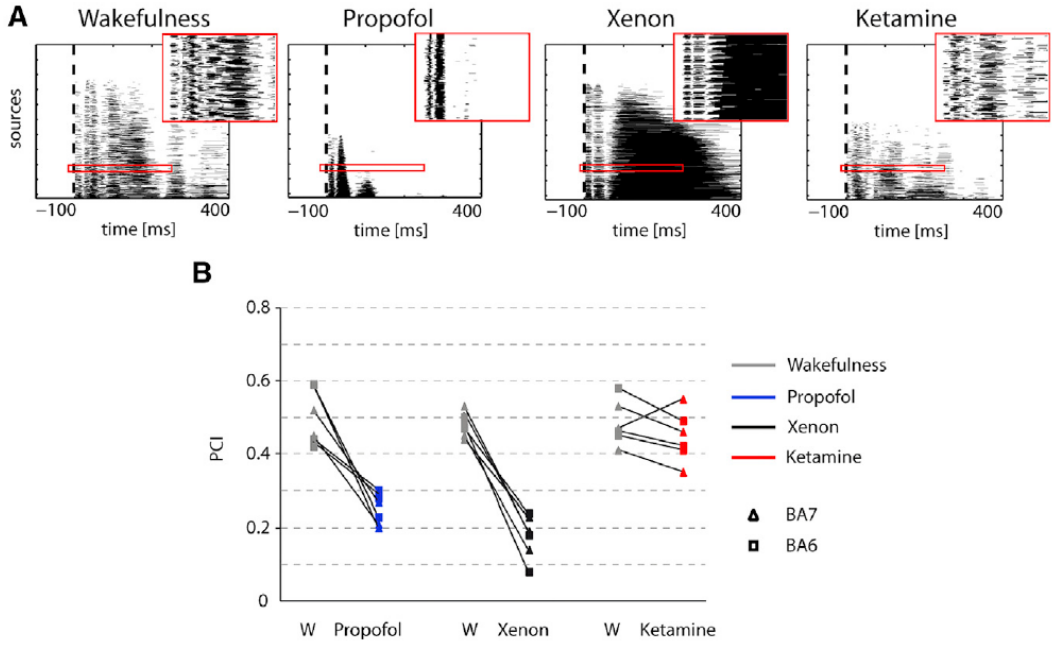

FIGURE 3.6: PCI in healthy patients under ansthesia. While propofol and xenon result in a loss of integration or differentiation, the behaviour of the cortical responses under ketamine anesthesia is very similar to the one in wakefulness. a) Spatio-temporal matrix of significant sources $SS(x,t)$ for representative healthy subjects in wakefulness and under propofol, xenon and ketamine ansthesia. The red frames show an extended set of the matrix. b) Individual PCI values for the three experiments in wakefulness and under anesthesia. The triangle and the square identify the different cortical targeting. (SARASSO, BOLY et al.)

More studies were conducted in 2015 by a research group lead by Simone Sarasso and Melanie Boly about PCI in healthy subjects under the effects of different anesthetics. [19] [20] As shown in figure 3.6, while the values of PCI for those who underwent xenon and propofol anesthesia are as expected significantly lower than those in wakefulness, the results obtained in ketamine-induced unresponsiveness are comparable to those obtained in a conscious state. This difference is explained by the retrospective reports of the subjects when waking up after the anesthesia: while those under propofol and xenon did not recall any conscious experience under the effects of the drug, subjects who underwent the ketamine experiments reported vivid, explicit dreams with a narrative structure, and some of them experienced hallucinations and distortions of the reality when awakening. This is particularly interesting because it shows that PCI is sensible to disconnected consciousness, also in those subjects who, based on their behavioural responses, are considered unconscious.

3.1.5 Limitations

As we have seen, the results obtained by estimating the PCI^{LZ} value in different consciousness conditions are very promising. However, this approach also shows some drawbacks and limitations.

The first problem resides in the algorithm itself. The PCI^{LZ} calculation requires a long and challenging processing of the TMS-evoked hd-EEG recordings. Indeed, obtaining the spatio-temporal matrix of the significant sources $SS(x,t)$ may be not so easy. The best would be to obtain a measure similar to the Perturbational Complexity Index that we described, but estimated directly at the EEG-sensors level.

Moreover, PCI^{LZ} is limited to TMS/hd-EEG evoked potentials: extending the algorithm to the usage of a traditional EEG recording system instead of the high density, or other types of recordings such as local field potentials, may lead to different conclusions.

3.2 PCI as State Transitions

Because of the limitations in the PCI^{LZ} approach explained earlier, a faster algorithm was proposed in 2018 by the same group researchers, led by Renzo Comolatti, and estimating the Perturbational Complexity Index directly from the EEG recordings of TMS-evoked potentials as the overall number of non-redundant **state transitions (ST)** caused by the perturbation. [3]

The proposed algorithm is shown more in detail in figure 3.7 and in the following subsections.

3.2.1 Recordings and Principal Component Analysis (PCA)

Details about the recordings of TMS/hd-EEG evoked potentials were already given for the PCI^{LZ} algorithm in 3.1.1.

Starting from the ERPs of these recordings, the first step is to perform **Principal Component Analysis (PCA)**, a method consisting in reducing the number of variables while limiting the information loss. In particular, being $\mathbf{X}$ the $N \times T$ matrix of the trial-averaged signals from $n = 1, \ldots, N$ different locations and $k = 1, \ldots, T$ time samples and $\mathbf{X}$ with $t = 0$ the instant of perturbation, the autocorrelation matrix of the response signal is calculated as:

$$\mathbf{A} = \mathbf{X}_{RES}(t) \times \mathbf{X}_{RES}(t)^T \tag{3.2}$$

where $\mathbf{X}_{RES}(t)$ is the response signal. The N_C **Principal Components (PCs)** are selected as the highest eigenvalues a_n of matrix $\mathbf{A}$ to account for at least 99% of the square mean field power of the response.

$\mathbf{X}(t)$ is then projected into the principal components, obtaining the $N \times T$ matrix $\mathbf{Y}(t)$.

3.2.2 Distance and Transition Matrices

At this point, each component $\mathbf{y}_n(t)$ is considered separately, and the **Distance Matrix** $\mathbf{D}_n$ is defined as the euclidean distance between timepoints:

$$\mathbf{D}_n(t_i, t_j) = \|\mathbf{y}_n(t_i) - \mathbf{y}_n(t_j)\| \tag{3.3}$$

This matrix is evaluated separately for baseline and reponse. Both are thresholded at different scales ϵ and for each ϵ a **Transition Matrix** $\mathbf{T}_n$ is calculated as the contour plot of the relative distance matrix, or in other words the non-recurrent states.

3.2.3 State Transitions

From the transition matrix for each ϵ, the **mean number of state transitions (NST)** across the T_c timesamples is calculated as:

$$NST_n(\epsilon) = \frac{1}{T_c^2} \sum_{i=1}^{T_c} \sum_{j=1}^{T_c} \mathbf{T}_n(\epsilon, t_i, t_j) \tag{3.4}$$

The temporal complexity of each component n is the maximized weighted difference at the optimal scale:

$$\Delta NST_n = T_R[NST_n^{res}(\epsilon_n*) - k \times NST_n^{base}(\epsilon_n*)] \tag{3.5}$$

where T_R is the number of response samples and $\epsilon_n* = \arg\max[NST_n^{res}(\epsilon) - k \times NST_n^{base}(\epsilon)]$ is the scale providing the maximum difference according to a weight parameter k (in Comolatti's work, it was set as $k = 1.2$).

Finally, the **State Transition Perturbational Complexity Index (PCI^{ST})** is defined as the sum of the optimal weighted differences across all the N_C components:

$$PCI^{ST} = \sum_{n=1}^{N_C} \Delta NST_n \tag{3.6}$$

Notice that PCI^{ST} can be read as the product between the number of principal components N_C and the average number of state transitions among all components $\Delta\overline{NST}$. N_C represents the spatial differentiation of the brain's response, since it is higher when different responses are given at different brain areas, while $\Delta\overline{NST}$ is the average temporal complexity of each component and is higher when for each component a high variety of patterns is evoked by the stimulation. High values of PCI^{ST} are obtained only when both parameters are high.

3.2.4 Results

In Comolatti's study, the method was firstly applied on 382 TMS Evoked Potentials from 108 healthy subjects both in alert wakefulness and in unconsciousness conditions due to NREM sleep or anesthesia. As shown

in figure 3.8, significant higher values of PCI^{ST} were obtained in a conscious condition and with a higher variability. The classification power was very similar to the one obtained by PCI^{LZ}, since the correlation between the two PCIs follows a linear model, but the biggest advantage was that the time required to perform the State Transitions algorithm took a much less inferior time to compute than the one with the Lempel-Ziv complexity.

The method was then applied also on 108 brain-injured patients, and for each of them different EEG settings were tried: the already used hd-EEG with 60 channels, the 10-20 system with 19 channels and a 10-20 system with only 8 channels. As shown in figure 3.9, the signs of consciousness in patients were detected with a very high accuracy while using the hd-EEG. The performance of state transitions PCI resulted just slightly inferior using 19 and 8 channels, proving that, on the contrary to the Lempel-Ziv complexity method, simpler EEG standards can be used having still good accuracy. Notice however that in UWS patients the absence of signs of consciousness cannot be considered as proof of absence of consciousness.

In conclusion, PCI^{ST} seems to be a method as reliable as the PCI^{LZ} in detecting signs of consciousness both in healthy and brain-injured patients, but with respect to the Lempel-Ziv complexity it provides a faster algorithm that can be applied without significant accuracy losses also in simpler EEG structures using an inferior number of channels.

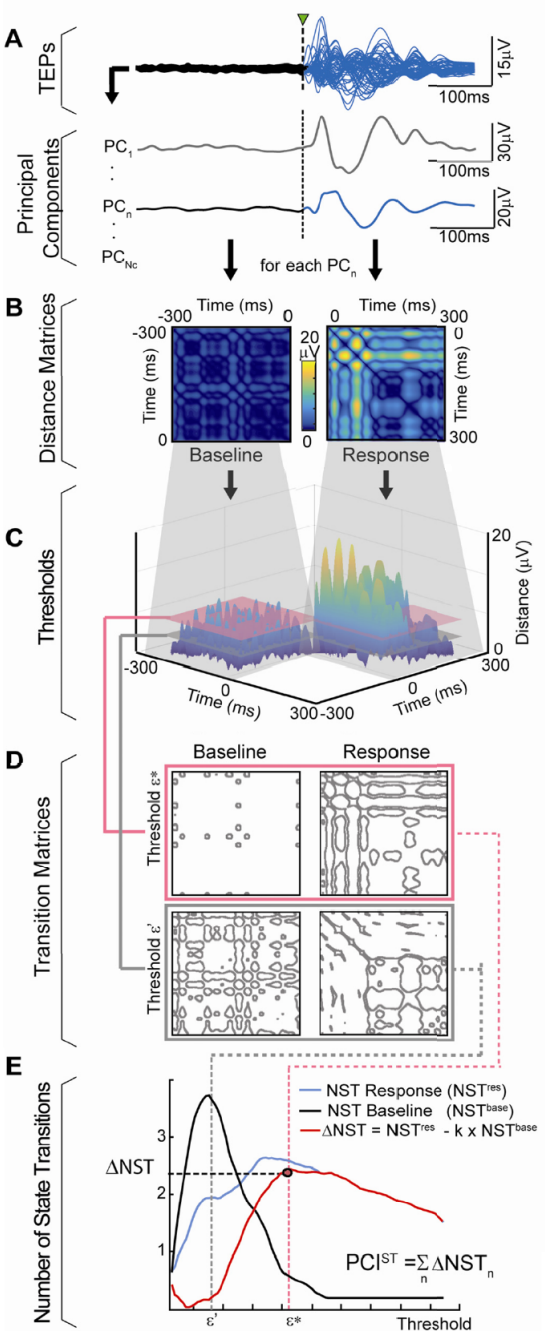

FIGURE 3.7: PCI calculation based on state transitions. a) From TMS/hd-EEG evoked potentials (TEPs), N_c principal components are extracted. b) For each single component PC_n, a distance matrix, based on the amplitude distance between timesamples, is extracted for both baseline and response time intervals. c) Distance matrices are thresholded at different amplitudes ϵ (in the figure: ϵ' and $\epsilon*$ are represented. d) For each ϵ, a transition matrix is extracted both for the baseline and the response. The number of state transitions for both of them is calculated (NST^{base} and NST^{res}). e) The complexity of each component is defined as the weighted difference between state transitions in baseline and in response (ΔNST_n). PCI^{ST} is finally obtained as the sum of the ΔNST_n for all NC principal components. (COMOLATTI et al.)

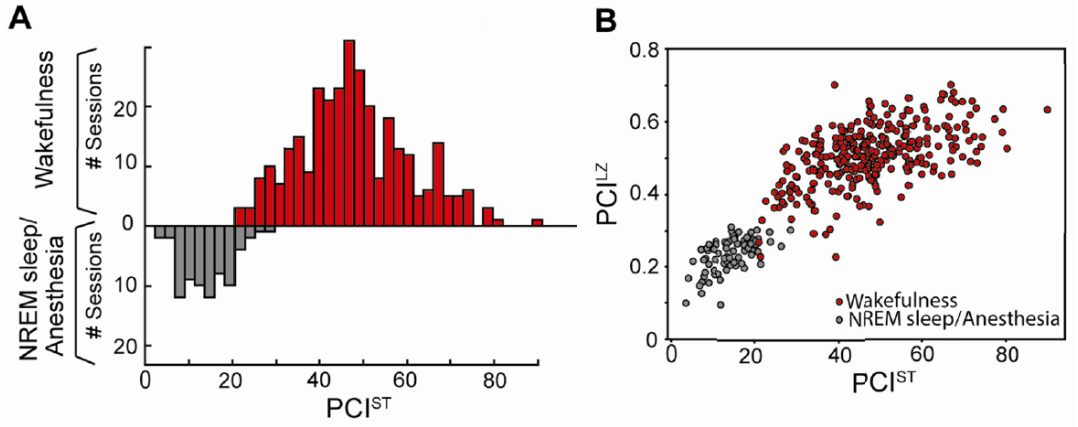

FIGURE 3.8: PCI^{ST} in healthy subjects in consciousness and unconsciousness.
a) Histogram of the values of PCI^{ST} in healthy subjects in alert wakefulness and under NREM sleep or anesthesia. b) Correlation between PCI^{ST} and PCI^{LZ}. (COMOLATTI et al.)

A

		LIS	EMCS	MCS	
hd-EEG	PCI^ST ≥	5 (100%)	11(100%)	45(91.8%)	61(93.8%)
	PCI^ST<	0	0	4	4
10-20	PCI^ST ≥	5 (100%)	11(100%)	44(89.8%)	60(92.3%)
	PCI^ST<	0	0	5	5
channels	PCI^ST ≥	5 (100%)	10(90.9%)	41(83.7%)	56(86.1%)
	PCI^ST<	0	1	8	9
		5	11	49	

B

		UWS		
		PCI^LZ ≥	PCI^LZ<	
hd-EEG	PCI^ST ≥	9 (100%)	4	13
	PCI^ST<	0	30(88.2%)	30(69.8%)
10-20	PCI^ST ≥	9 (100%)	6	15
	PCI^ST<	0	28(82.3%)	28(65.1%)
8 channels	PCI^ST ≥	8 (88.9%)	5	13
	PCI^ST<	1	29(85.3%)	30(69.8%)
		9	34	

FIGURE 3.9: States of consciousness in brain-injured patients using PCI^{ST} in different EEG setups.
The tables represents the number and percentages of patients classified as high and low complexity with respect to the classification cutoffs for different EEG setups (60 channels for hd-EEG, 19 for 10-20 system, 8 channels). a) In LIS, EMCS and MCS patients, PCI^{ST}'s sensitivity does not change much according to the EEG setups. b) Contingency tables for UWS patients in low and high complexity subgroups according to PCI^{ST} and PCI^{LZ}. (COMOLATTI et al.)

Chapter 4

Project Development

When visiting a rehabilitation center or the department of an hospital, it may happen to observe that a bed patient diagnosed with minimum consciousness or even in a vegatative state slightly moves a finger or a leg. This experience usually lights up hope inside the patient's relatives, who interpret this as a sign of consciousness and an attempt of the subject to communicate with the external world. However, things may be more complex than this, and behind this action there may be a variety of different explanaitions: there is a possibility that it truly is an effort to interact with the outside, but it could also be an involuntary response to a stimulus, or even a muscle spasm.

In this thesis, we propose to use the Perturbational Complexity Index to distinguish among voluntary and involuntary movements in a subject. In particular the approach was tested on healthy subjects and we chose to use the PCI^{ST} as defined in 3.2, since it is a fast and easy way to obtain the parameter that can be used without significant losses even in easier EEG systems. In the following chapter, we better describe the algorithm used to calculate the PCI and show the obtained results for voluntary, semivoluntary and involuntary movements. We also tried to apply the found results to different classifiers, and we will define the error rates of the outcomes.

4.1 EEG Recordings and ERP extraction

To apply the algorithm previously defined and calculate the PCI, the ERPs of each subject in response to the subject's movement are needed. This section defines the steps applied to obtain them.

4.1.1 Protocol definition

Three different experiments were conducted, and for each of them distinct recordings were taken.

The experiments were so defined:

1. **Voluntary Actions**: Subjects were requested to move a finger of their hand after hearing an acoustic signal. The stimulus were presented at random intervals and the subjects could autonomously decide on their own when to perform the action.

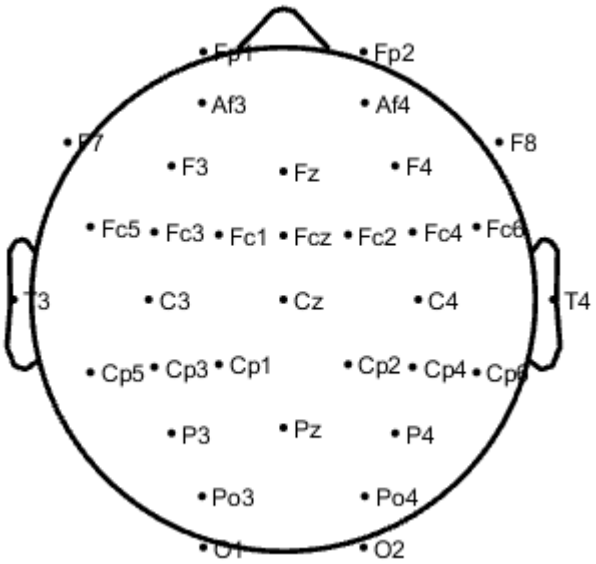

FIGURE 4.1: Channel locations of the EEG recordings.

2. **Semi-voluntary Actions**: Subjects were requested to perform the same action of the previous experiment, but immedately after hearing the acoustic signal. Again, the stimulus were presented at random intervals, but this time the subjects had to move as fast and as accurately as possible.

3. **Involuntary Reflexes**: The experimenter applied a percussion stimulus under the subjects' knee, in order to elicit the phasic patellar reflex and obtain a totally involuntary movement of the leg. This time, no will to perform the action was requested from the subject.

4.1.2 Materials

To acquire data, a Galileo Suite from EB Neuro with Brain Explorer amplifiers was used, and the Galileo Software was adopted to visualize traces.

A total of 34 EEG electrodes were applied on the scalp of each subject as shown in figure 4.1. In particular, the scalp was firstly cleaned with an abrasive cream called Nuprep, and then the electrodes were set on the head through a cap. Inside the electrodes, an electroconductive gel was put, in order to enhance the adherence to the scalp and improve the electrode impedence. The threshold of the electrodes impedence was set to be $10k\Omega$.

Electrodes were also placed on the subject's hand or leg, in order to record the movement for the EMG trace. More than 40 trials for each subject and each experiment were performed.

All recordings were taken on healthy subjects in Centro Puzzle, a cooperative center for brain injured patients in Turin.

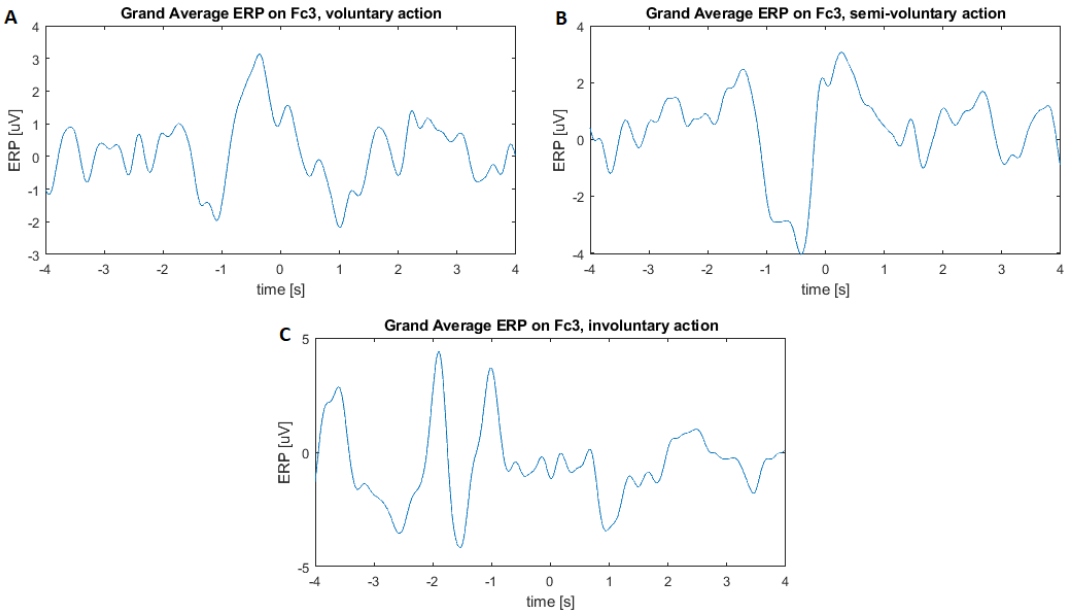

FIGURE 4.2: Grand Average of ERP on Fc3 for the experiments. a) For the voluntary actions, the grand average on Fc3 shows a behaviour similar to that of the Readiness Potential. b) For the semi-voluntary actions, the grande average on Fc3 shows a behaviour similar to that of the Contingent Negative Variation. c) For the involuntary reflexes, the grand average on Fc3 does ot a show a particular behaviour.

4.1.3 ERP extraction

After recording EEG and EMG traces for each subject and each experiment, they were analyzed by using **eeglab**, an interactive and opensource toolbox for Matlab, provided by the Swartz Center for Computational Neuroscience (SCCN). In particular, an eeglab program, already developed in 2018 by Simone Rumac for his thesis work, was used to extract the ERPs in response to the movement. [17]

At first, the onset instant of the action was detected from the EMG trace by using the *Teager-Kaiser Energy Operator (TKEO)*, as explained in Rumac's thesis. From the onset, the EEG data from all channels was divided into epochs, from 4 seconds before the movement to 4 seconds after the movement. Finally, the ERP averages for all eletrodes were extracted and exported into a text file, in order to use them for developing the wanted algorithm.

4.1.4 Chosen subjects

The measurements were performed on a high number of healthy subjects. Unfortunately, two problems arose:

- The first problem appeared while taking EEG measurements in the third experiment. In fact, every subject is different, and not all of them showed a visible and clear phasic patellar reflex. For this reason, it

wasn't possible to conduct the involuntary reflex experiment on all of the subjects.

- The second problem emerged during the signal processing analysis. Indeed, many measurements appeared to be badly taken or corrupted by an high electrode impedence and had to be discarded.

For this reason, a low number of subjects were still left for testing the hypothesis. In order to enlarge the dataset, we obtained some fictitious measurements by considering the valid ones and taking a lower number of trials for calculating the ERPs. Thanks to this expedient, we were able to obtain a total of 18 subjects: for all of them, the measurements for the first two experiments (voluntary and semi-voluntary actions) were present, and for half of them there were also the measurements for the involuntary reflex.

The validity of the experiments and chosen measurements are proven by the calculated Grand Averages on Fc3 channel shown in figure 4.2. While no particular behaviour is elicited for the involuntary reflexes, as expected the voluntary actions evoke a behaviour which appears as a Readiness Potential (RP), and the semi-voluntary actions bring out a potential similar to the Contingent Negative Variation (CNV) due to the imperative nature of the stimulus (both the RP and CNV were already mentioned in section 2.1).

4.2 PCI calculation

In this section, we will explain the applied algorithm more in detail and the principal differences implemented with respect to the procedure already defined in 3.2.

As an example, only figures from one subject (subject #3) and for experiments 1 and 3 (voluntary and involuntary actions) are reported. Similar considerations to the ones reported can be given for the other subjects and for experiment 2 (semi-voluntary actions).

4.2.1 Data cleaning

Before starting the algorithm, data needs to be cleaned.

First of all, we applied a low-pass filter to the imported ERPs, in order to have a smoother behaviour and better visualize the signals to work with. The filter we chose was the same for all channels, a **low-pass Butterworth IIR filter** of the third order with $-3dB$ normalized frequency at 0.05.

Moreover, as already mentioned, because of electrode detachment and high electrode impedence, in some experiments some channels resulted invalid to be used and had to be discarded. Using one channel less in some experiments didn't result in any particularly different behaviour. The channels to be discarded were selected by visual inspection.

Examples of ERPs after the data cleaning step are shown in figure 4.3. What could be already seen from this firs step was a prominent RP component in the first experiment for voluntary actions, while no particular behaviour

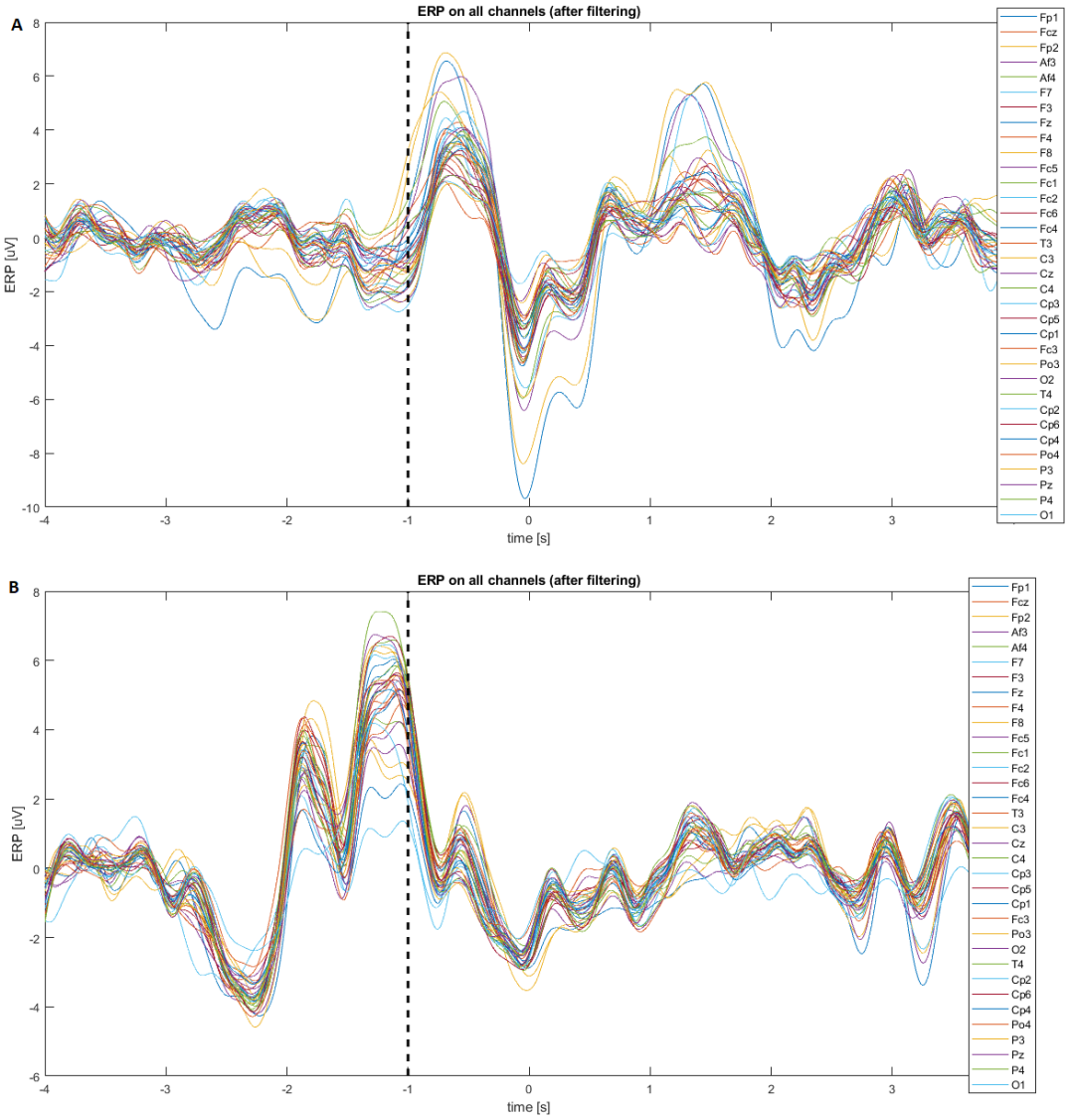

FIGURE 4.3: ERPs of subject 3 after filtering and data cleaning. The imported ERPs from all channels were low-pass filtered and invalid channels were deleted. In the figures, the black dotted line represents the onset of the stimulus (the intention to perform an action), set to one second before the onset of the movement. a) ERPs from experiment 1, voluntary actions. It's already possible to notice the RP component starting around the onset of the movement, or intention to move. b) ERPs from experiment 3, involuntary reflex. Neither the RP nor the CNV are visible, and the voltage behaviour seems to be similar for all electrodes.

was shown for the involuntary reflexes, which instead showed a sterotypical voltage for all the electrodes. In the second experiment, for semi-voluntary actions, most of the subjects showed a potential similar to the CNV, but in some cases there also appeared to be a stereotypical behaviour.

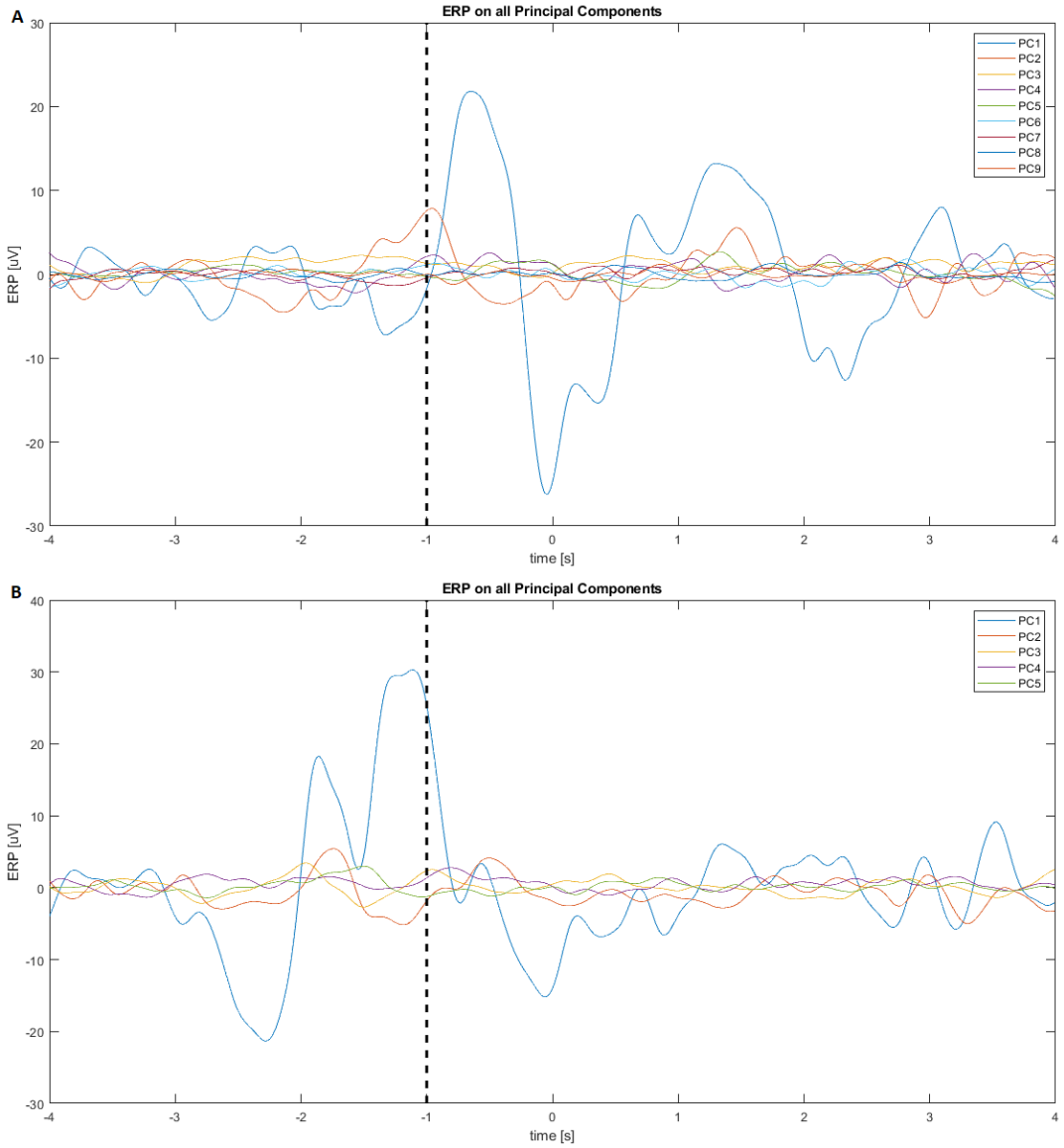

FIGURE 4.4: Principal Components of subject 3.
a) Principal components from experiment 1, voluntary actions. Notice that the first component follows the behaviour of the RP. b) Principal components from experiment 3, involuntary reflex. The number of extracted PC is lower than the previous case because the voltage across all the electrodes is more stereotypical.

4.2.2 Principal Component Analysis

The second step is the Principal Component Analysis (PCA).

Differently from what was described in 3.2.1, in which the weights of the PCA were extracted only from the reponses, in the algorithm we developed we considered the whole time interval. It was chosen to keep the eigenvalues providing for at least the 99% of the total square sum.

Applying the PCA means *linearly tranforming* the variables and keeping only those providing for the largest eigenvalues. What happens, as shown

in figure 4.4, is that for those measurements following a stereotypical trend in all electrodes, like in the involuntary reflex experiment, the number of principal components is strongly reduced. On the contrary, for the voluntary action experiment an higher number of PCs is provided, and interestingly the first component follows the RP behaviour. Again, for the second experiment (semi-voluntary actions) the situation in most cases appeared to be similar to that of the first experiment, with the first component following the trend of the CNV, but in some cases the number of principal components was smaller due to a stereotypical response.

4.2.3 Distance Matrix

After extracting the principal components, each of them is analyzed separately. In particular, for each of them a Distance Matrix is calculated for both the baseline and the response, in which each element $D(i,j)$ represents the absolute difference between the value of the component at timesample i and timesample j.

Instead of considering $t = 0$, corresponding to the onset of the movement, as the instant of response start, we considered $t = -1s$, which ideally represents the beginning of the RP, and so the *"intention"* of the movement is taken as the stimulus of the experiment. We also took off 1 second at the beginning and 1 second at the end of the considered time interval, in order to streamline the calculation process. The baseline interval was then defined as $t_{base} = [-3; -1]s$ and the response interval as $t_{res} = [-1; 3]s$.

An example of obtained Distance Matrix is shown in figure 4.5. For experiment 1, voluntary actions, the baseline Distance Matrix for the first component tends to have lower values with respect to the response Distance Matrix, having wide areas of high value elements. This difference between the two matrices tends to be less evident for the other components. The same happens for those measurements of the second experiment for semi-voluntary actions in which the CNV component is prominent and clearly visible. For those experiment 2 measurements that are less clear and the involuntary reflexes of experiment 3, instead, for all components the elements of the two matrices tend to have a similar range of values.

4.2.4 Transition Matrix

At this point, different thresholds ϵ are set: we can define a *"binary state"* as "below" or "over" the threshold. The Transition Matrix for each ϵ represents the transitions between one state and the other.

A simpler way to interpret the Transition Matrix, as already explained, is by considering the matrix of the states (below and over) and, as shown in figure 4.6, taking the contour plot. In general, the transition matrix of the response presents an higher number of transitions than the baseline.

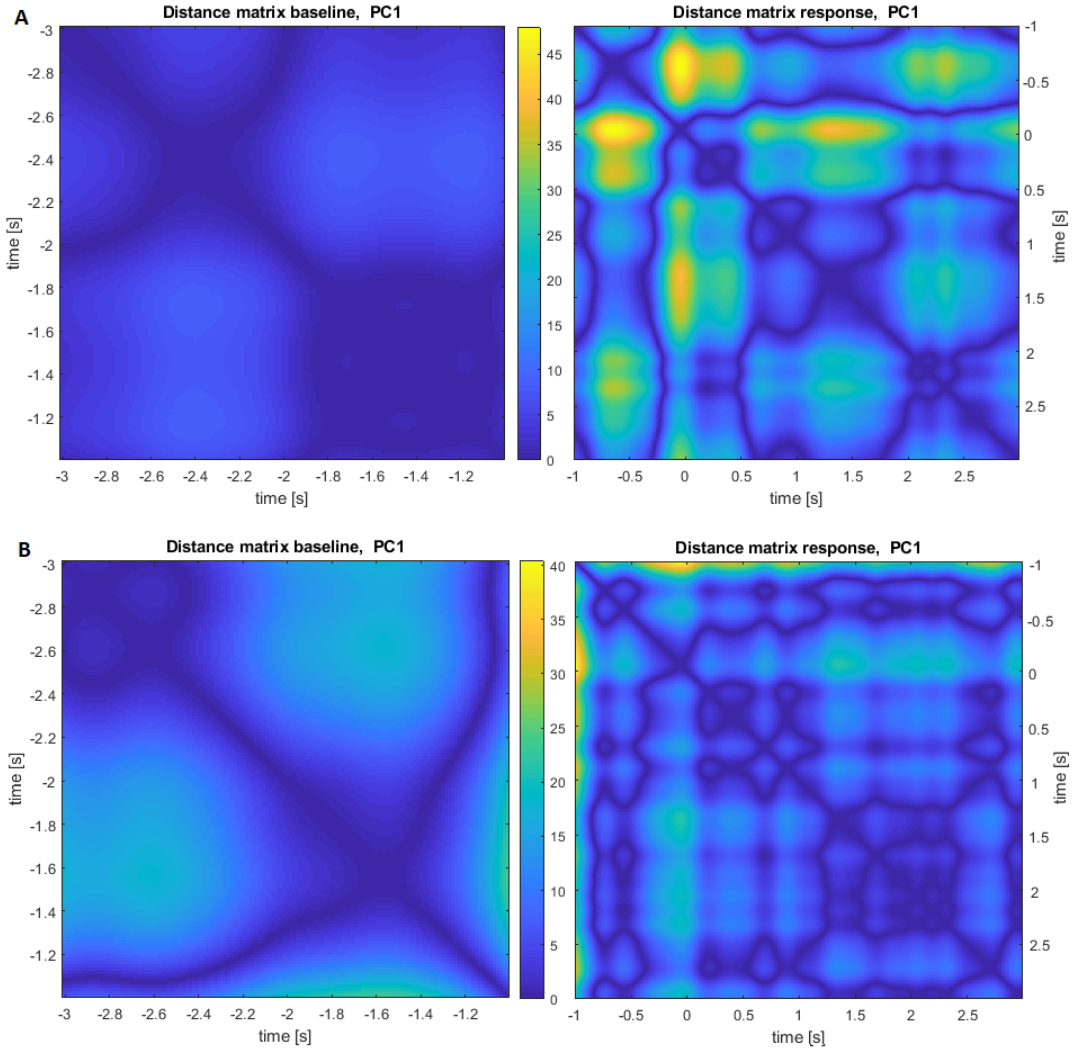

FIGURE 4.5: Distance Matrices of subject 3 for Principal Component 1.
On the left side of the figures, distance matrix of the baseline, on the right side the distance matrix of the response. a) Distance matrix for baseline and response from experiment 1, voluntary actions. Notice that the matrix of the baseline presents much smaller values, while the response one presents wide areas of high distance. b) Distance matrix for baseline and response from experiment 3, involuntary actions. Notice that baseline matrix and response matrix are more similar than in the previous experiment.

4.2.5 Number of State Transitions and PCI calculation

From the Transition Matrix for a given threshold ϵ, the number of state transitions can be simply calculated by summing up the non-zero elements of the matrix. In order to compare the NST from baseline and response the number needs to be normalized for the number of timesamples.

As shown in figure 4.7, the weighted difference between the NST from response and NST from baseline is calculated for each ϵ, and in particular we

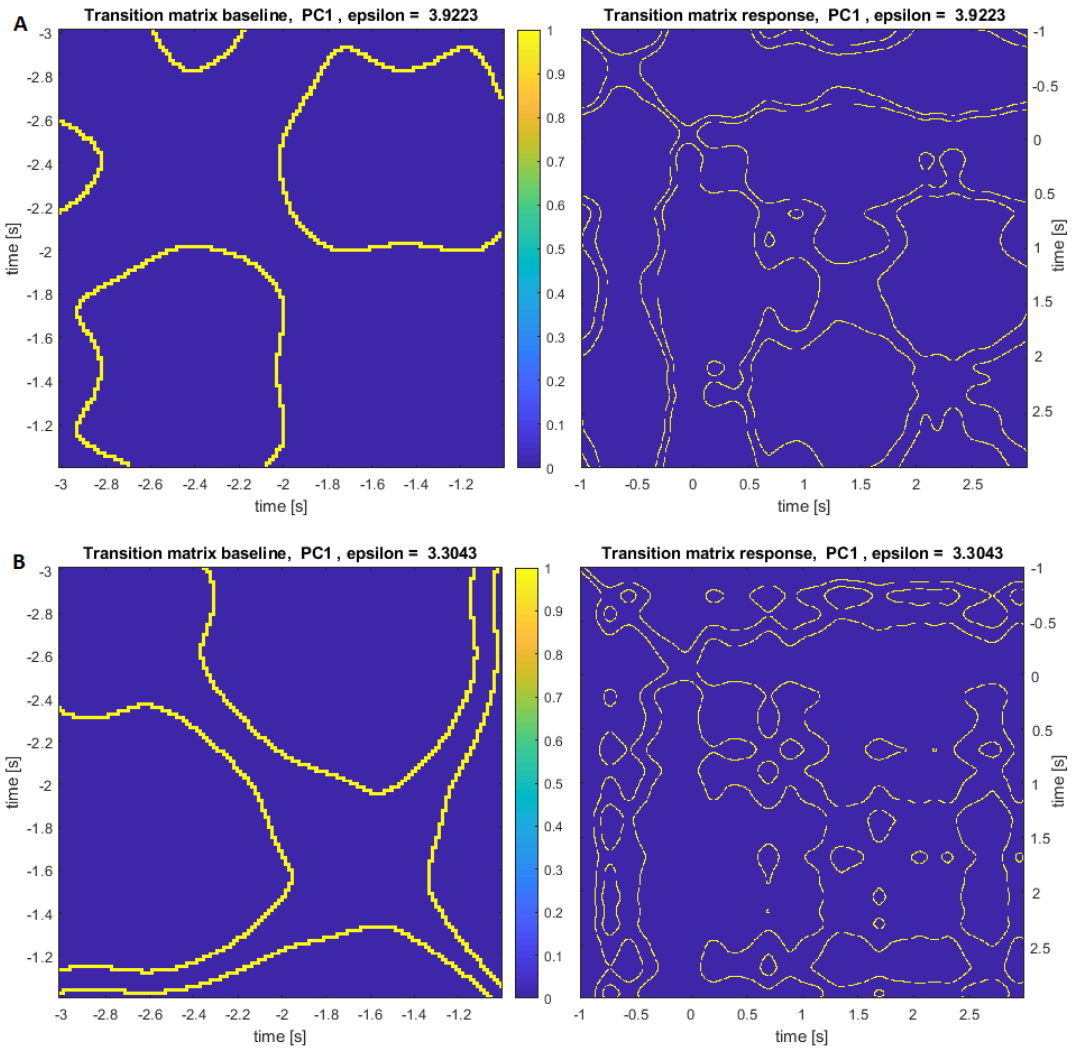

FIGURE 4.6: Transition Matrices of subject 3 for Principal Component 1.
On the left side of the figures, transition matrix of the baseline, on the right side the transition matrix of the response. For both shown experiments the number of transitions in the response matrix is higher than in the baseline matrix. a) Transition matrix for baseline and response from experiment 1, voluntary actions, and $\epsilon = 3.9223$. b) Transition matrix for baseline and response from experiment 3, involuntary reflex, and $\epsilon = 3.30430$.

decided to keep $k = 1.2$ as already chosen in Comolatti's experiment. The maximum NST difference tends to be higher for the voluntary actions rather than the involuntary movements.

Finally, the PCI for the given subject and experiment can be calculated with the expression 3.6, by summing the maximum ΔNST for each component and multiplied by the square number of timesamples. This means that the PCI value results higher if the number of extracted principal components is high and if for each of them the difference of state transitions is high (or, in other words, if the response to the stimulus is both integrated and differentiated).

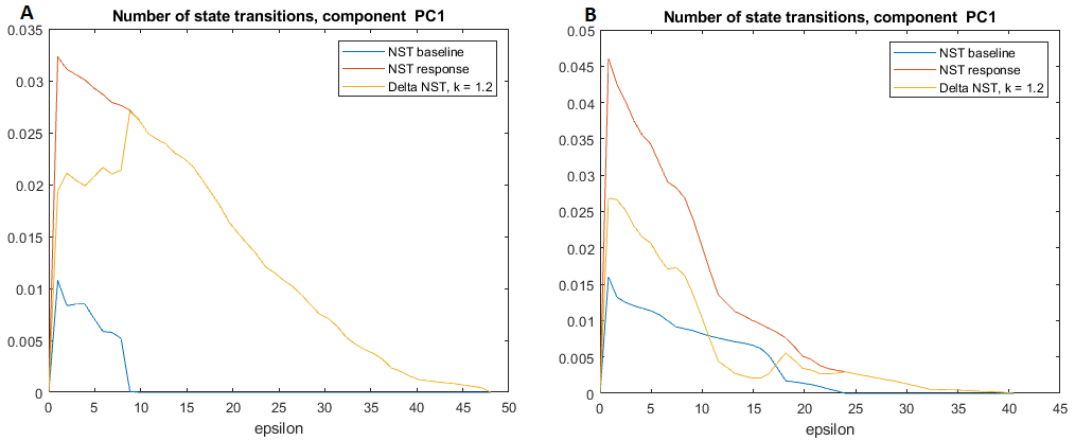

FIGURE 4.7: Normalized Number of State Transitions of subject 3 for Principal Component 1.
a) Normalized NST in function of the threshold ϵ for experiment 1, voluntary actions. b) Normalized NST in function of the threshold ϵ for experiment 3, involuntary reflexes.

4.3 Collected results and classification methods

The algorithm defined above was applied to all the valid subjects and experiments. In this section we will discuss the results obtained and try to apply classification methods to demonstrate if PCI is a valid measure for detecting voluntary and involuntary movements in healthy subjects.

4.3.1 Obtained PCI measures

The PCI obtained for each subject and experiment is enumerated in detail in table 4.1, and is visually represented in figure 4.8.

As it is possible to see from the figures, in the voluntary action experiment the PCI tends to be higher. However, the variance of the values is quite big, and the range of obtained measures varies between more than 100 and about 260. This variance gets even bigger in the case of the second experiment, in which it is possible to identify two macro-groups: the first one for those subjects in which, similarly to the first experiment, the visible CNV component in the ERP measurements determined an higher PCI value, and the second group for those in which the very low number of principal components, due to the prominent stereotipical behaviour among electrodes, affected the PCI with a lower value. In the third experiment for involuntary reflexes, as expected, the obtained PCI resulted in lower values and lower variance.

Clearly, the behaviour of the PCI in the second experiments tells us that probably a multi-class classification is not a good idea using this data. The semi-voluntary movements, indeed, determined a measure that cannot be fitted in the between of the other two experiments, but rather incorporates them both. For this reason, we don't expect to have good results in using non-binary classifiers.

	Perturbational Complexity Index (PCI)		
#Subject	exp1: Voluntary	exp2: Semi-voluntary	exp3: Involuntary
1	126.4298	142.1391	119.0455
2	151.0882	190.308	
3	172.6657	59.0406	80.4885
4	240.3718	76.3888	139.93
5	186.6717	91.8866	102.3606
6	246.7878	69.2528	100.4594
7	169.4423	64.7171	86.651
8	176.8913	73.658	125.5887
9	122.6754	152.9294	
10	155.0153	182.3589	
11	105.7229	97.2142	108.4465
12	151.3627	167.892	96.722
13	137.5652	167.7437	
14	183.6098	223.1921	
15	255.1215	225.24	
16	200.686	189.5639	
17	173.5344	205.7463	
18	199.5932	189.0868	

TABLE 4.1: Table of the obtained Perturbational Complexity Index (PCI) for each subject and experiment.

Even the binary classification using only experiment 1 and experiment 3 PCI, however, seems to be problematic: there is a relevant overlap among the two groups of measures, and so a valid thresholding with a low error rate does not seem possible.

4.3.2 Obtained Number of Principal Components

It is interesting, however, to have a look also at the number of principal components extracted in the PCA stage of the algorithm, determined by the stereotipical behaviour of the ERPs across electrodes. The obtained number of components, in particular, are enumerated in table 4.2 and visually represented in figure 4.9.

You can see that in the voluntary action experiment the number of components (NC) tends to be higher and for involuntary reflexes it is smaller. Again, for the second experiment the resulted NC is a wide range variable following the behaviour of both the other two experiments. This time, however, the variance of the measure seems to be similar for all three experiments.

Using this number together with the PCI as features to classify the will of the action might be a solution to obtain better results in the classification phase.

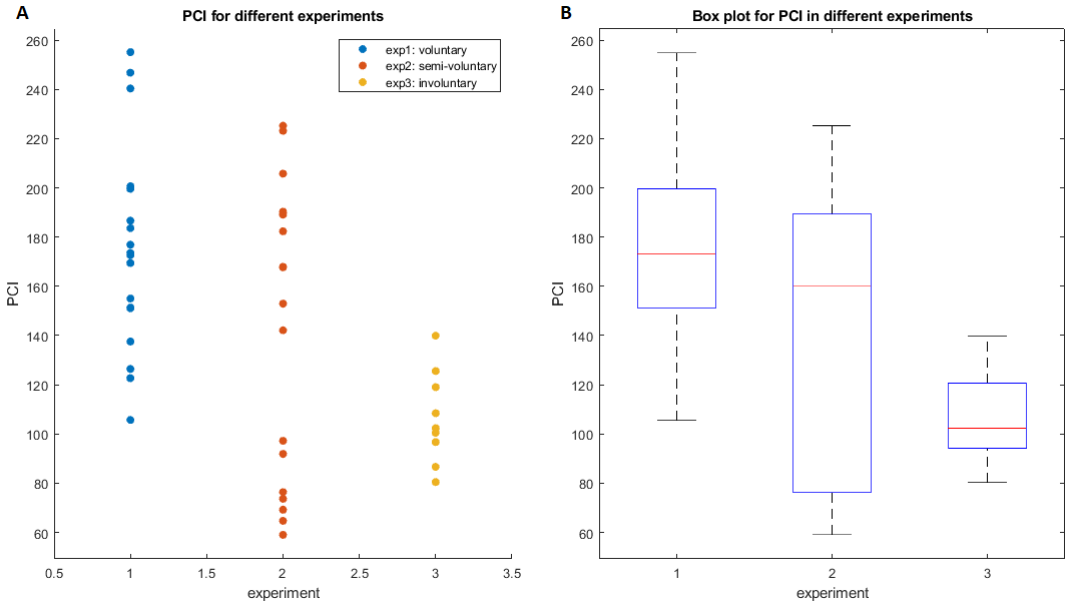

FIGURE 4.8: Obtained PCI for each experiment. As expected, PCI is smaller for involuntary reflexes and higher for voluntary movements, while semi-voluntary movements determine a wider range of possible values with high variance. a) PCI obtained by applying the algorithm for each subject in different experiments. b) Box plot of the obtained PCI. The red central mark represents the median of the group variable, the blue box identifies the 25th and 75th percentile, while the black whiskers extend to the most extreme data points.

4.3.3 Classification with Minimum Distance Criterion

The first method we applied to classify the obtained data is based on the **Minimum Distance criterion**.

Calling x_i the representative vector of class i, obtained by taking the mean values of the normalized features (PCI and number of principal components), the decision regions of the classifier according to the criterion can be written as:

$$R_j = \left\{ \mathbf{u} : \| \mathbf{u} - \mathbf{x}_j \|^2 \leq \| \mathbf{u} - \mathbf{x}_k \|^2 , \forall k \neq j \right\} \tag{4.1}$$

In other words, the classifier assigns to a given measurement the class that is the most *"similar"*.

Considering only experiment 1 and 3, we can apply a binary classifer, distinguishing only between voluntary actions and involuntary reflexes. Results can be summed up in a **Confusion Matrix**, in which each element c_{ij} is equal to the number of cases belonging to class i and classified as j. In the binary case, the elements of the matrix correspond to:

- The **True Positives**, those voluntary actions correctly classified as voluntary;

	Number of Principal Components		
#Subject	exp1: Voluntary	exp2: Semi-voluntary	exp3: Involuntary
1	8	7	10
2	10	10	
3	9	4	5
4	13	6	7
5	9	5	6
6	13	4	5
7	9	5	6
8	10	5	7
9	8	7	
10	8	10	
11	7	6	9
12	9	8	9
13	8	8	
14	11	11	
15	14	11	
16	11	10	
17	11	11	
18	11	9	

TABLE 4.2: Table of the obtained number of extracted principal components for each subject and experiment.

- The **True Negatives**, involuntary reflexes correctly classified as involuntary;
- The **False Positives** or **Type I errors**, voluntary movements mistakenly classified as involuntary;
- The **False Negatives** or **Type II errors**, those involuntary actions classfied as voluntary.

The true positive rate, or the probability that a voluntary movement is correctly classfied as voluntary, is called **sensitivity**, while the true negatives rate or probability that an involuntary reflex is classified as involuntary is called **specificity**. Of course, both sensitivity and specificity in a good classifiers are wanted to be as close as possible to 1, in order to have reliable test results.

The classifier was compiled and tested taking all elements of the dataset as training samples.

Figure 4.10 represents the confusion matrix obtained by the binary minimum distance classifier. As you can see, we obtained an 88.9% specificity and 77.8% sensitivity, which is a quite good result. More interestingly, the **precision** of the classifier, or probability that an element classfied as voluntary is actually voluntary, is 93.3%: this means that if using this method we classify a movement as voluntary, we can be sure with a probability very close to 1

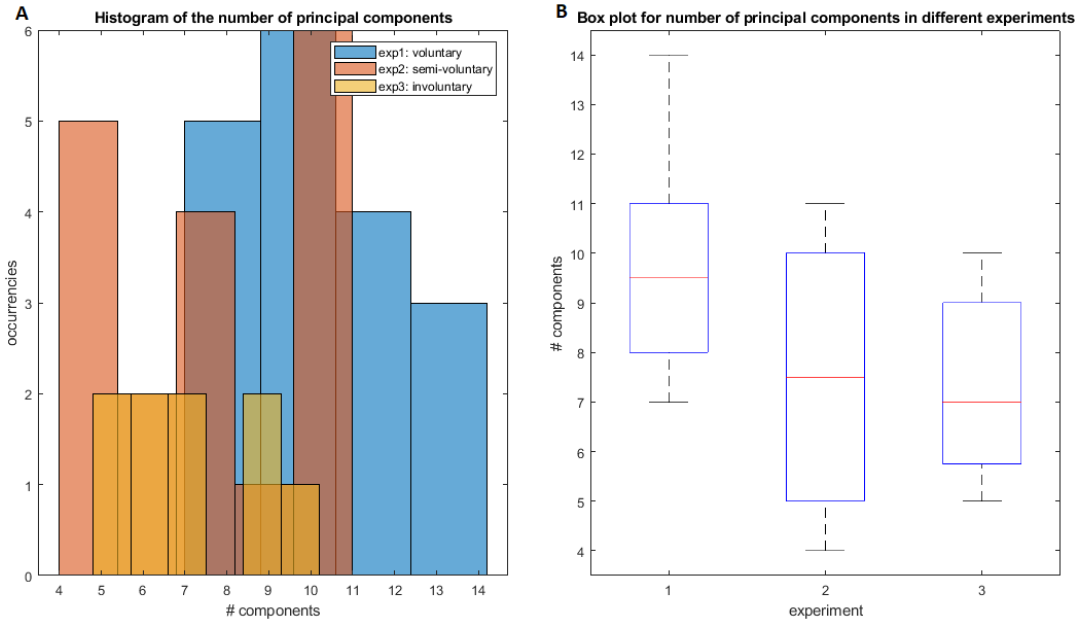

FIGURE 4.9: Obtained number of principal components for each experiment.
As expected, the number of components is smaller for involuntary reflexes and higher for voluntary movements, while semi-voluntary movements determine a wider range of possible values with high variance. a) Histogram of the number of principal components for each experiment. b) Box plot of the obtained number of principal components. The red central mark represents the median of the group variable, the blue box identifies the 25th and 75th percentile, while the black whiskers extend to the most extreme data points.

that the action was actually voluntary. Still, a relevant portion of voluntary actions are mistaken as involuntary.

We also applied the criterion to test a multi-class classifier, considering again all three experiments. As it is shown in figure 4.11, as expected, the classifier fails in identifying the semi-voluntary actions: due the fact that in this second experiment two groups can be identified, one with a visible CNV component giving results similar to the first one and the other in which like in the involuntary reflexes the ERPs are strongly stereotypical along all electrodes, the 77.8% of the semi-voluntary movements are labelled as one of the other two classes. The percentages of correctly classified involuntary and voluntary actions, nevertheless, are respectively 77.8% and 66.7%. It's important to notice, however, how low the percentages of how many estimated classes were correct are: saying that only the 46.7% of the estimated involuntary reflexes were actually involuntary, means that if we estimate a movement as involuntary there is less than half chance that is truly involuntary, and this percentage drops to 36.4% in the case of the semi-voluntary actions. Only in the case of the voluntary experiment this percentage is still relevant, and it goes up to 63.2%: this is particularly interesting if we notice that none of the involuntary reflexes were classified as

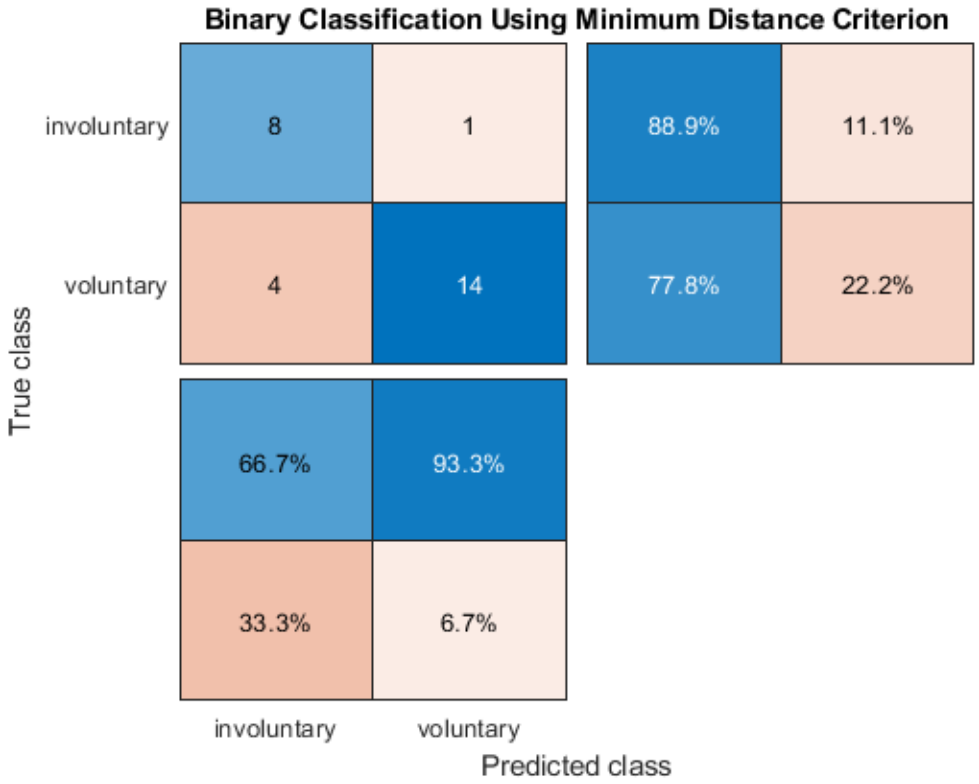

FIGURE 4.10: Confusion matrix in binary classification using the minimum distance criterion.
Blue cells represent correct classifications, red cells represent incorrect classifications. In the top left of the figure, the confusion matrix is represented. The true positives and true negatives are respectively 14 and 8, while the false positives (type I error) and the false negatives (type II error) are 1 and 4. In the top right of the figure, the percentages on the rows of the confusion matrix, or in other words how many true classes were correctly and incorrectly classified. Notice that 88.9% corresponds to the specificity of the classifier while 77.8% is the sensitivity, and 11.1% and 22.2% are respectively the false negative and false positive rates. In the bottom left of the figure, the percentages of the columns of the confusion matrix, or in other words how many estimated classes were correct or incorrect. Notice that 66.7% is the negative predictive value and 93.3% is the precision, while 33.3% and 6.7% correspond respectively to the false omission and the false discovery rate.

voluntary. This means that all of the actions labelled as voluntary are either truly voluntary or semi-voluntary, so a movement in which a certain level of *"will"* to act is included.

In conclusion, the results obtained by the minimum distance classifier may be interesting in the binary case, considering only purely voluntary and involuntary actions, but the multi-class classifier is still not reliable, since it fails to identify the semi-voluntary actions.

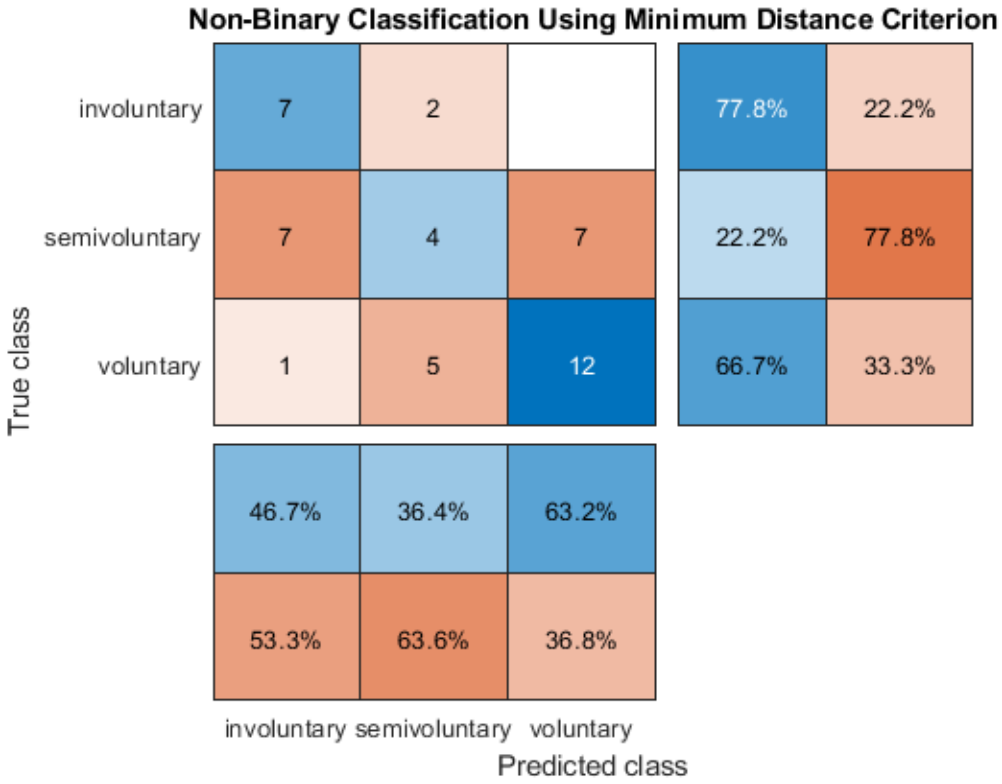

FIGURE 4.11: Confusion matrix in non-binary classification using the minimum distance criterion.
Blue cells represent correct classifications, red cells represent incorrect classifications. In the top left of the figure, the confusion matrix is represented. In the top right of the figure, the percentages on the rows of the confusion matrix, or in other words how many true classes were correctly and incorrectly classified. In the bottom left of the figure, the percentages of the columns of the confusion matrix, or in other words how many estimated classes were correct or incorrect.

4.3.4 Classification with Support Vector Machines (SVM)

The second classification method we applied is based on **Support Vector Machines (SVM)**.

The idea behind the binary SVM is to split the dataset in two hyperspaces of features, separated by an hyperplane. Of course, more than one hyperplane can exist allowing the wanted split between classes, but the SVM algorithm tries to find the solution at maximum distance from the closest points, called *support vectors*. The hyperplanes can be seen as the decision regions of the algorithm.

The classifier was trained with the leave-one-out technique: for each sample of the dataset, the algorithm was trained with all the other elements and tested on the excluded sample. The results reported in the following refer to the testing samples.

Figure 4.12 represents the confusion matrix obtained in the binary classification using SVM. In particular, the given results are comparable to

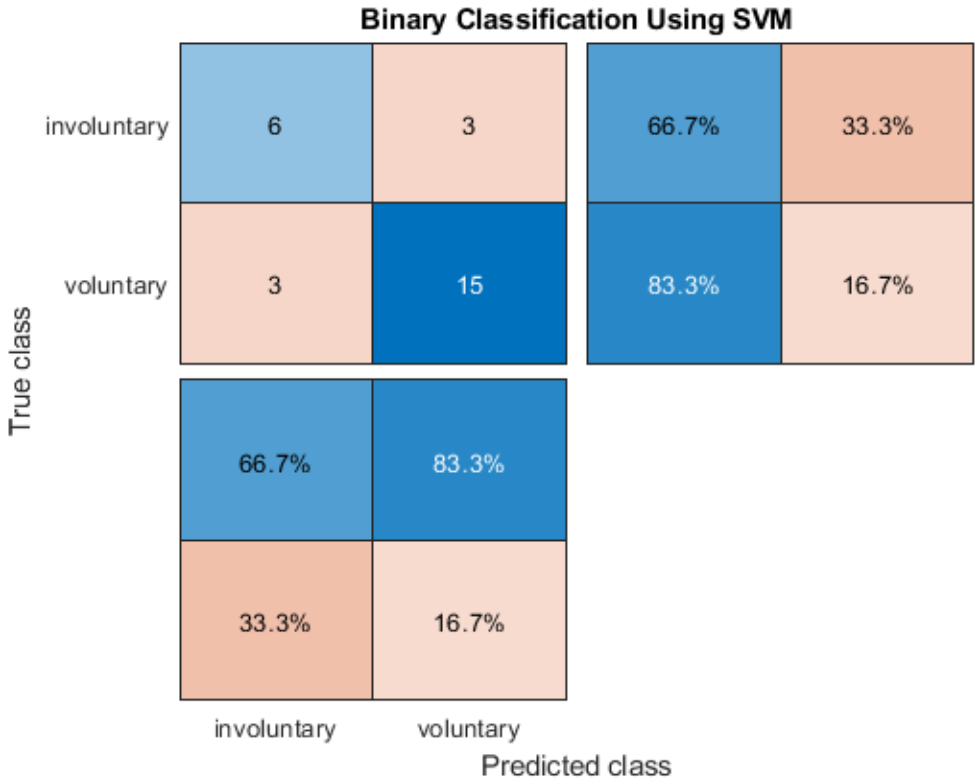

FIGURE 4.12: Confusion matrix in binary classification using Support Vector Machines.
Blue cells represent correct classifications, red cells represent incorrect classifications. In the top left of the figure, the confusion matrix is represented. The true positives and true negatives are respectively 15 and 6, while the false positives (type I error) and the false negatives (type II error) are 3 and 3. In the top right of the figure, the percentages on the rows of the confusion matrix, or in other words how many true classes were correctly and incorrectly classified. Notice that 66.7% corresponds to the specificity of the classifier while 83.3% is the sensitivity, and 33.3% and 16.7% are respectively the false negative and false positive rates. In the bottom left of the figure, the percentages of the columns of the confusion matrix, or in other words how many estimated classes were correct or incorrect. Notice that 66.7% is the negative predictive value and 83.3% is the precision, while 33.3% and 16.7% correspond respectively to the false omission and the false discovery rate.

those with the minimum distance method, with an 66.7% specificity and 83.3% sensitivity.

For the multi-class SVM classifier, the **One-Versus-One (OVO)** design method was applied, in which the binary SVM is run $K(K-1)/2$ times, with K equal to the total number of classes (in our case $K = 3$), to distinguish among all the couples of classes.

The obtained confusion matrix is shown in figure 4.13. As it is possible to see, the most interesting thing about this classifier is that all of the second

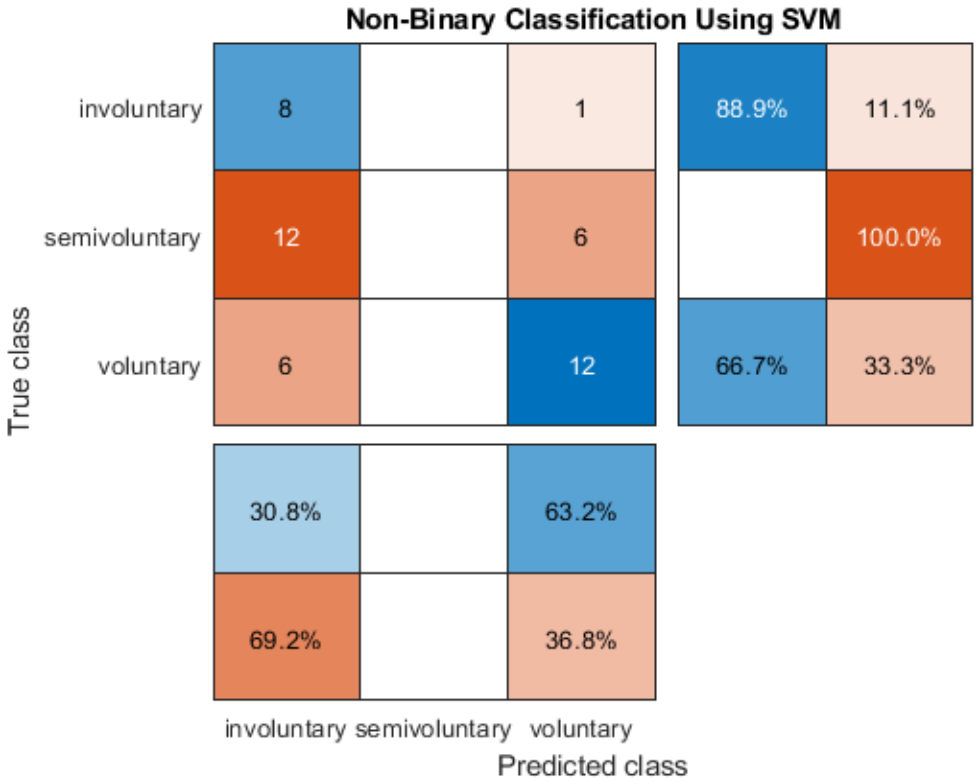

FIGURE 4.13: Confusion matrix in non-binary classification using Support Vector Machines.
Blue cells represent correct classifications, red cells represent incorrect classifications. In the top left of the figure, the confusion matrix is represented. In the top right of the figure, the percentages on the rows of the confusion matrix, or in other words how many true classes were correctly and incorrectly classified. In the bottom left of the figure, the percentages of the columns of the confusion matrix, or in other words how many estimated classes were correct or incorrect.

experiment elements are wrongly labelled as either voluntary or involuntary. This is because semi-voluntary actions, due to their double nature that we already discussed, are not considered as a distinct class but as belonging to one of the other two. Again, for this reason the non-binary classifier with SVM does not seem to be reliable, just as much as the minimum distance one. However, it might be interesting to try to fit the second experiment into one of the other two categories.

Chapter 5

Conclusions

The validity of calculating the PCI^{ST} with transcranial magnetic stimulation to detect signs of consciousness in healthy subjects and in patients was already discussed in 3.2.4.

The project developed for this thesis, however, took a step forward in the study, demonstrating that:

- **The Perturbational Complexity Index is a measure that can be calculated also on the ERPs of a subject, and not only from TMS-evoked potentials.** The Transcranial Magnetic Stimulation is an expensive technique, not accessible to all researchers, and proving that common EEG measurements and ERPs relative to other stimuli can be used to calculate PCI means that the measure could become a method that is both cheap and available to everyone.
- **The Readiness Potential, or intention of movement, can be used as a stimulus to set the baseline and response while estimating the PCI.** This is particularly important because it could lead the reaserch into calculating the PCI also in absence of a real physical movement and using only the "intention" to act, useful for those subjects who are unable to move.
- **PCI can be used to distinguish among a purely voluntary action from an healthy subject and an involuntary reflex.** Applying the measure in brain-injured patients, we could be able to identify a voluntary movement and assess the presence of consciousness in that patient.

This means, of course, that further researches on the subject, especially on brain-injured patients, may lead to more interesting conclusions and give a hand to move across the vast spectrum of *"consciousness"*.

5.1 Issues and limitations of the approach

Before concluding this thesis and articulate some ideas for the future, however, it is important to examine the main problems encountered while developing the project and the limitations of the described approach.

5.1.1 The dataset

We already mentioned in 4.1.4 the issues about the testing dataset. In particular, the available measurements were mostly badly taken or corrupted by noise, and on a large portion of subjects it was not possible to experience the involuntary patellar reflex, necessary for the third experiment. Due to this, the dataset had to be enlarged with *"fictious"* measurements, how we already explained in the previous chapter.

Despite being an approach widely used in literature, using *"fake"* subjects obtained from the real ones also brings some limitations. The results obtained from the fictious measurements, in fact, tend to be similar one another, and using them to test a classifier might lead to erroneous assumptions.

Moreover, the issues encountered in taking the EEG measurements, like electrode detachment and high impedence, would only get worse on brain-injured patients. Identifying the ERP components could be very challenging on patients, and so the calculated PCI might have a distorted value.

5.1.2 The classification

The most important issue with the used approach, however, lies in the results obtained by the algorithm and in particular in how they were used in the multi-class classification methods tested.

As we already discussed, indeed, both of the proposed classifiers are able to distinguish between voluntary and involuntary movements with good sensibility and sensitivity. This is a good starting point, but the real world is not only *black* or *white*, and it includes also a set of *greys*. And unfortunately, at the moment the non-binary classifiers we developed are not able to distinguish this *grey* area, the **semi-voluntary actions** .

What happens is that the elements from that class, instead of being considered as a new class are assigned to one of the other two. This may be due to two different possibilities. The first one, and the most desirable for our purpose, is that this is a problem restricted to the currently used dataset. In other words, enlarging the dataset and avoiding to include fictious measurements may lead to different conclusions and results, identifying more clearly a well defined class of semi-voluntary movements. The second option, however, is that the PCI is a measure that fails to capture this *"grey area"* between the purely voluntary movements and involuntary reflexes. This, of course, would be a critical problem in this approach: we cannot enclose the semi-voluntary actions into the involuntary ones, because they include a certain level of *willingness* that is not present in reflexes, but this *"willingness"* is also on a different level than the one that causes the voluntary actions.

5.1.3 Actions and consciousness

Finally, an essential limitation of this approach is in the main idea behind it.

At the center of this work, in fact, there is the idea that identifying a voluntary movement in a brain-injured patient is a sign of consciousness in the patient. And even if we can take for real this assumption, however, it is

important to notice that the opposite would never be true. After more studies, the method proposed in this approach could lead, in a near future, to take off many subjects from the darkness of a consciousness incommunicable to the external world, by demonstrating that a performed action included a certain level of willingness to move and was more or less voluntary. Nevertheless, if we state that a patient moved only due to an involuntary reflex, we could not assert that he/she is unconscious. In other words, the will to move is a sign of consciousness, but we cannot say that the absence of this will is a sign of unconsciousness.

Moreover, this approach is based on the concept of performing an action, while many brain-injured patients are unable to move, for example those in a locked-in state. It's true that with more studies on the subject it could be possible to extract the ERPs on the mere *"will"* to act without a real movement and extract the PCI measure even in absence of an action from the subject, but this would still be not enough. A relevant number of brain-injured patients, in fact, develop a form of **rejection** for the external world and refuse to communicate and respond to the examiner's stimuli. For these subjects and other similar cases, despite their conscious being, the proposed method would fail to find signs of consciousness, since they not only would not perform any action, but also not express any will to do it.

5.2 Ideas for the future

The limitations reported in the previous section must not induce the reader in diminishing the progresses obtained so far. This thesis proved that there is a difference in the amount of integrated information extracted from the ERPs of an healthy subject performing a voluntary movement or during an involuntary reflex, and the Perturbational Complexity Index may become an essential instrument to identify the will of patient to communicate with the external world.

This was just the first step of the study, and further researches may lead to overcome the issues and limitations we came across while writing this thesis. As a final discussion on this work, we propose here some ideas to continue the study.

5.2.1 Enlarging the dataset

The first and most important step is to enlarge the dataset with real good measurements and take off the fictious ones. More specifically, special attention must be taken in selecting for the study subjects with a visible phasic patellar reflex, since the number of available elements for the third experiment resulted particularly limited.

With a larger testing dataset, the study on both the PCI measure itself and the classification methods may become more precise. Consider that in the algorithm we defined in 4.2 the value of the PCI measure and the number of extracted principal components strongly depend on two parameters:

- the percentage of the total square sum of eigenvalues taken in the Principal Component Analysis phase (in this work, taken as 99%);
- the weight of the weighted difference of the number of state transitions among the baseline and the response (here, taken as $k = 1.2$).

In this thesis, we left the parameters as they were defined in the Comolatti's study, but changing one of them (or both of them) brings important changes in the resulting measures. Enlarging the dataset, these variables could also be studied and analyzed in order to obtain some settings providing a better differentiation among the classes.

5.2.2 Studying the semi-voluntary actions

Once the dataset is enlarged, particular attention should be given to the second experiment, the semi-voluntary actions.

If the double nature manifested by these elements in this work is attributable uniquely to the currently used dataset, then this problem would eventually disappear by using more real measurements and discard the fictious ones. If instead the PCI truly fails to capture the area in between a purely voluntary movement and an involuntary reflex, some changes to the protocol or the approach could be generated to overcome this limitation.

5.2.3 Testing on brain-injured patients

Finally, the approach could be tested on brain-injured patients: this would be the final and most important phase of the project, but also the most critical.

As we already said, the EEG measurement phase was already a demanding step in the healthy subjects. Most of them had to be discarded due to the high electrodes impedence, and this noise would only get worse in patients. Moreover, contrary to what happens in healthy subjects, patients are not able to repeat the trials as many times as needed, and for this reason the ERP components might become very difficult to be interpreted.

In the end, however, with some strong pre-processing of the EEG signals, we could be able to demonstrate if this approach actually works. We could prove to have found a new method to identify signs of consciousness by using a mathematical measure of easy and fast calculation.

Bibliography

[1] Adenauer G. Casali, Olivia Gosseries, Mario Rosanova, Mélanie Boly, Simone Sarasso, Karina R. Casali, Silvia Casarotto, Marie-Aurélie Bruno, Steven Laureys, Giulio Tononi, and Marcello Massimini. *A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior*. in Science Translational Medicine, 5(198):198ra105, (2010).
DOI: 10.1126/scitranslmed.3006294

[2] Mike X Cohen. *Analyzing neural time series data: Theory and Practice*. The MIT Press, Cambridge, Massachusetts, (2014).

[3] Renzo Comolatti, Andrea Pigorini, Silvia Casarotto, Matteo Fecchio, Guilherme Faria, Simone Sarasso, Mario Rosanova, Olivia Gosseries, Melanie Boly, Olivier Bodart, Didier Ledoux, Jean-Francois Brichant, Lino Nobili, Steven Laureys, Giulio Tononi, Marcello Massimini, Adenauer G Casali. *A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations*. Pre-print, not peer-reviewed, (2018).
DOI: 10.1101/445882

[4] Thomas M. Cover and Joy A. Thomas. *Elements of information theory – 2nd ed*. Wiley-Interscience, Hoboken, New Jersey, (2006).

[5] Arnaud Delorme, Toby Fernsler, Hilit Serby, and Scott Makeig. *EEGLAB Tutorial*. University of San Diego, California, (2006).

[6] Joseph T. Giacino, Kathleen Kalmar, and John Whyte. *The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility*. in Physical Medicine and Rehabilitation, 85(12):2020–2029, (2004).
DOI: 10.1016/j.apmr.2004.02.033

[7] Yvonne Höller, Aljoscha Thomschewski, Jürgen Bergmann, Martin Kronbichler, Julia S. Crone, Elisabeth V. Schmid, Kevin Butz, Peter Höller, Raffaele Nardone, Eugen Trinka. *Connectivity biomarkers can differentiate patients with different levels of consciousness*. in Clinical Neurophysiology, 125(8):1545-1555, (2014).
DOI: 10.1016/j.clinph.2013.12.095

[8] Hans Helmut Kornhuber and Lüder Deecke. *Brain potential changes in voluntary and passive movements in humans: readiness potential and reafferent potentials*. in Pflügers Archiv - European Journal of Physiology, 468 (7): 1115–1124, (2016).
DOI: 10.1007/s00424-016-1852-3

[9] Steven Laureys. *The neural correlate of (un)awareness: lessons from the vegetative state*. in Trends in Cognitive Sciences, 9(12):556-559, (2005).
DOI: 10.1016/j.tics.2005.10.010

[10] Benjamin Libet. *Unconscious cerebral initiative and the role of conscious will in voluntary action*. in Behavioral and Brain Science, 8(4):529–539, (1985).
DOI: 10.1017/S0140525X00044903

[11] Benjamin Libet, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. *Unconscious cerebral initiative and the role of conscious will in voluntary action*. in Brain, 106(3):623–642, (1983).
DOI: 10.1093/brain/106.3.623

[12] Steven J. Luck. *An introduction to the event-related potential technique – 2nd ed*. The MIT Press, Cambridge, Massachusetts, (2014).

[13] Oriano Mecarelli. *Manuale Teorico Pratico di Elettroencefalografia*. Wolters Kluwer Health Italia, (2009).

[14] Kwun Kei Ng, Simon Tobin, and Trevor B. Penney. *Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation*. in Frontiers in Integrative Neuroscience, 29(5):77 (2011).
DOI: 10.3389/fnint.2011.00077

[15] Rodrigo Quian Quiroga and Stefano Panzeri. *Extracting information from neuronal populations: information theory and decoding approaches*. in Nature Reviews Neuroscience, 10:173–185, (2009).
DOI: 10.1038/nrn2578

[16] Giacomo Rizzolatti and Corrado Sinigaglia. *So quel che fai: il cervello che agisce e i neuroni specchio*. Raffaello Cortina Editore, Milano, (2006).

[17] Simone Rumac *Single trial analysis of Readiness Potentials using Empirical Mode Decomposition*. (Master Thesis) Rel. Gabriella Olmo, Vito De Feo. Politecnico di Torino, Corso di laurea magistrale in Ingegneria Biomedica, (2018).
https://webthesis.biblio.polito.it/8512/

[18] Marco Sarà, Francesca Pistoia, Patrizio Pasqualetti, Fabio Sebastiano, Paolo Onorati, and Paolo M. Rossini. *Functional isolation within the Cerebral Cortex in the Vegetative State: a nonlinear method to predict clinical outcomes*. in Neural Rehabilitation and Neural Repair, 25(1):35–42, (2011).
DOI: 10.1177/1545968310378508

[19] Simone Sarasso, Melanie Boly, Martino Napolitani, Giulio Tononi, Steven Laureys, and Marcello Massimini. *Consciousness and Complexity during Unresponsiveness Induced by Propofol, Xenon, and Ketamine*. in Current Biology, 25(23):3099–3105, (2015).
DOI: 10.1016/j.cub.2015.10.014

[20] Simone Sarasso, Mario Rosanova, Adenauer G. Casali, Silvia Casarotto, Matteo Fecchio, Melanie Boly, Olivia Gosseries, Giulio Tononi, Steven Laureys, and Marcello Massimini. *Quantifying Cortical EEG Responses to TMS in (Un)consciousness*. in Clinical EEG and Neuroscience, 45(1):40–49, (2014).
DOI: 10.1177/1550059413513723

[21] Hiroshi Shibasaki and Mark Hallett. *What is the Bereitschaftspotential?*. in Clinical Neurophysiology, 117(11):2341–2356, (2006).
DOI: 10.1016/j.clinph.2006.04.025

[22] Giulio Tononi. *An information integration theory of consciousness*. in BMC Neuroscience, 5:42, (2010).
DOI: 10.1186/1471-2202-5-42

[23] Giulio Tononi e Marcello Massimini. *Nulla di più grande*. Baldini&Castoldi, Milano, (2013).

[24] William Grey Walter, R. Cooper, V. J. Aldridge, W. C. McCallum, and A. L. Winter. *Contingent Negative Variation: an electric sign of sensori-motor association and expectancy in the human brain*. in Nature, 203:380–384 (1964).
DOI: 10.1038/203380a0