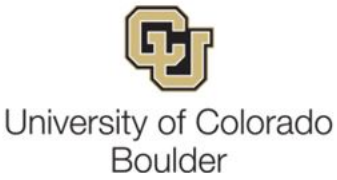
Master thesis : Individual differences across neural basis of pain

Auteur : Vandeleene, Nora
Promoteur(s) : Phillips, Christophe; 5198
Faculté : Faculté des Sciences appliquées
Diplôme : Master en ingénieur civil biomédical, à finalité spécialisée
Année académique : 2017-2018
URI/URL : http://hdl.handle.net/2268.2/4581



University of Liege

Faculty of Applied Sciences

Individual differences across neural basis of pain

Supervisors:
Pr. Christophe Phillips
Dr. Tor Wager

A Master thesis submitted by

Nora Vandeleene

in fulfillment of the requirements for the degree of

Master in Biomedical Engineering

Academic year 2017-2018

Abstract

Pain perception is a phenomenon qualified as ordinary in the mind of most people, but it is at the same time very poorly understood in a neurological point of view. Although a pain sensation is easily forgotten when felt for a short period of time, some patients are subject to different types of chronic pain, and this can indubitably lead to great discomfort in the everyday life. The neural basis of pain remain poorly known in the neuroscientific field, and each new discovery allows to make this knowledge grow.

The goal of this project was to determine whether or not different biotypes of pain co-exist. The idea is that different people correlate to pain in different manners. Thus, a large dataset grouping a total number of 433 participants coming from 13 different pain studies was constructed in order to conduct a stratification process and discover possible subgroups. For each participant individually, a pain-predictive weight map was created based on several single-trials maps, each of them associated with a pain rating given by the participant at the moment of the experiment. The weight maps were constructed with two different machine learning approaches: multivariate with PCR and univariate with OLS regression. A clustering algorithm, k-means in this project, was then applied to find four clusters in those weight maps, either directly in the maps or after some dimensionality reduction technique. The two dimensionality reduction methods leading to the best results were the application of two pain signatures patterns, NPS and SIIPS, and a dimensionality reduction using PCA, and both were better for the multivariate procedure. The different brain activation patterns associated with the clusters were then visualized using a one-sample t-test, and they in fact showed different and unique templates for each cluster. Both outcomes constitute interesting findings: the first one, based on predefined patterns (i.e. the signatures), reflects different behaviors in subjects towards the level of activation of each signature, in the idea of defining if some subjects track one signature more than the other. On the other hand, the second one exhibits different activation patterns that are more difficult to interpret.

Future studies should be dedicated to a deeper analysis of the different brain patterns discovered; in this project only their visualization is available, but it is not sufficient to draw accurate conclusions. A numerical analysis to determine statistically significant differences between them is the priority before claiming the discovery of four pain biotypes. Also, an expert's opinion is essential for the analysis of the activation patterns discovered for each cluster, especially in the PCA reduced maps, as this topic surpasses the content of the project.

Acknowledgments

I would like to thank different people that have contributed, from near or far, to the realization of this project.

First, I would like to express my gratitude to my supervisor and professor from the University of Liège, Mr. Christophe Phillips, for giving me the opportunity to get in contact with the CAN Lab and spend four months in Colorado.

I also want to thank my local supervisor, Mr. Tor Wager, for giving me the chance to work on a stimulating project, but mostly for the kind and cheering conversations all along the internship.

I am grateful for the help given by the researchers of the CAN Lab, especially to Bogdan for his constant help at the beginning of the project, and to Yoni for his advice at the end of it. I would like to thank also the ICS members, especially Mrs. Jean Bowen and Mrs. Kate Bell, for the numerous administrative details.

I would like to express my thanks to my parents, who spent a lot of time encouraging me and reviewing this project, but also all the smaller projects conducted for the past five years, even when they did not understand a word of it. I also thank them for giving me the resources to perform an internship abroad.

Finally, I would like to thank my classmate, projects partner and friend Valentine, with whom I shared a continent and a time zone for several months, for all her help, advice, but mostly encouragement during the entire realization of the project. I would like to thank her for taking the time to listen to the (many) problems encountered in this work, and always trying to find a solution with me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

The Cognitive and Affective Neuroscience (CAN) laboratory from the University of Colorado, Boulder, has been studying pain for several years. Various projects regarding the coding of pain in the brain have been conducted, involving a great number of scientists from different fields: engineering, psychology, medical doctors... Most of the measurements under study come from a functional Magnetic Resonance Imaging (fMRI) machine. Those measurements are used to depict brain patterns of activation when the participant is subject to a kind of pain.

Different studies have been carried with the CAN Lab team, for example, they have developed two different pain signatures, corresponding to two different patterns of activity in subjects under some specific painful stimulus [10, 12]. Those signatures, along with the corresponding study conducted, are developed in a following section. Different types of pain have been investigated, such as the back pain. At the moment, the researchers in the CAN Lab are working on various topics, for instance the craving sensation or the reaction to "pain sounds".

There are many aspects of pain inside the brain, and many ways to study them. In the studies conducted by the CAN Lab researchers and used for this project, the type of pain under study is usually a physical type of pain, more specifically, a thermal pain: the stimulus consists of a thermal stimulation on the forearm of the subject with a Peltier thermode end-plate [10], and depending on the level of heat, it can be qualified as painful or not. At this point, it is already clear that the participation of the subject is crucial; indeed, different people tolerate different levels of heat before calling them "pain", and thus, it is important to calibrate the levels labeled as "warmth" and "pain" for each participant.

There exist different types of pain. In the studies used for the following work, several have been investigated. For example, in the paper investigating the neurologic pain signature (NPS), in addition to the thermal pain described here above, they have examined another type of pain, called the romantic pain. To do that, the subjects chosen are people that had recently been through a romantic break up. The stimuli were simply the visioning of a picture of the ex-partner [10]. The goal here is to determine whether the different types of pain are coded in the same way within the brain or not. This could be interesting for future studies examining chronic pain or clinical pain, for populations that are not able to inform the surrounding people about their painful state.

1.1 Motivation

The neurologic response patterns to pain are still a great mystery in the scientific field nowadays. Although a neurologic pain signature (NPS) [10] has been developed by the CAN Lab team, there are still many things to learn on how each individual brain responds to different types of pain. It is thought that different brains correlate to pain in different ways. We wish to say that one specific brain activation pattern leads to a sensation of pain, but that is not true. Instead, pain sensations are coded in different manners in every person.

The goal of this study is to investigate individual differences in the neurological behavior towards pain sensation. This was accomplished by the means of stratification tools, in order to find different and unique patterns of brain activation related to pain within a large dataset (433 participants coming from 13 studies). The idea is to be able to find different subgroups, or clusters, of subjects which share the same pain-predictive pattern, and at the same time which have different patterns from subjects in other groups. Before this project, no one could tell whether these clusters even existed, and if they did, how many they were. Maybe some people's activation pattern looks very similar to the NPS signature pattern described in a following section, but maybe some other people react in a totally different way. And maybe these different ways could be similar in a number of participants. The discovery of different brain biotypes linked to pain would constitute an important tool for future studies. For example, a consistent succeeding research would be to study how each of these brain biotypes patterns are susceptible to respond to reappraisal, or a kind of medication for pain.
Every new discovery, not only in the understanding of pain processes but also in any medically related study, makes the general knowledge of the subject increase by bringing new information. This information alone might not seem extremely gainful, but most of the time they are crucial for the elaboration of future studies.

Drysdale and al. [25] have conducted a similar task in the idea of finding different biotypes linked to depression. Models of depression have been developed where the two main contributions to this mental state are the disinhibition of the central pain regulatory system and inhibition of the central pleasure system and of the psychomotor facilitatory system [26]. A depressive state is defined by nine symptoms, characterized by changes in mood, appetite, sleep, energy, cognition and motor activity, and a person is diagnosed with depression when subject to at least five of these symptoms. The detection of depression biotypes had already been investigated before, by identifying clusters of the nine symptoms mentioned above that seemed to co-occur within the patients. Several subtypes of depression have been determined with this method, they were characterized with changes in neuroendocrine activity, circadian rhythm and other potential biomarkers. However, this discovery was proved to be inconsistent and variable at the individual level, and it was not useful to distinguish patients from healthy controls or to reliably predict treatment response at the individual level. The idea behind the present study (Drysdale and al. [25]) is quite similar: the final goal is to determine several depression subtypes, but in this case the subtypes would correspond to neurophysiological biotypes constructed by clustering the subjects according to shared patterns of brain dysfunction.
This study consisted in a large, multisite dataset. The data under study were the measurements of a resting-state functional Magnetic Resonance Imaging (rs-fMRI) method. This is an application of the functional MRI where we can measure spontaneous, low-frequency fluctuations in the BOLD signal to investigate the functional architecture of the brain [27]. The analysis led to the revelation of four distinct biotypes linked to depression, which were characterized by homogeneous patterns of dysfunctional connectivity in frontostriatal and limbic networks. Those appear in Figure 1.1. One other interesting thing discovered during this study is that patients from different clusters respond

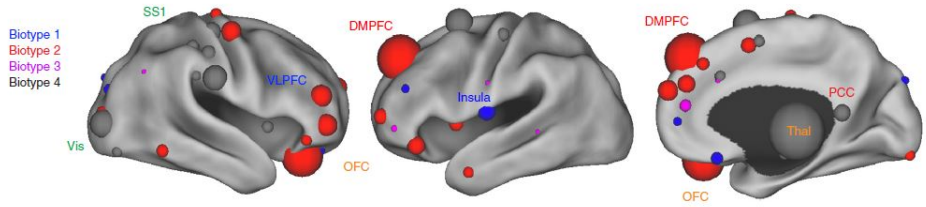


Figure 1.1: The four biotypes linked to depression determined by Drysale and al. [25]

to transcranial magnetic stimulation (TMS) in different ways. TMS is a treatment by neurostimulation, non-invasive, that is usually used for depression types resistant to medication. It works by modulating the functional connectivity in cortical networks [25]. The study has shown that biotype 1 subjects have a better response to TMS, compared with the other biotypes.

Although it was impractical to reproduce the exact same method because it would have required to reprocess the raw signal data, the finality of this work constituted the goal wished to be achieved in this pain biotyping project. Here, the data were taken with a functional MRI machine, and the idea was to work with the single-trials data for each participant individually. Each of the 433 subjects is described by a certain number of single-trial maps, each of these maps associated with a pain level (usually a temperature) and a pain rating (the evaluation of pain given by the participant, usually on a scale to 1 to 100). Single-trials maps contain parameter estimates for each voxel of the brain. They are an estimate of multiple repetition times (TR), TR being the pulse repetition interval, i.e. the time interval between two pulses applied to a slice [5]. In this case, they correspond to multiple time points in the fMRI signal timeseries. A deeper description of the fMRI signal appears in the next chapter.

The actual clustering was not applied directly on the single-trials maps, but instead on pain-predictive weight maps. Those were constructed by the means of supervised machine learning tools; they represent a linear combination of the component voxels weights from the single-trials in the shape of a vector of size equal to the number of voxels, and they are trained to predict pain rating of unseen data. Two different approaches were used for the construction of weight maps: multivariate, with a Principal Components Regression (PCR), and univariate, with an Ordinary Least Squares (OLS) regression. The difference between these approaches is developed in section 4.7.

Then, the clustering step, performed with k-means algorithm, could be applied to both data types separately. However, the very high dimensionality of the data (equal to the number of voxels, in this project it was 352 328) was very likely to give poor results. Therefore, several dimension reduction techniques were employed: the first one is the application of two pain signatures which were developed by the CAN Lab team, NPS and SIIPS. NPS tracks more the nociceptive information while SIIPS is able to track the contribution to pain beyond nociception. These different aspects to pain and a deeper description of the signatures appear in section 2.1 and 3.1, respectively. The two signatures are defined by another pain-predictive weight map, which can be applied to our maps. After this process, each participant is described by two values, corresponding to the regression parameter for each signature. The clustering is therefore applied in a two-dimensional space. The second dimensionality reduction technique is a simple reduction using Principal Components Analysis (PCA), with a number of components chosen to lead to the most reliable results. The third one is the reduction of the brain activation into seven canonical networks developed by Yeo

and al. from the Buckner Lab [70]. For the last two methods, the clustering algorithm is applied on the reduced maps, but the t-Distributed Stochastic Neighbor Embedding (t-SNE) method is used for the two- or three-dimensional visualization of the clusters [55].

The accuracy of the clusters discovery is assessed with several concepts. First of all, it seemed important to test the reproducibility of the clusters. If one participant belongs to one specific subgroup, he/she should always belong to this subgroup. Thus, two distinct datasets were created, each one containing half of the single-trials data for each participant. Each dataset is then used to construct pain-predictive weight maps; the first one used to determine the clusters and the second one used to test if the same participants form the same groups. The cluster reproducibility is equal to the percentage of subjects satisfying this condition. The second accuracy measure was the cluster consistency. Silhouette values were employed, a tool giving a numerical evaluation on "how well the subject should belong to this cluster". This consistency metric did not carry much weight in the final decision about which data type leads to the most accurate results, as data points with higher dimensionality are usually associated with lower Silhouette values. Finally, the within- and between-cluster spatial correlations were assessed. For each cluster, we want to find a high correlation with its corresponding cluster (in the two distinct datasets mentioned here above) and a low correlation between non-corresponding clusters. Those three tools are applied to each type of weight maps (multivariate or univariate), and after each dimension reduction technique (signatures, PCA and the canonical networks), in order to determine which type of maps leads to the best results.

A visualization of the brain patterns determined by the clusters is performed at the end of the process, for the two data types that gave the best results. This is done by the means of t-tests. Therefore, it was possible to compare the activation patterns, by observing the brain regions activated in each clusters, to make sure that they in fact exhibit different and unique activation templates.

The entire project was based on the search of four clusters, but other numbers of clusters were investigated. The choice for the optimal number was based on the cluster reliability and consistency value, and on the visualization of the dendrogram linked with hierarchical clustering.

1.2 Collaborating researchers

Considering that the number of fMRI measurements is limited for each study, the purpose of this work, pain biotyping, was inconceivable for one study alone. Thus, the need for collaborators is essential. Several researchers from the CAN Lab have conducted a number of pain studies, and they made the collected data available for everyone. This collaboration made it possible to construct a large dataset (here we found 433 subjects) and the clustering technique can have a meaning. Of course, this solution is not as optimal as it would be if all the subjects came from the same study, but for now, it is the only way. Indeed, every specialist has its own protocoles and preferred methods. For example, the study design might differ: some researchers prefer to keep it simple and induce only 3 levels of heat, but other ones prefer to use a little bit more tricky design, like with several levels of pain and sometimes with previous cues, i.e. a "warning" that informs the participant whether the following stimulus is going to be painful or not (and the cue is sometimes right, sometimes false). The study design varies a lot depending on the purpose of the study. Some researchers decide to study different types of pain at the same time, as it is explained in the NPS signature section. Also, different fMRI machines might be used. Obviously, when someone decides to conduct a study of pain, the measurements are used in many ways, in the purpose of finding different conclusions about them. The 13 studies brought for this clustering work were not conducted in the simple idea of the

clustering, but each scientist has different ideas on the results they want to emerge. All the data coming from the 13 studies was transformed to be as similar as possible: for example, the number of voxels composing the image is not a fixed value, thereby the images were resampled to have an equal number of voxels, here 352 328.
The thirteen studies, along with their names and the study or studies they are associated with, appear in Table 1.1. Some information are missing because a number of these studies were conducted very recently and all the information was not available.

	Name	**Number of participants**	**Short description of the study type**	**Prior publications**
1	BMRK3	33	*Use of cognitive self-regulation strategy*	Woo and al. (2015), PLOS Biology [15] Wager and al. (2013), NEJM, [10]
2	BMRK4	28	*Expectation effects on somatic and vicarious pain intensity*	Krishnan and al. (2016), eLIFE [16]
3	BMRK5	78	*Painsound study*	Unpublished
4	IE	50	*Two expectation manipulations*	Roy and al. (2014), Nature Neuroscience [17]
5	IE2	19	Unpublished	Unpublished
6	ILCP	29	*Perceived control and expectancy*	Liane Schmidt and al, (In prep.) [19]
7	SCEBL	26	*Cue-induced expectancy*	Koban and al. (In prep.) [20]
8	NSF	26	*Cue- and placebo-induced expectancy*	Atlas and al. (2014), Pain [14] Wager and al. (2013), NEJM, [10]
9	REMI	17	Unpublished	Unpublished
10	ROMANTIC PAIN	30	*Passive experience of a single stimulus level*	Unpublished
11	STEPHAN	40	Unpublished	Unpublished
12	EXP	17	*Cue-induced expectancy*	Atlas and al. (2010), Journal of Neuroscience [18]
13	LEVODERM	40	Unpublished	Unpublished

Table 1.1: Table summarizing information from the 13 studies used for the clustering

All the Matlab codes implemented in the context of this project are stored in the following platform: `https://github.com/canlab/pain_biotyping`

Chapter 2

Background

Contents

2.1 Pain

Although it seems logical to believe that the painful sensations are coded with the same receptors as the somatic sensations, it is not the case. Actually, the perception of painful stimuli, called "nociception", is linked with specific receptors at the surface of the skin: the nociceptors. Also, it should be noted that the processing of painful stimuli is multimodal, i.e. it implies discriminative, affective and motivational components. In addition to that, the brain regions contributing to the coding of pain sensations are multiple, and still not well understood. Those factors make the medical investigation of pain a difficult subject [3].

2.1.1 The nociceptors

The nociceptors are the nerve endings responsible for the transmission of pain sensations. They are located both on the skin and under it, and they proceed the transduction of the painful stimulus by triggering action potentials. Nociceptive fibers consist in a cellular body located inside the spinal ganglia, and an axonal extension with two endings: one towards the periphery and one towards the spinal cord or the brain stem.
The nociceptors can be divided into two categories, characterized by the amount of myelin surrounding the axon. This amount will in turn describe the velocity of the signal transmission, and it is always lower than for the somesthesic receptors that treat somatic sensations in a very rapid way. The subdivision can be made between $A\delta$ fibers, which are poorly myelinized and defined by

a conduction speed between 5 and 30 m/s, and C fibers, not myelinized at all, with a conduction speed usually smaller than 2 m/s. In addition, C fibers show a smaller diameter, which increases the slow conduction effect. A representation of the physiology of pain sensation, exhibiting both fibers types, appear in Figure 2.1.

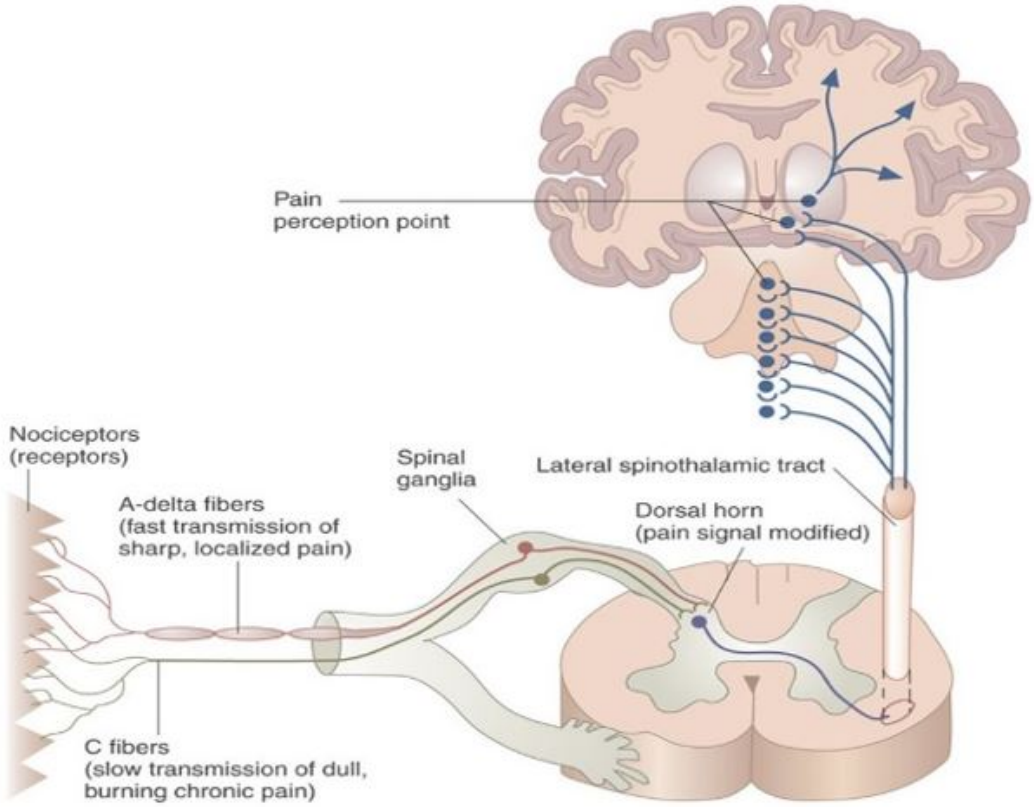


Figure 2.1: Physiology of pain sensation [85]

$A\delta$ fibers can once again be subdivided into two categories: type I $A\delta$ fibers possess high activation threshold for thermal stimuli but respond faster to mechanical and chemical painful stimuli, while type II $A\delta$ fibers are characterized by a very low threshold to heat but a quite high one to mechanical provocations. Knowing that, it is easily understood that the fact that the same type of stimulus, i.e. a thermal one, is used for the thirteen studies used in this project is important.
Experiments both on humans and animals have shown that somatic nerve endings are not responsible for the transmission of pain. Indeed, when several stimuli are applied to the subject, the nerve endings related to thermal and mechanical stimuli do not increase their firing frequency when the stimulus goes from nonpainful to painful. However, nociceptive fibers start to fire only when the level of the stimulus, for example a thermal level, reaches relatively high values. This proves the co-occurence of nociceptive and non-nociceptive thermal receptors inside the skin.

Usually, we define two types of pain: one is called "rapid", well localized, and the second one is called "slow" and lasts longer. Considering their respective amount of myelin, it is easily derived that $A\delta$ fibers are responsible for the rapid pain, usually defined as "acute", and C fibers for the slow pain, characterized by a longer and duller pain [3].

2.1.2 The transduction and propagation of nociceptive signals

The temperature threshold at which a thermal stimulus becomes painful is about 43°C, which corresponds to the nociceptive $A\delta$ and C fibers activation threshold. Note that this threshold may vary from one person to the other. The thermal receptors present in the skin come from the Transient Receptor Potential (TRP) channels family, more precisely they are called "vanilloids receptors", VR1 or TRPV1. Those are the receptors that are responsible for the thermal sensitivity, and they can also be activated with capsaicin, a substance present in hot peppers that confers their "burning taste". The structure of TRP channels is similar to regular potassic channels: they present 5 or 6 transmembrane domains with one pore between 5th and 6th domain. When it is at rest, this pore is closed. But when it is activated, and thus open, flows of calcium and sodium enter the receptors cells and this triggers the emission of action potentials from the nociceptive fibers.
So far, the receptors linked to mechanical and chemical stimuli remain less understood. It seems that other receptors from the TRP family are involved, for example TRPV2 and TRPA1, responsible for the detection of environmental chemical irritants. Receptors from another family have also been identified: Acid-Sensing Ions Channels (ASIC) family, which are thought to be the source of squeletic and cardiac pain following a pH change due to ischemia. However, since the stimulus in question in this work is exclusively thermal, this topic is not investigated in details.

The graduated potentials generated at the nociceptive endings need to be converted into action potentials in order to be transmitted to the spinal cord synapses. The voltage-dependent sodium and potassium channels play a major role in this transmission. More precisely, a sodium channel subtype, called "Nav1.7", seems crucial for the propagation of nocicpetive information. Indeed, Nav1.7 anomalies are very often linked to pain sensation impairments in human beings. Mutations of the Nav1.7 gene can induce either loss of function and incapacity of feeling pain, or hyperexcitability which leads to intense burning sensations. On the other hand, Nav1.8 gene is believed to play a role in the transmission of mechanical and chemical painful stimuli [3].

2.1.3 Pathways specific to pain

The pathways specific to pain transmission originate in the sensitive neurons from the spinal ganglia: just as the other sensitive neurons, the nociceptive neurons axons enter the spinal cord through the dorsal root. When they arrive in the dorsal horn, a bifurcation takes place, the fibers separate between an ascending and a descending branches. This bifurcation is called "Lissauer's dorso-lateral tract". Then, the fibers travel a short track inside the spinal cord before arriving inside the grey matter of the dorsal horn. The dorsal horn is divided into 5 layers, called "Redex layers". The first and fifth layers are equipped with projection neurons with quite long axons that reach the brain stem and thalamus targets. Layer 2 contains interneurons. Fibers that enter the dorsal horn are not fairly shared between the five layers: for example, C fibers only reach layers 1 and 5. Knowing that, it is clear that studying the effects of pain is already a difficult subject. The spinal ganglia and the dorsal horn are shown in Figure 2.1 from the previous section.
In the spinal ganglia, we find two types of fibers: primary and secondary. The primary fibers enter the spinal cord through the dorsal horn, emit collateral fibers that extend onto several segments of the spinal cord which also end in the dorsal horn. On the other hand, secondary fibers emit, from the dorsal horn, long axons that reach up the lateral spinothalamic tract, as shown in Figure 2.1. Those are the ones that reach the brain stem and thalamus targets [3].

2.1.4 Parallel pathways

The secondary fibers mentioned above project onto several structures of the brain stem and forebrain. This proves the diversity of the neural network coding for the nociceptive information. The precise role of this structure is still unknown, but it is clear that each projection target deals with particular aspects of the sensorial and behavioral responses to a painful stimulus.
Some components of this system are responsible for the processing of discriminative aspects of the pain sensitivity, such as location, intensity and nature of the painful stimulation. This information seems to be present in the ventro-postero-lateral core in the thalamus before it reaches primary and secondary somatosensory cortices neurons. The nociceptive data remain distinct and isolated until they arrive in the cortical area.
Other portions of the system treat motivational and affective pain aspects, such as bad feelings, anxiety, fear etc. The brain areas targeted by those projections are the following: several portions of the reticular formation and the superior colliculus, the midbrain central grey matter, the hypothalamus and the amygdala [3]. These aspects of pain are responsible for the person-dependent level of tolerated heat before calling it pain. The feelings and beliefs about pain might differ a lot from one person to the other, and this leads to a direct consequence on pain sensation [4].
This section implies that different aspects of pain are coded in different brain areas, still with some assumptions made, which is a subject that was deeply investigated in the CAN Lab. This activation has been found by the means of two medical imaging tools: electrophysiology and function MRI measurements. In this work, the data are exclusively of the second kind.

2.1.5 Other cognitive processes related to pain

Beyond nociception, pain perception depends on several cognitive processes able to influence the painful sensation in a better or worst manner. For example, attentional processes play an important role: indeed, one person might feel less intense pain if distracted from it. One other process is the formation of expectations about pain and reappraisal of the experience, which both depend on previous situations. For instance, a person who is resistant to pain medication and aware of it might feel more pain than other persons. In the same spirit, if a person knows that they react very well to some medication, the simple action of taking this medication might release a portion of the pain sensation. This second effect is know as the "placebo effect", when pain is reduced with the only idea of having received a potent painkiller.
More complex cognition processes play a role in the pain regulation, such as the ones related to the perceived threat of pain. In this case, it depends on what the person subjected to pain thinks this pain means, which is subjective. For example, if the person subjected to pain believes this sensation is related to some pathological process, the perception will be increased.
All of these processes might influence pain sensation in both ways: either by amplifying it or, on the contrary, by attenuating it. A high perceived threat value of pain will tend to a more intense painful sensation, while a reappraisal of pain will induce the opposite effect.
Expectations about an upcoming event allow the individual to prepare and to adjust different systems in the body, such as the sensory, cognitive and motor systems, to produce the adequate responses. In the case of thermal stimulations, a cue that prefaces the stimulus will induce an increased signal in some brain areas related to pain during the time separating this cue and the actual stimulus. Perceived control refers to the extent to which a person can perceive their pain as controllable: people who perceive a high degree of control will tend to persist more in the face of failure, while people who perceive a low degree will withdraw more rapidly. Perceived control is thought to trigger the reappraisal process, because an aversive event will seem less threatening if the idea of control is possible. In an experimental way, perceived control can be induced to participants by allowing them to stop the thermal stimulation when they cannot stand it anymore. This

is internal control, as opposition to external control when the painful stimulus is induced by the scientist conducting the study, and it has been shown to reduce pain intensity ratings [4].

2.2 Functional MRI studies

Functional Magnetic Resonance Imaging (fMRI), which was introduced in the early 90's, is a powerful tool in the neuroscience field in order to study brain activity and connectivity. It is a fast, non-invasive instrument that provides brain images and allows scientists to study the responses of the neurons to different stimuli, or simply in their rest state (in this second case it would be called "resting state MRI", or "rsMRI", mentioned in the Introduction). Functional MRI experiments are always linked with some specified study design that reflects the brain activation phenomenon the laboratory wants to study. For example, if a laboratory wants to investigate the brain responses to the visioning of happy v.s. angry faces, they will follow a design where the subject inside the machine has to look at happy and angry faces, in a randomly selected order, every stimulus separated with a period of rest. The goal of the fMRI experiment is to collect images from all three states (rest, angry faces and happy faces) in order to be able to compare them and discover patterns of activity that are specific to them.
Although, a number of pre-processing steps have to be followed in order to obtain correct and interpretable results [6]. The different steps observed to construct the data used for this work are described in a following section.
Functional MRI uses a similar technique and the same equipment as the regular MRI [7]. The basic concepts of MRI are not developed in this thesis, while those specifically linked to functional MRI are explained in the following sections. For a review of MRI basic principles, see Ref. [5], *Medical Imaging Signal and Systems* from Jerry L. Prince and Jonathan M. Links.

2.2.1 The BOLD signal

The signal measured with the fMRI is called the Blood Oxygenation Level-Dependent (BOLD) signal. It corresponds to an indirect measure of the neurons activity within the brain. Actually, it measures the changes in regional cerebral blood flow, volume and oxygenation, which follows a certain modification in neuronal activity induced, for example, by a stimulus or a task.
The basic concept underlying this phenomenon is the following: when a region in the cortex is activated due to a stimulus/task, the demand for oxygen and nutrients in this region increases. The only way to bring those requests is to increase the blood flow in the area in question. Figure 2.2 shows a representation of the blood flow increase induced by the growing neural activity. The blue dots represent the oxygen molecules that are sent to the neurons to allow them to perform their proper tasks.
The blood flow increase is usually more important than what is actually needed for the neurons, and thus, at the capillary level, we end up with more oxyhemoglobin molecules than needed. This results in the balance of oxygenated arterial blood to deoxygenated venous blood showing a net increment. The BOLD contrast measures inhomogeneities in the magnetic field due to changes in the level of oxygen in blood. In terms of magnetic susceptibility, oxygenated and deoxygentaed blood have dissimilar behaviors. The oxygenated blood is diamagnetic and does not induce any signal loss, while the deoxygenated blood is paramagnetic and induces one [8]. This comes from the following phenomenon: when the hemoglobin is not bound to oxygen, the difference between the magnetic field around this hemoglobin molecule and the magnetic field induced by the MRI machine is much greater than for hemoglobin that is bound to oxygen. This means that oxygenated blood makes the local magnetic environment more uniform. The magnetic state of the environment is felt

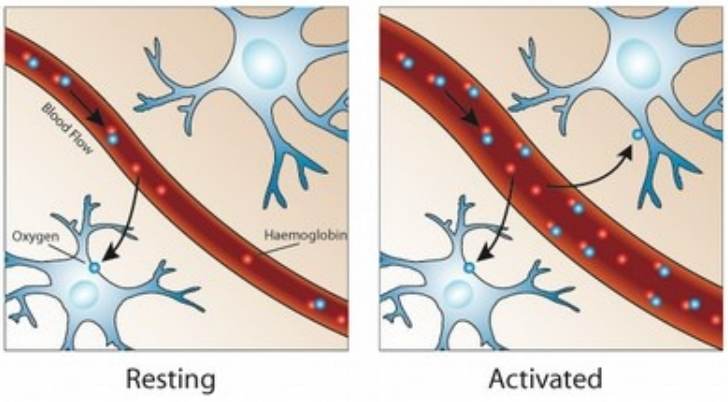


Figure 2.2: Origin of the BOLD signal [81]

by the water molecules in the tissues, and they are responsible for the longevity of the signals used to produce magnetic resonance images: if the field is less uniform, if there are more inhomogeneities induced by the deoxygenated blood, there will be an increase in the mixture of different signal frequencies that come from the sample, and this will induce a faster decay of the overall signal. Thus, when the increase in blood flow described right above appears, along with the decrease in the ratio of deoxygenated blood to oxygenated one, the environment is more uniform in the area in question and the MRI signal from that area decays more slowly. This results in a stronger signal when recorded in a typical magnetic resonance image acquisition, and this is what constitutes the BOLD signal recorded in fMRI. It should be noted that this increase is quite small, in the order of 1 % of the original signal [6, 7].

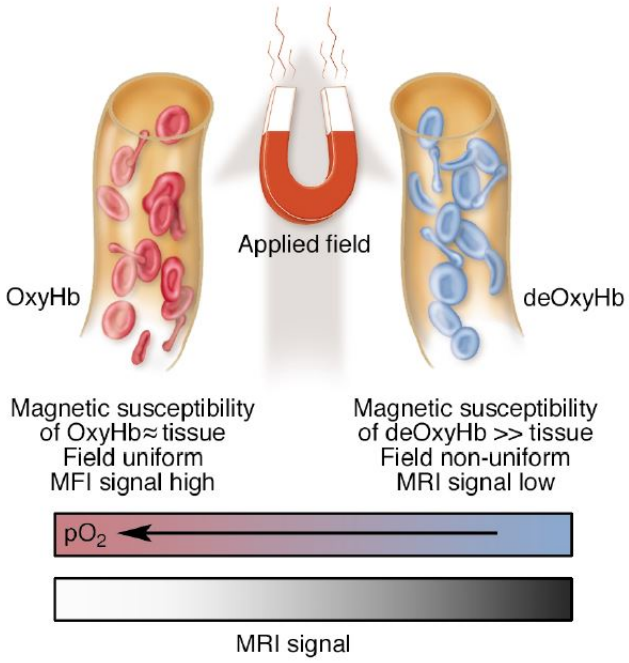


Figure 2.3: Illustration of the origin of the BOLD effect in fMRI [7]

Figure 2.3 shows a more precise representation of the phenomenon described here.

The change in magnetic resonance signal, which is triggered by instantaneous neuronal activity,

is called the Hemodynamic Response Function, or HRF. This constitutes a theoretical representation of the BOLD signal, as shown in Figure 2.4. In this Figure, it can be noted that the curve is initially at a resting state, called the baseline, when no stimulus is applied. Just as the stimulus appears, the first portion of the curve is a small decrease of the signal, called the "initial dip", which is thought to come from the initial decrease of deoxyhemoglobin, although it has not been proven yet. This decrease in deoxyhemoglobin comes from the fact that when the neuronal activity suddenly increases, triggering at the same time an increase in blood flow, the first blood molecules entering the region in interest will release their oxygen in order to fill the need for oxygen from the activated neurons. This appears for a very short period of time where the deoxyhemoglobin molecules are more present. Then, as explained above, there is an over-compensation of the blood flow, and thus the oxygenated blood remains in the vessels and finally dilutes the deoxygenated blood and exceeds it in proportion. This phenomenon explains the peak in the BOLD signal that follows the initial dip. The peak lasts about 4 to 6 seconds following the activation. Just after this period of time, the signal decreases and reaches a value under the baseline level. This portion, called "undershoot", comes from the combination of reduced blood flow and increased blood volume.
Figure 2.4 shows a theoretical hemodynamic response to a brief event, i.e. a task or a stimulus [1].

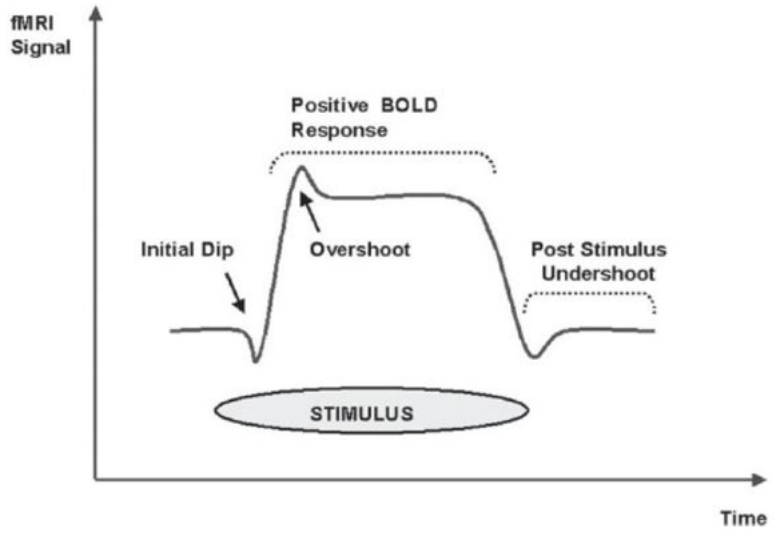


Figure 2.4: Model of the Hemodynamic Response Function (HRF) explaining the BOLD effect [82]

This specific shape is important to take into account to continue the study as the easiest measurable effect, i.e. the peak, appears a few seconds after the event we want to recover. However, the model presents a few limitations. First of all, as said before, the magnitude of the signal at the peak is very small (usually between 0.1 and 5 %), and it can be hardly detected. Secondly, the response is delayed, as mentioned above, and quite slow. Short events, such as clapping hands, can induce very long responses, sometimes until 30 seconds. Finally, studies have demonstrated that the shape of the response may vary with different people, but also in different regions within the brain, which makes the analysis of the signal even more hazardous [1].

2.2.2 Noise

One other problem linked with this HRF model is the presence of noise. The source of this noise might be the machine itself, the hardware, or the subject under study. It induces several effects on the measurements, such as high-frequency "spikes", image artifacts and distortions, low-frequency

drift and periodic fluctuations over time [1]. It is crucial to separate the true signal from noise in order to analyze and interpret the results. The first step is to be able to identify those sources of noise, and then they might be removed from the BOLD signal.

The four main categories of noise sources are the following [1]:

- **Thermal motion of free electrons in the system**: This type of noise is sometimes called the "background noise", and it is independent of the signal of interest, i.e. it appears even if there is no activity-related signal of interest. It produces a thermal noise which arise from the thermal agitation of charge carriers in the subject, but also in the MRI system electronics. One way to measure this type of noise is to acquire the data without exciting any magnetization. The thermal noise is always present, but its effect might be reduced by the means of low-pass filtering. Indeed, it is a wideband noise source, i.e. which has the same power at all temporal frequencies, and filtering allows to eliminate noise from frequencies that are outside of the signal band of interest [9].

- **Gradient and magnetic field instability**: The instabilities and imperfections in the gradient and magnetic field can lead to variations in the spatial frequency components acquired over time, which can in turn induce temporal variations in the reconstructed image. One way to reduce this contribution to noise is to use a well designed and maintained MRI system [9]. A noise which is always present, and mostly due to the scanner instabilities and not the physiological noise (since it is present even in cadavers or in phantoms), is the drift. It represents slow changes in voxel intensity over time in the fMRI signal, it can be categorized as a low-frequency noise. Since it cannot be avoided, this drift must be taken into account during the preprocessing steps or the statistical analysis. In order not to confuse drift with task-related signal, it is important to use high-frequency in the experimental design, e.g. more rapid alternations of stimulus on/off states [1].

- **Head movement and its interaction with magnetic field**: Usually, head motion is corrected with preprocessing techniques. However, some remaining effects of motion cannot be suppressed, which are usually called "spin-history" artifacts. Those are due to complex interactions with the magnetic field, a subject that is not addressed in this project. Those effects might be reduced with statistical analysis but could never be entirely suppressed. One thing to be careful with is the task-related head movement: you might find some activation pattern by studying a specific task, but if this task is related to a specific movement each time it is performed (e.g. the task under study is back pain, you might use a device that will push the back of the participant and this will probably produce an unwanted displacement of the participant's head), you might be confusing the real task-based activation pattern with a regular head-movement noise [1].

- **Physiological noise**: The two main sources of physiological noise, without considering subject motion which is explained in the previous item, are cardiac pulsations and respiratory activity. Cardiac pulsations lead to dynamic changes in the relative distribution of brain tissue, blood and cerebrospinal fluid. The respiratory activity produces expansions and contractions of the chest cavity and walls, which leads to dynamic variations in bulk magnetic susceptibility and related perturbations of the main magnetic field. Both of them will induce distortions in the acquired images that vary across space and time. Respiratory activity may also have an effect on the signal due to the regulation of carbon dioxide (CO_2) that is function of the respiratory system. Indeed, carbon dioxide might play a role of vasodilator, and then the variation in the level of CO_2 might lead to dynamic variations in cerebral blood flow, which constitutes an additional source of noise [9].

Those two types of physiological noise are linked with aliasing problems: the periodic signals that occur more rapidly than the sampling rate will often be aliased back to a lower frequency. This sampling rate is dictated by the Repetition Time, or TR [1]. If for example you use a TR of 2.5 seconds, the sampling rate is equal to 0.4 Hz, and the cardiac noise with a frequency of 1 Hz will appear at a frequency of 0.2 Hz [9]. To avoid aliasing, the sample rate must be at least twice as fast as the fastest frequency in the signal [1].

Those noise and artifacts might be mitigated on the one hand during the acquisition, and on the other hand during the analysis of the results. During the acquisition, some necessary precautions have to be set in place, such as: implementing a good quality control process for the fMRI machines, to make sure it works correctly, choosing the right acquisition sequence and parameters for your goal in neuroimaging, i.e. what you want to study, using specialized IRM sequences such as spin-echo, simultaneous multislice etc. which might reduce some type of noise, and finally limiting the head movement of the participant inside the scanner. For the analysis, there exist other ways to reduce the effect of noise, for example: identifying and correcting the outliers and artifacts, using standard pre-processing techniques to adjust for the head movement and slow drift, invoking some helpful statistical procedures such as robust regression or hierarchical modeling, and modeling low-frequency drift and periodic fluctuations over time [1].

2.2.3 Predicted response

For the continuity of the study (e.g. for the General Linear Model (GLM) design matrix), scientists usually need to construct an expected BOLD signal. To do so, they simply perform a convolution between the study design and the HRF model developed here above. In the example in Figure 2.5, it can be seen that the study design is task-based (as opposition to resting-state) and corresponds to a block design.

Figure 2.5: Generation of expected BOLD signal using HRF [83]

The block design consists of presenting consecutive stimuli as a series of epochs, or blocks, with stimuli from one condition being presented during each epoch, followed by an epoch of stimuli from another condition, or with rest/baseline epochs [6]. Indeed, in this example, each block corresponds to an epoch of a certain stimulus, or task, separated by epochs of rest. The two other types of task-based designs are the event-related design, which is composed of randomized sequence of discrete, short events separated by an inter-stimulus interval, and the mixed block/event related design, which is a combination of the two first designs [6]. A representation of the three task-based designs appear in Figure 2.6.

Figure 2.7 shows an example of the construction of an expected signal by taking into account the low-frequency noise, i.e. the drift. In this figure, the blue curve represents the data at is has been recorded with the fMRI system. The goal here is to construct the expected signal, the green curve, as similar as possible to the data. To do that, we use the HRF function convoluted with the study design, as explained right above and represented as the red curve, but also a representation of the

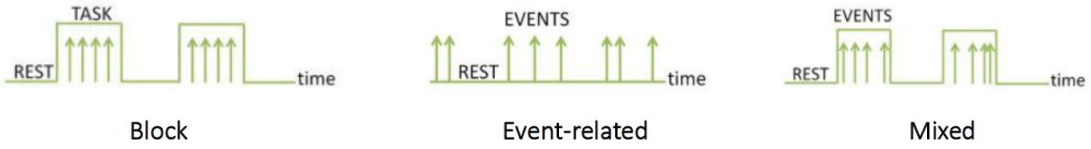


Figure 2.6: The three task-based designs [6]

drift, shown in Figure 2.7 as the black curve. The result is quite similar to the expected signal following the trend induced by the drift [1].

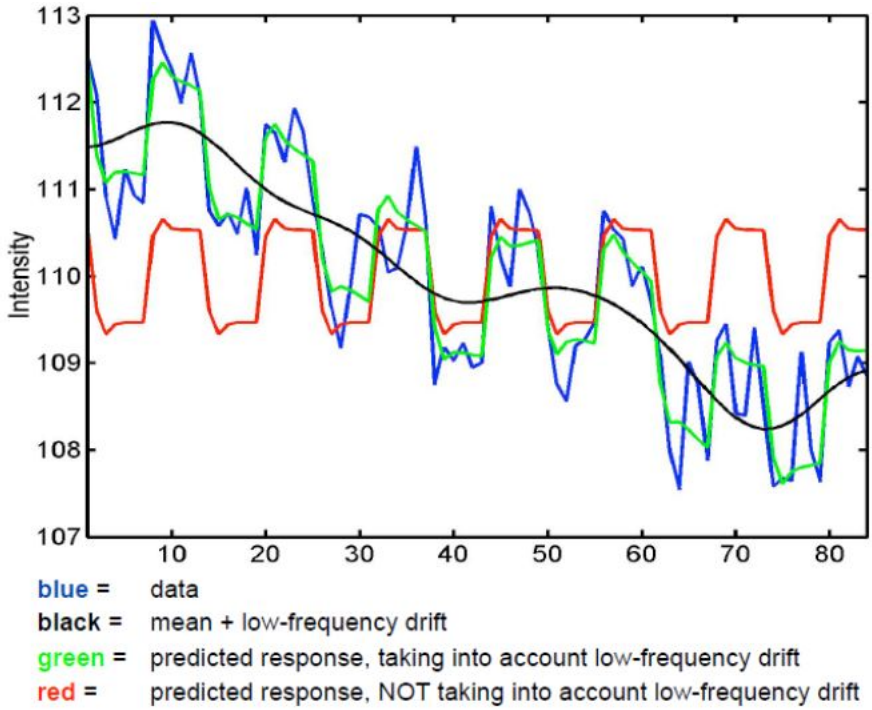


Figure 2.7: Construction of the expected signal [84]

2.2.4 Limitations

Due to the very nature of fMRI studies, the number of images is usually limited. First of all, the number of participants is usually quite small, around 30 for most studies. Secondly, the number of images taken for each participant is limited to the time it takes for the MRI machine to capture one image. A typical fMRI study consists of several tasks, or stimuli, separated by time periods of rest. Depending on the device, the construction of one brain volume takes approximately 2 seconds, and the idea is to take "as many as possible" images in various states (rest state and stimuli corresponding, in our case, to the different levels of induced pain). The number of constructed volumes could be maximized by the means of two methods: either by reducing the spatial resolution, which would result in an increase in the temporal resolution. The spatial resolution is usually around $1mm^3$, which corresponds to the size of the voxels, i.e. a small cubic volume corresponding to a group of neurons within the brain. Reducing this resolution even more would lead to extremely large voxels and the results would be less interpretable, some activation patterns could be missed. This explains the difference in resolution between functional and regular MRI. In the first one, the goal is to study the brain function, i.e. how it responds to different stimuli. The anatomy here is

not the priority, although it should not exceed some minimum so the results are still interpretable. It is crucial to take a certain number of images in different states to study the time variation of this response. On the other hand, the regular MRI focuses on the brain anatomy, in this case the spatial resolution is maximized and there is no need to take the temporal resolution into account. Although, by increasing the spatial resolution, the scan time is longer and longer, and this could induce more and more movements from the subject inside the machine, which could lead to false results. In both types of MRI, a trade-off must be taken between temporal and spatial resolution [2]. Also, by increasing the spatial resolution in fMRI experiment, physiological noise effects of aliasing explained here above are increased [9].
The other solution would be to make the whole study last longer, then the number of constructed volumes could be higher with the same resolution. However, the tasks running should not exceed some specified time limitation, or else it could result in the apparition of unwanted activation within the brain, such as regions related to fatigue or boredom. These could distort the results and should be avoided.

These limitations make the analysis of functional MRI difficult, on the one hand because the number of observations is poor; on the other hand since those measurements are naturally very noisy, as explained in the previous section.

Chapter 3

Previous work

Contents

As stated in the Introduction, two signatures patterns called NPS and SIIPS, developed by the CAN Lab team, have been used to reduce the dimensionality of the initial data and investigate the behavior of subjects towards those signatures.

3.1 Pain signatures

One major discovery conducted in the Can Lab is the determination of different brain signatures related to pain. In 2013, the article "An fMRI-Based Neurologic Signature of Physical Pain" was released, describing the activation pattern within the brain related to the physical pain. This signature is called "Neurologic Pain Signature" or "NPS". Three years later, a second paper is released : "Quantifying cerebral contributions to pain beyond nociception". It presents a second type of pain signature, called the "Stimulus Intensity Independent Pain Signature" or "SIIPS1". This is quite different from the first one, NPS, since it is associated with neuronal activity which is not due to the activation of the nociceptive receptors inside the skin. In addition, the second signature allows the prediction of variation in pain intensity, which is not the case for the NPS. Those two signatures are more described in the following sections.

3.1.1 Neurologic Pain Signature (NPS)

The neurologic pain signature, or NPS, is an fMRI-based measure which predicts the pain intensity at the level of the individual person. To develop this signature, four studies have been conducted, consisting in a number of 114 patients.

It should be highlighted that until now, pain assessment is mostly driven by the means of self-report, i.e. an estimation of the level of pain given by the participant after each single-trial test, within some scale. The problem with this method is that it is highly subjective; indeed, two persons might not be sensitive to pain in the same way, and one same pain test could lead to a quite high assessed level for one person and a lower one for the other. Knowing that, it is easy to understand the difficulties of assessing pain in vulnerable populations, such as small children, people that are minimally conscious or people subject to some disease that affects cognitive functions.
The idea behind this study is to combine fMRI measures with machine learning to develop a brain-based neurologic signature, which could then provide direct measurements of pain intensity, and be used to study and compare analgesic treatments. This signature is developed by the means of experimental measures of thermal pain.

3.1.1.1 Study design

The whole work was divided into 4 studies. Each of the studies was characterized by a certain number of randomized sequences of thermal stimuli of varying intensity, applied on the left forearm of each participant. The participants groups were as mixed as possible, comporting a certain proportion of men and women, and within a quite large range of age. This is an important factor to take into account to be certain not to take brain activation related to subgroups of patients as an activation to pain.

In study 1, involving 20 subjects, 12 trials of pain stimuli were applied corresponding to four intensities which were calibrated individually for each participant, since their sensitivity to pain might vary (something that would be qualified as painful for one person may not be the case for the next one). The four intensities corresponded to one non-painful warmth and three levels of painful heat.

In study 2, involving 33 subjects, 75 trials were applied. The different stimuli levels consisted in six temperatures, and each application of a stimulus was followed by an assessment of the subject on whether the stimulus was painful or not. The ratings given by the patient were coded between 0 and 100 if the event was non-painful and between 100 and 200 if it was.

In study 3, involving 40 subjects, 32 trials were applied, consisting of 8 trials with each of four stimulus types. Study 3 was used to detect if 'social pain' had the same neurophysiological signature as the noxious pain. In this idea, the subjects were chosen such as they had recently been through a romantic breakup, and continued to feel rejected. The four stimuli consisted in 1) a noxious (i.e. painful) heat, 2) a warmth that was near the pain threshold, 3) the viewing of a picture of their ex-partner and 4) the viewing of a picture of a close friend. The two first stimuli were again calibrated for each participant.

Study 4 was conducted in the idea of assessing the effects of remifentanil on pain. Remifentanil is a potent ultra-short acting synthetic opioid analgesic drug [11]. The subjects, in the number of 21, went through two series of trials (36 in total), where they each received the drug. In the first series of trials, they were aware that they received the remifatril, and in the second one, they were told that no drug was delivered, even if it had been administrated. Once again, the right drug amount delivered to each participant was previously calibrated individually. There were 36 trials: 18 of them were characterized by a painful stimulus and the other 18, a non-painful warmth. The design was such as the drug concentration changed over time during each infusion series.

3.1.1.2 Methods

Study 1 was used to derive the signature. For this purpose, they used a machine-learning-based regression technique called LASSO-PCR (least absolute shrinkage and selection operator-regularized principal component regression) to predict the outcomes, i.e. the pain reports, from the fMRI activity. The signature development was characterized by five steps:

- **Feature selection**: they selected the voxels ($3mm^3$ measurement) within an a priori mask of pain-related brain regions using previous litterature, more specifically the NeuroSynth meta-analytic database (`www.neurosynth.org`). This was applied in order to detect the relevant brain areas involved with pain.
- **Data averaging**: the data collected during pain sensation from each voxel (inside the mask) were averaged within each stimulus intensity, which leads to the generation of four distinct pain-related activation maps for each subject.
- **Machine learning**: the LASSO-PCR method, a linear algorithm, was run using the four maps developed here above to predict the pain ratings. The activation maps from each condition within participants were used as the predictor, and the average pain reports from each condition within participants were used as the outcomes. The researchers used leave-one-out cross-validation to estimate the prediction error on new trials.
- **Bootstrapping**: this technique was used to determine P-values for voxels weights to find a threshold for the signature weights, for display and interpretation. They constructed 5 000 bootstrap samples, with replacement, that consisted in corresponding brain and outcome data and ran the LASSO-PCR technique on each.
- **Permutation tests**: they were used to guarantee that the procedure is unbiased. They permuted the data 5 000 times, and for each permuted dataset, the whole cross-validate LASSO-PCR technique was applied. An unbiased procedure should lead to a correlation between predicted and observed pain rating which is symmetrically distributed around 0. The permutations tests allowed to confirm this theory.

The study 1 subjects were used as a training set for the development of the NPS pattern, and the remaining subjects (from studies 2, 3 and 4) were used as a testing set, in order to validate the results determined in study 1.

Study 2 was used, as said right above, to test the neurologic signature developed in study 1. No further model fitting was applied, and the data came from a different scanner. The goal is to predict pain in individual participants, using the signature.

In study 3, they simply applied the signature to the activation maps resulting from the 4 trials types (painful stimulus, non-painful warmth, ex-partner pictures and friend pictures).

In study 4, they tested the effects of stimulus intensity (painful vs warmth), the administration of the remifentanil drug and the way the drug was administrated (when the participant is notified and when they are not) on the signature response.

3.1.1.3 Results

The results from study 1 led to a signature pattern with positive weights in regions including the bilateral dorsal posterior insula, the secondary somatosensory cortex, the anterior insula, the ventrolateral and medial thalamus, the hypothalamus and the dorsal anterior cingulate cortex. These

activated regions are shown in Figure 3.1.
The discovery of the activation of various brain regions, not necessarily correlated with each other in other tasks, confirms that pain is a distributed process within the brain. The signature response

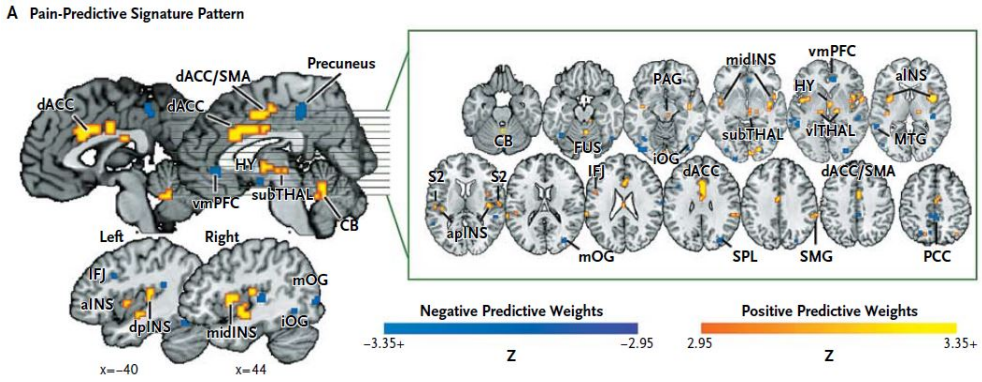


Figure 3.1: Pain-predictive signature pattern developed in study 1 by Wager and al. [10]

increased in a non linear way when increasing the stimulus intensity during thermal stimulation. The discrimination between painful stimulus and non-painful warmth conditions, which were chosen so that the intensity level is above and under a specific temperature (45° Celsius) known to activate specific nociceptors, gave quite high specificity and sensitivity, around 94%.
Studies 2, 3 and 4 have been conducted in the idea of testing the NPS signature developed in study 1, and they all delivered reliable results.
In study 2, they found that the signature increases monotically, and non-linearly, across the six temperatures tested. It showed a good correlation with the reported level of pain and the stimulus temperature. However, it has been found that the signature strongly predicted pain intensity (based on pain intensity ratings), but only weakly predicted warmth intensity (based on warmth intensity ratings) during trials involving nonpainful heat, as shown in Figure 3.2. With these results, it shows that the signature is mainly associated with the sensation of pain, but in some way also reflects the somatic stimulation.

In study 3, it has been showed that the signature response was significally stronger for physical pain than for the other type of pain under study, i.e. social pain. It can be seen in Figure 3.3 that the signature response is higher than the response for the three other conditions: warmth, view of rejecter, view of friend. This is consistent with the notion that different groups of neurons code for different affective events.

Study 4 showed that the signature response decreased with the increasing drug affect-size concentration, both in the open and hidden infusions.

3.1.2 Stimulus Intensity Independent Pain Signature (SIIPS1)

The idea behind the development of this signature was that the brain processes related to pain are multiple and not only due to nociceptive input, as explained in section 2.1.5. Indeed, the cerebral processus mediate psychological and behavioural influences that come along with pain sensation. The problem is that the cerebral contributions beyond nociception remain poorly understood nowadays. In the study presented in this section, the CAN Lab team have developed a pain signature that predicts pain above and beyond nociceptive input by performing a multivariate pattern analysis on the data [12].
Multivariate pattern analysis (MVPA), as opposition to the univariate pattern analysis used, for example, in the General Linear Model (GLM), is a technique that has proven to be more sensitive

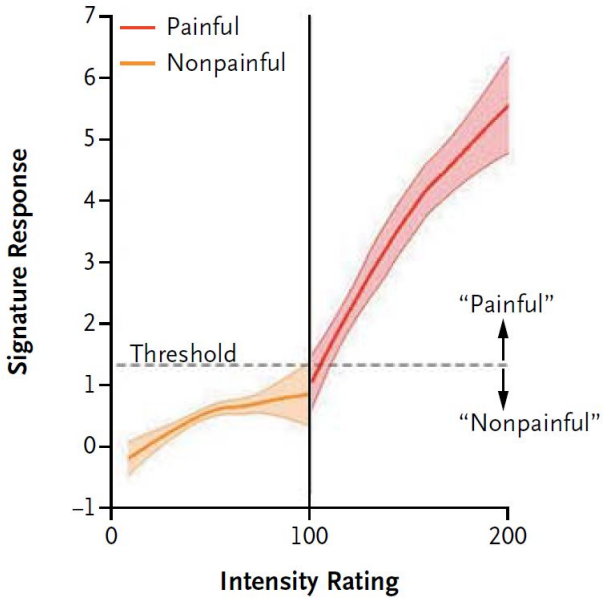


Figure 3.2: Signature response according to reported intensity [10]

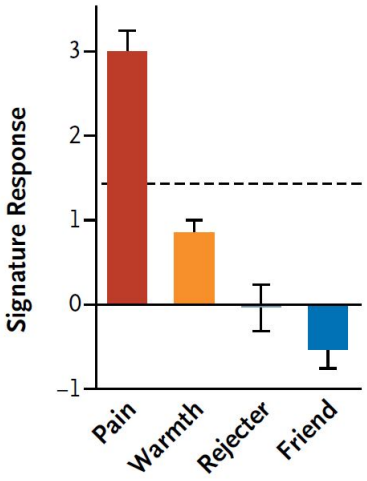


Figure 3.3: Signature response for different conditions [10]

and more informative about the functional organization of the brain. Univariate measurements indicate the global activation of a cortical field or system, while multivariate techniques analyze neural responses as patterns of activity, which in turn gives an idea of the different brain states a cortical field or system can produce. It allows to gain information about the way brain areas encode different types of information. There exist several methods using multivariate pattern analysis. One drawback to MVPA is its unfamiliarity and perceived complexity, which is why it took some time for researchers to adopt the method. A different model of cortical organization had to be considered [13].
The signature was developed using four datasets as a training set, and two as independent test datasets, with cross-validation. It is called "stimulus intensity independent pain signature-1", or "SIIPS1", and its responses allow to interpret the variation in trial-by-trial pain ratings, which was not possible with the previously developed pain fMRI-based marker, usually focusing on nociception, such as NPS. Also, it is shown that the SIIPS1 responses arbitrate the pain-modulating effects of three psychological manipulations of expectations and perceived control [12].

The brain areas that have been considered as non-nociceptive are the following: the dorsolateral prefrontal cortex (dlPFC), the hippocampus, the ventromedial prefrontal cortex (vmPFC), the nucleus accumbens (NAc). These regions are thought to have two roles, either by modulating activity in nociceptive circuits but also by directly influencing pain experience independently of these circuits. For example, it has been shown that chronic pain pattern of activity involves a shift from brain regions related to nociception to a type of activation mostly present in frontal-limbic networks[12].

The goal of the development of the SIIPS1 signature is to be able to identify a multivariate pattern of fMRI activity which is able to predict pain unrelated to nociception, i.e. after removing the effects of noxious stimulus and the NPS presented here above. Also, it would be interesting to discover which brain areas are involved in this process. Finally, the purpose of this study is to investigate whether a model using independent contributions from non-nociceptive brain regions would predict pain better or worst than one using only the noxious stimulus-encoding regions, and also to explore the effects of psychological interventions such as expectancy and perceived control with this model [12].

3.1.2.1 Study design

The SIIPS1 signature development was realized on 4 separated studies (Studies 1-4), with 137 participants in total which pooled a number of 6 740 images (about 50 trial-level images per person) studied in the form of single-trial estimates of brain responses during individual epochs of noxious heat.

Since the SIIPS1 signature is likely to track pain contributions from endogenous cerebral processes (such as the ones presented in the introduction of the SIIPS1 paper), it seemed interesting to investigate whether or not the signature was sensitive to pain modulation induced by psychological interventions. The mediation analyses were conducted on two datasets (Studies 5-6) to test whether the SIIPS1, the NPS or both were able to mediate effects of expectancy and perceived control. More precisely, Study 5, including 17 participants, was dedicated to examine expectancy effects. Study 6 , with 29 participants, helped studying psychological manipulations in 2×2 factorial design. Basically, it allowed participants to have some degree of control in the noxious stimulus delivered. This design is quite complex and surpasses the topics wished to be highlighted for a good understanding of the rest of the project, and is thus no further described.

3.1.2.2 Methods

For the signature development, the first step was to remove the effects of stimulus intensity (by helding the stimulation constant) and the NPS response from each participant's single trial-level brain images. This was performed with a set of regressors modeling all possible differences among intensities. After that, a Principal Component Regression (PCR) technique was used to estimate a multivariate fMRI pattern able to predict the residual pain ratings. The correlation between prediction and outcome for each individual was evaluated through ten-fold cross-validation. Then, a population-level pattern was constructed from a weighted average of the 137 participants' predictive maps (prediction-outcome correlations were taken as a weight). Finally, the brain regions with consistent contributions among all participants were identified by the means of weighted t-tests.

Study 5, investigating expectancy effects, was characterized with auditory cues: participants were told than one cue was predictive of high pain and one other one of low pain. In the training phase, the stimuli were in accordance with the prior cue. High- and low- intensity levels were calibrated for each subject. In the testing phase, on another portion of the skin, the 'high pain' cues were followed by either high- or medium- intensity noxious stimulations (50% each) and the 'low pain' ones were followed by low- or medium- intensity stimulations (50% each).

3.1.2.3 Results

Results showed that there are actually brain systems that contribute to pain beyond nociception, by showing that participants shared activation in those areas. The areas could be subdivided into three groups:

- Positive pain-predictive weights in targets of nociceptive afferents: The brain areas in question are the insula, the cingulate cortex and the thalamus. They showed positive weights in the SIIPS1 after the noxious stimulus intensity and the NPS response effects had been previously removed.
- Positive pain-predictive weights in other regions than the ones targeted by the nociceptive afferents: Those regions are not known to be related to nociceptive information. They are thus likely to form extra-nociceptive contributions to pain. The regions in question are the dorsomedial prefrontal cortex (dmPFC), the middle temporal gyrus, the caudate and the ventrolateral prefrontal cortex. These regions are most probably involved in the mediation of internal thought processes that increase pain independently of nociceptive inputs.
- Negative pain-predictive weights (i.e. a higher activation in those regions corresponds to a less intense pain sensation): The areas in question are the vmPFC, the NAc, the parahippocampal cortex, the posterior dlPFC and others. Once again, these regions make extra-nociceptive contributions to pain. Those regions are thought to contribute to cognitive, evaluative or motivational aspects of pain instead of sensory ones. It is also suggested that they play a role in chronic pain.

All of these regions appear in Figure 3.4.

With these results, it was possible to evaluate the joint contributions of both signatures, NPS and SIIPS1, in predicting trial-by-trial pain ratings. A multilevel GLM was used to quantify the unique and shared contributions from the signatures. The tests were initially performed on the train dataset used for the SIIPS1 development, using leave-one-out cross-validation. Then, the same analyses were

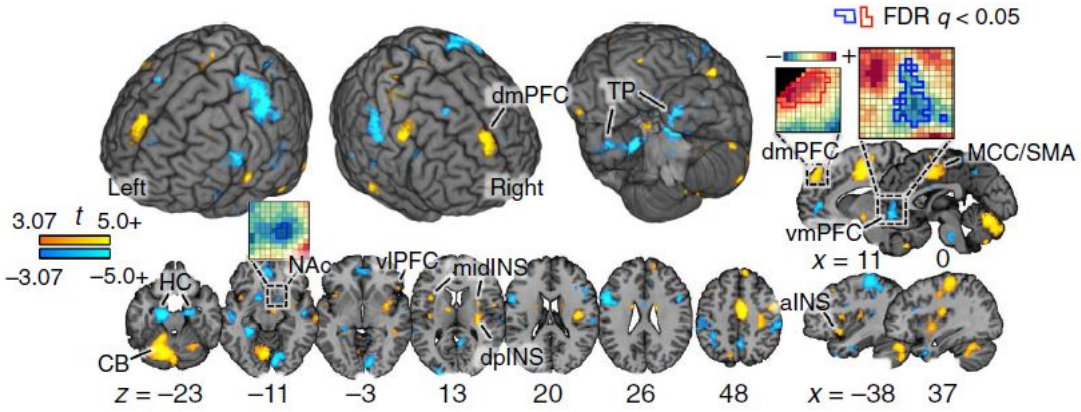


Figure 3.4: Multivariate fMRI signature [12]

conducted on two unknown datasets (Studies 5-6) gathered as a testing dataset. The results of this step showed that the SIIPS1 and the NPS each make unique, significant contributions to predicting pain on individual trials [12].

In Study 5, investigating expectancy effects by the means of "high pain" and "low pain" cues, an analysis of the medium-intensity noxious stimulations was conducted, and it showed that the cues strongly biased pain ratings towards the cued values.
Finally, mediation analyses allowed to prove that the SIIPS1 signature was able to partially mediate the effects of the three psychological manipulations on pain tested in this study, while the NPS signature showed limited evidence for mediation of these effects.

One section of this pain stratification project is dedicated to the responses people have with those two types of signatures. The goal is to show that there exist four categories of subjects, depending on their SIIPS1 and NPS responses: people having low responses for both patterns, people having high responses for both patterns, and people having higher responses for one pattern than the other.

Chapter 4

Data acquisition and analysis

Contents

Before arriving to the point where the clustering is practicable, it seems important to describe the type of data we are working on, and which steps have been applied to acquire them. This section first describes what has been done before the project, the pre-processing techniques and the previous analysis that were performed on the raw data. Then, one section is dedicated to a description of the data available for the project, which results from the previous steps described. Finally, some of the steps that had to be added prior to clustering are explained. The starting point of this project begins right after the description of the data available.

In each study, the measurements are taken as functional MRI data, which operation has been described in the Introduction chapter. The images were taken with a Siemens Tim Trio 3T MRI scanner in the Intermountain Neuro-imaging Consortium facility at the University of Colorado in Boulder [16]. The thermal stimuli were induced with a 16 mm Peltier thermode end plate.

Since the number of studies is quite high, and the description of the procedure is not indispensable for the understanding of the work performed, only one study was chosen to illustrate the steps performed prior to clustering. Except for the study design, the stages of pre-processing and analysis are approximately the same for every study. For a description of each study separately, see Table

1.1 for references to the written articles linked to them. The study presented here is "BMRK4", see Ref. [16].

4.1 Study design

This study includes 28 participants. For each of them, the scanning session was separated into eleven runs and lasted about one hour. The thermal stimulations were separated into three levels, corresponding to three temperatures. Before each stimulus, a cue was given to the patient. This cue could also be of three types, corresponding to the three levels of stimulations. The experimental was 3×3, meaning that the actual stimulations and the cues were completely crossed with each other, i.e. a certain cue was not necessarily linked to the corresponding stimulation level. Among the eleven runs, the first two were used as a conditioning task, where the participant learned the association between the cue and the level of stimulation. Those two runs were not used in the further analysis. The remaining nine runs were dedicated to the experimental task. Each run consisted in 9 trials, pooling a number of 81 single-trials per session, and thus, per participant. The participants performed the same actions for all the runs without being aware of the different types of runs. After each cue-stimulation pair, the participant was asked to make a rating about the sensation they felt, on a visual analog scale [16].

4.2 Data acquisition

The data were acquired as structural images using high-resolution T1 spoiled gradient recall images (SPGR) for anatomical localization, and were then warped to the MNI space (Montréal Neurological Institute). Functional images were acquired with an echo-planar imaging sequence (TR = 1300 ms, TE = 25 ms, field of view = 220 mm, 64x64 matrix, 3.4 x 3.4 x 3.4 mm3 voxels, 26 interleaved slices with ascending acquisition, parallel imaging with an iPAT acceleration of 2) [16].

4.3 Pre-processing

All the pre-processing steps were performed with SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK), for this particular study. The other studies might have used different versions of the software. SPM, or Statistical Parametric Mapping, is a software used for the analysis of brain images data sequences. It uses tools for testing hypotheses about functional imaging data, by constructing and assessing different statistical processes [50].

4.3.1 Suppression of outliers

Before implementing the other pre-processing steps, the outliers were previously removed from the dataset. Removing outliers is an important step, leaving them would lead to distortions in the estimation of population parameters [47]. In each image, for all slices, the mean and standard deviation (across voxels) of values were computed in order to detect the "spikes" in the signal, corresponding to the global outlier time points. Mahalanobis distance was the metric chosen to detect those outliers, it was computed for the matrix of mean and standard deviation per slice. The distribution of Mahalanobis distances for a multivariate normal distribution is chi-square with p degrees of freedom, p being the dimensionality of the data, the number of variables. The Mahalanobis distances are ordered from lowest to highest, and plotted against their corresponding chi-square values [48, 49]. Here, any value with a significant χ^2 value was identified as outlier [16].

4.3.2 Co-registration

For each participant, a structural T1-weighted image was computed, from all imaging session. Those were used to localize anatomical regions. They were then co-registered to the first functional image for each participant, using once again the SPM8 software, with an automated registration iterative procedure using mutual information. Co-registration is a technique used in neuroimaging to combine information from different modalities or at different time points. Here, we are in the case of different modalities: structural images are co-registered with functional ones. The goal is to find the transformation T that maps 3D points from one image onto the anatomically corresponding point in the other. It helps compensating for variation in patient's position inside the scanner, or for the scan plane selection. When the transformation is known, the second image can be resampled so that it has the same size and dimensions of the first image. Thereafter, the anatomical points stand in the same locations, i.e. same voxels, in both images. The metric used for image registration, in this case, is the mutual information. The goal is to maximize this mutual information, corresponding to the information that one image contains about the other, which is based on measures of entropy (marginal entropy for each image and joint entropy between them) [51]. A manual adjustment was added to the SPM8 automated registration process. Then, the structural images were normalized in the MNI space, once again using SPM8, interpolated to $2\times2\times2$ mm^3 voxels [16].

The functional images, after correction (see next section), were wrapped into a normative atlas from the SPM8 software.

4.3.3 Correction

The functional images were corrected for slice-acquisition-timing and motion using SPM8 [16]. Functional imaging data are usually acquired using sequential 2D imaging techniques. Thus, there is a certain time delay between individual slices, and the full 3D volume constructed with those slices might show a significant shift between the expected and actual measured hemodynamic response following a stimulus. This effect might lead to a decreased sensitivity to detect activations. Slice-timing correction is the solution to this problem, it works by temporally realigning individual slices to a reference slice, based on their relative timing. The resampling methods are data interpolation methods such as linear, sinc or cubic spline interpolation [52].
Motion correction is applied to correct the subject's motion in the scanner, during the experiment. It is usually modeled by a time series of 3D rigid body transformations. Those transformations align each volume of the time series to a reference volume, by maximizing some similarity measure such as the mutual information mentioned above [53].

4.3.4 Smoothing

The smoothing step is performed on the normalized images. It is performed by convolving a 3D Gaussian function, more precisely, a full-width-at-half-maximum (FWHM) Gaussian kernel. This step leads to a higher signal-to-noise ratio, although the resulting images are more blurred. It also allows inter-subject averaging and increases the validity of SPM [16, 2].

4.4 Previous analysis

4.4.1 First-level analysis

First-level general linear model (GLM) analysis were conducted with SPM8. The theory behind the GLM outreaches the content of this project. For a complete review, see Ref. [54]: *Generalized linear models and extensions*, from JW Hardin and JM Hilbe. The GLM used in this case was

based on boxcar regressors, convolved with canonical hemodynamic response function, constructed to model the periods of thermal stimulation and pain rating. This procedure is briefly explained in section 2.2.3. To the GLM, other regressors are added as nuisance variables, such as the intercepts for each run (called "dummy" variables), the linear drift over time within each run described in the section 2.2.2, the six estimated head movement parameters (the three spatial directions, roll, pitch and yaw), their mean-centered squares, their derivatives and squared derivatives for each run (24 columns in total), indicator vectors for the outliers that have been identified, indicator vector for the first two images of each run, and signals from white matter and ventricule [15, 16].

4.4.2 Single-trial analysis

For the single-trial analysis, another GLM was constructed, with a design matrix with one regressor for each trial. The model was used to compute the single-trial response magnitudes. As in the first-level model, boxcar regressors convolved with canonical hemodynamic response function were built to model the periods of stimulation and rating. Trial-specific regressors were added, one for each single trial. Nuisance variables such as the ones mentioned above were also added [15].

4.5 Data available for the project

At the end of all these steps, we end up with the data on which the rest of the work is done. At this point, it should be noted that the previous steps (pre-processing and previous analysis) were already implemented and there was no way to modify them. The data available at the beginning of the project were the single trials data for each participant of the thirteen studies. The last steps, i.e. normalization and predictions, were performed using CAN Lab or Matlab tools.

The data is presented as a *fmri_data* object in Matlab, which belongs to a class developed by the researchers from the CAN Lab team. This object contains several fields. The ones that were used for the rest of the study are resumed in Table 4.1. Each study is composed of several par-

	Size	**Content**
.dat	$nb\_voxels \times nb\_single\_trials$	Parameter estimates for each voxel
.Y	$nb\_single\_trials \times 1$	Pain ratings
.images_per_session	$nb\_participants \times nb\_images$	Number of images taken for each participant
.additional_info	$nb\_participants \times nb\_images$	Pain levels

Table 4.1: *fMRI_data* object characteristics

ticipants, usually between 15 and 40. For each of them, there is a certain number of single-trial maps, corresponding to the number of images taken during the session. This number is usually contained between 50 and 80. The single-trial map contains a parameter estimate for each voxel in the image. Parameter estimates are the coefficients resulting from the single-trial analysis performed with GLM, mentioned in section 4.4.2. They result from the analysis of time series for each voxel by fitting an experimental design matrix such as the predicted response described in section 2.2.3. They can be viewed as a scaling coefficient of each regressor of the experimental data (i.e. each voxel) which minimizes the distance between the actual data and the constructed model. In the following equation, the coefficients composing the `.dat` field are the β:

$$y = X_{design}\beta + \epsilon \tag{4.1}$$

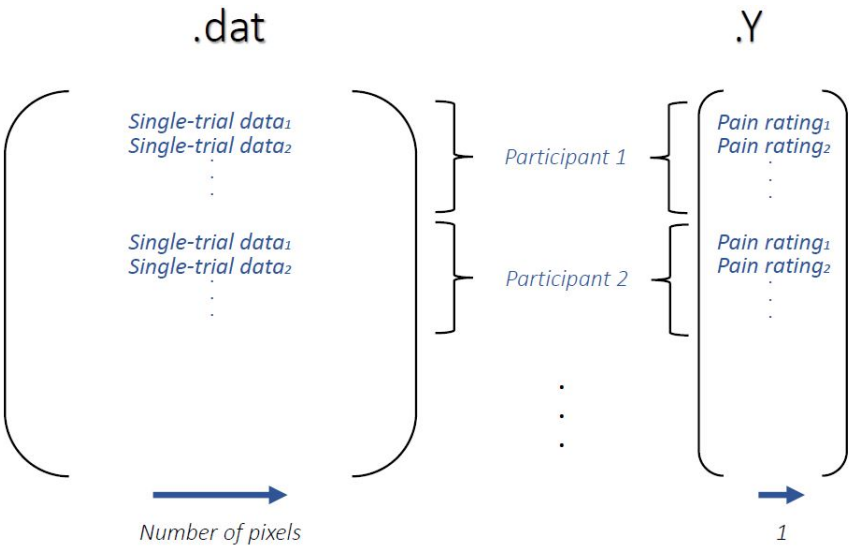


Figure 4.1: Representation of the data arrangement in a `fmri_data` object

Where y are the time series from each voxel, X_{design} is the experimental design matrix and ϵ are the residuals [21].

The single-trial maps are associated with a random pain level induced as a thermal stimulus. The type and number of pain levels may vary from one study to the other. For example, it might consist of three levels, corresponding to three different temperatures. It is important that this pain level is randomly chosen, and the order of pain stimuli must vary from one participant to the other. If this is not respected, one might find some kind of brain activity which is not related to the task under study.
Note that for certain studies, the number of single-trials data is the same among all participants, and sometimes it is not. That is why it was important to use the `.images_per_session` information to prepare the data for the clustering step.
Each single-trial image is also linked with a pain rating, which is an estimation of the pain felt by the participant after each trial, on a scale from x to y. One complication is that each researcher uses its own scale values and it may vary a lot from one study to the other. Some people use a 1 to 10 scale while others use a 1 to 100 one. In addition, different people might evaluate a same amount of pain with very different scales. Since the entire stratification process is based on the pain ratings given by the participants, the clusters finally found could reflect different behaviors of assessing pain instead of actual different activation patterns. Indeed, in this situation, one cluster wcuild contain the subjects which gave a high rate for the pain felt, and another cluster would contain the subject corresponding to lower pain ratings, even if the pain is in fact the same. Both of these scaling issues are overcome by the means of the normalization step explained here under.

The data is ordered in the following way: in each of the matrices mentioned in Table 4.1, the participants, along with their respective number of single-trials data, are stacked one after the other. A visual representation of the arrangement of the data in the different matrices is shown in Figure 4.1.

4.6 Data standardization

A normalization step is usually recommended in order to make variables comparable to each other. Indeed, every person has a unique brain and the same stimulus may induce different responses in different people. Sometimes, the same region is activated during a stimulus (here: pain) in two different subjects, but one shows higher activation rate than the other. That is why it is necessary to reduce the responses to the same scale, in order no to have higher contributions from people who show higher amplitude of activation. Doing that, we can compare them in a more accurate way [1, 23].
There are two different ways of scaling the data: normalization and standardization, which are quite different. After a normalization step, the data are scaled so that their value is comprised between 0 and 1. A possible way to do that is to compute the following formula on every data points:

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4.2}$$

On the other hand, a standardization step leads to a dataset that has zero mean and unit variance. This can be achieved with the following formula:

$$x_{new} = \frac{x - \mu}{\sigma} \tag{4.3}$$

Where μ is a mean value and σ a standard deviation [24].
There exist three different ways of performing this standardization step, depending on the mode used to compute the mean and standard deviation of the previous equation. Indeed, it is possible to standardize the data along the whole dataset, but also along the rows or the columns of the data (if the data is represented as a matrix). In our pain study case, the rows standardization would correspond to a standardization within subject, while the columns one would give a standardization within pixel.
The standardization step is chosen for this study, and it is preferred to compute the new values with the mean and standard deviation among the whole dataset. This is performed on the parameters estimates matrix `.dat` from now on referred as X, to remove the effects of differences in activation amplitude among subjects, and on the outcomes, the pain ratings, Y, to remove the effects of subjective experience mentioned in the previous section.

4.7 Machine learning for predictions

There are several ways to determine whether the clusters exist or not, and whether they are reliable. The data available may be employed in various forms, and by applying a clustering algorithm we can find any number of groups in each of them. In any kind of experimental neuroimaging study, different goals usually require specific methods for the data analysis, as they usually need to focus on specific processes [29]. Here, the real challenge is to test whether the clusters mean something or not. And it would be interesting to test whether they appear in different data analysis methods.

To be able to make those clusters appear, the first step is to construct pain-predictive weight maps. The data collected for the project constitute a "curse-of-dimensionality" or "small-n-large-p" problem [40], meaning that the number of parameters, i.e. the parameter estimates for each voxel, exceeds the number of observations in several orders of magnitude (about 300 000 voxels for approximately 80 observations per subject). Thus, it is important to construct weight maps, which captures the information of interest within the brain and allows to represent each participant as one vector, which constitutes a linear combination of all the contributions [31].

The construction of weight maps is done using standard machine learning techniques. Due to the continuous nature of the outcomes to be predicted, i.e. the pain ratings given by the participants, machine learning tools must be for regression, not classification [40]. Supervised machine learning tools would be used in this case, since we want to approximate the outputs using the inputs. On the other hand, the discoveries of clusters will be performed using unsupervised machine learning tools, such as k-means and k-medoids, as it will be showed in the results chapter. Unsupervised machine learning tools are used to find regularities within the data points without any information about the inputs or outputs [33]. In supervised machine learning, predictive modeling is accomplished by minimizing the error of a model and making predictions as accurate as possible. Initially, linear regression is a tool for statistical analysis used to understand the relationship between input and output variables, but it is not the first time that the machine learning field borrows tools from other fields and uses them towards its ends. A linear regression model tries to estimate the outputs using the following formula:

$$\hat{y}(o) = w_0 + \sum_{i=1}^{n} w_i a_i(o) \tag{4.4}$$

Where o is the observation, $\hat{y}(o)$ is the predicted outcome, $a_i(o)$ is the ith coefficient of observation o and w_i are the model parameters. The goal is to find the right model parameters, the ones that fit well the learning sample available and have a good generalization to unseen observations [34, 41]. The way this is accomplished is explained in the next two sections.

The weight maps are created by constructing a predictive model, such as the one presented above, using the set of observations we have access to. With this model, we can in turn estimate pain report from brain activation for each subject. The resulting equation is the following:

$$\hat{Y} = \boldsymbol{w}^T \boldsymbol{x} \tag{4.5}$$

Where $\hat{Y}$ is the predicted outcome, $\boldsymbol{x}$ is the corresponding set of features and $\boldsymbol{w}$ is the brain weight found using machine learning techniques [22]. In this project, the clustering is done on the weight maps $\boldsymbol{w}$, one for each participant of the study, obtained with this method, either directly or after some other dimensionality reduction technique (see section 5.1). Weight maps may be built for different purposes, and thus with different analysis methods. Two different manners have been implemented in this project: multivariate and univariate methods.

Multivariate and univariate approaches constitute the basis for the following study. Both represent a way of constructing the pain predictive weight maps for each subject, on which the clustering will be applied, directly or not. For a long period, mass-univariate methods were the only ones used to analyze neuroimaging data, although medical images are multivariate by nature. For example, in a fMRI study of pain related activity, an univariate analysis method will associate each individual voxel to a target variable, which is represented by the pain stimulus in our case, considered as a model for brain activation [29]. Statistical analyses used in fMRI studies are applied on each individual voxel, for example with a GLM. It should be noted that some local spatial dependencies are actually considered in the data analysis due to the previous smoothing (pre-processing step) and the spatial spread of the hemodynamic response. Univariate methods were not able to describe how specific and individual cognitive information was encoded in the neurons. Multivariate approaches were developed to overcome these limitations. For example, in a study revising memory, multivariate methods have been able to show that the lateral prefrontal brain regions had an increased activity when switching from higher to lower working memory load, meaning that prefrontal activity illustrates the storage of working memory contents with a certain delay [30]. Multivariate approaches combine information coming from several channels, i.e. several voxels. Univariate methods mostly bring information about interpretability of the brain regions, frequencies or time intervals reflecting

a specific cognitive process. On the contrary, multivariate methods allow to estimate/decode brain states from neuroimaging data, and vice versa. In this case, the interpretability of the model parameters is less important [29].

For this project, pain-predictive weight maps were built in a similar way using univariate or multivariate approaches. As mentioned in the previous section, the single-trials data for each individual are stacked upon each other as a list. The first step to implement was to separate those data, so that we can perform the machine learning algorithm for each individual, based on their own single-trials data. The methods differ for the construction of univariate or multivariate maps. An Ordinary Least-Square (OLS) regression was used for the first one, and a Principal Component Regression (PCR) for the second. Both of them lead to the creation of weight vectors, one for each subject, in the shape of $(1 \times nb\_voxels)$, corresponding to a certain linear combination of the components voxels weights. Each of the thirteen studies was analyzed in the same way and, once again, all the weight vectors (from all the participants in the thirteen studies) were stacked upon each other. At the end of the process, the weight matrix has a shape of $(nb\_participants \times nb\_voxels)$, where each row corresponds to one subject. Several clustering methods are then applied on the different weights matrices, directly or not. This will be developed in the results chapter.

4.7.1 Univariate approach: ordinary least squares regression

The pain-predictive weight maps constructed using univariate methods will be from now on referred as "univariate maps".
Different methods could be employed to build those maps, the one chosen here is an Ordinary Least Squares (OLS) regression, which is the most common method for fitting linear statistical models. The OLS regression model takes the following form, for one observation i:

$$Y_i = w_0 + w_1 X_{i1} + w_2 X_{i2} + ... + w_k X_{ip} + \varepsilon_i \tag{4.6}$$

Where Y_i is the outcome value for observation i, w_0 is the regression constant, X_{ij} is the score for observation i at the jth predictor variable, w_j is the partial regression weight for predictor j and ε_i is an error term for observation i. This can be written in a matrix form:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{w} + \boldsymbol{\varepsilon} \tag{4.7}$$

With the following structure, for k explanatory variables and n observations:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_k \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ 1 & X_{13} & X_{23} & \dots & X_{k3} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

In our situation, the Y_i correspond to the pain ratings evaluated by the participant under study, i being one single trial, from a number n in total. The X_{ij} correspond to the parameter estimates described in the previous section. Note that the same procedure was used to derive those parameter estimates for each voxel in Eq. (4.1). The β from that equation correspond to the X_{ij} is this one.

The k elements in $\boldsymbol{w}$ are called "regression coefficients", and they bring information about each predictor variable's unique or partial relationship with the outcome variable [32]. In OLS regression, the goal is to estimate those unknown parameters by creating a model that will minimize the

sum of the squared errors between the observed data and the predicted one [42]. The total squared error is computed using the following formula:

$$TSE(LS, \boldsymbol{w}) = \sum_{i=1}^{N} \left(y_i - \boldsymbol{w}^T \boldsymbol{X}_i \right)^2 \tag{4.8}$$

Where LS is the learning sample containing the observations and y_i is the output corresponding to observation i. OLS regression will derive the $\boldsymbol{w}$ which minimizes this statement. An example is shown in Figure 4.2: the goal is to find the linear function of X that minimizes the sum of squared residuals from Y [34, 41].

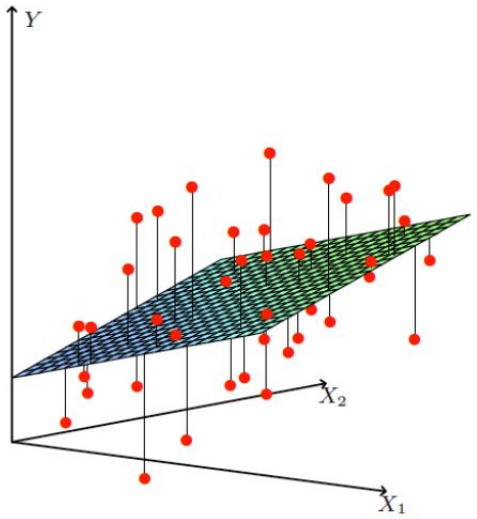


Figure 4.2: Visualization of the OLS regression method [34]

This method is applied to the data available for this project described earlier, and a weight matrix is constructed, containing univariate weight maps for each participant. The weight maps correspond to the regression coefficients. The method has to be applied to each patient individually, in order to obtain a weight map unique for each subject. The income matrix X (`.dat`), containing the parameter estimates for each voxel (columns) for each single trial (rows), is divided so that each participant is described by their own matrix. If the number of single trial for participant i is s, the X matrix for participant i would have the following shape: $(s \times nb\_voxels)$. Then the regression step is applied for each participant separately. It is done with the `regress.m` function developed by the CAN Lab team, which is specifically adapted for fMRI data. In this function, the data parameters forming the X matrix are regressed on the outcomes, i.e. the pain ratings.

4.7.2 Multivariate approach: principal component regression

The pain-predictive weight maps constructed using multivariate methods will be from now on referred as "multivariate maps".
Different methods could be employed to build those maps, the one chosen here is a Principal Component Regression (PCR) technique. It is an interesting technique especially when high multicollinearity exists among the data, effect that occurs when there is an approximate linear relationship among independent variables. When this effect is present, the least squares estimates described earlier are unbiased, but they might have a value quite far from the true value, which means that their variance is increased. Using PCR should allow to reduce this variance and lead to more reliable

estimates, although they would be biased in this case [44]. PCR method combines linear regression with principal component analysis, or PCA. PCA is a technique able to gather highly correlated independent variables into a principal component, and all principal components are independent of each other [63, 64]. This way of proceeding allows to move from a set of correlated variables to a set of non-correlated principal components [43].
The first step for the PCR technique is to standardize the data, which is done by subtracting their means and dividing by their standard deviations. Then, the independent variables have to be transformed into their principal components, using PCA. By doing that, we create new variables that are weighted averages of the original variables X, and since they consist of principal components, the correlation between them is always zero. At this point, we can solve the multicollinearity issue mentioned above, by omitting the components with very small eigenvalues, which correspond to the high multicollinearity. This induces dimensionality reduction by discarding the smallest eigenvalue components. After that, we can regress the outcomes to these components, and the multicollinearity is not an issue anymore. The regression is conducted using the first m dimension reduced principal components (where $p - m$ components have been discarded, p being the number of independent variable). The results are transformed back to the original scale and we find the coefficient estimates in the right scale. As already said, the bias has increased, but it should still lead to better estimations by reducing the variance [44, 45].

This step is implemented for each participant individually, in the same spirit than for the univariate weight maps. The function used was `predict.m` function developed by the CAN Lab team. The previous separation into the different participants was performed just as for the univariate maps. The method is associated with a 5-folds cross-validation, which is a technique used to evaluate and to compare learning algorithms. In cross-validation, the dataset is separated between two sets: a train set used to train the model and a test set used to validate it. The train and test sets must cross-over in successive rounds so that each data point is trained and tested at least once. The accuracy of the model is computed at each of the k rounds and at the end of the process, it will lead to a mean accuracy. This is particularly useful when the number of observations is small, which is the case in this project [46, 34]. For a deeper description of the k-folds cross-validation, see Ref. [46]. In this case, the cross-validation is used with 5 rounds, in order to make a more accurate prediction than without cross-validation.

At the end of the construction of the two types of weight maps, the subjects that contain one or more NaN values in any of those maps are removed from all of them. Indeed, a lot of methods do not function correctly when they have to deal with these values.

Chapter 5

Stratification: Procedure

Contents

For this project, the clustering is performed on several types of data. The clustering algorithm chosen for this project is k-means. A short reminder of its operation appears in a further section. To test whether the clusters obtained are accurate or not, for each subject, the single-trials were separated into two approximately equal groups (depending if the total number is an odd or even number). The idea is to find the clusters on the first half of the single trials for each subject, and then test whether the same participants belong to the same clusters. Once again, the whole technique is described later.

This chapter presents the different methods that were used to determine clusters in the dataset, and the methods used to evaluate the quality of those clusters. The actual results are presented and discussed in the following chapters.

5.1 Data types

As explained in section 4.7, two different machine learning tools were applied to the original data in order to obtain two types of weight maps used for the clustering: multivariate and univariate maps. As a reminder, those maps consist of a number *nb_participants* observations, i.e. 433, for a number *nb_voxels* of variables, i.e. 352 328. The high-dimensional nature of the data evokes that the weights are highly unstable, and the clustering directly applied on those maps is unlikely to give good results. In order to solve that, several techniques can be employed in order to reduce the dimensionality and stabilize the weights at the same time. Note that the stratification on the entire maps is conducted anyway, in order to compare the results with the different dimensionality reduction techniques presented here under.

5.1.1 Application of signatures

A first way of reducing the clustering space dimension is to use predefined patterns such as the ones presented in Chapter 3, i.e. the signatures patterns NPS and SIIPS.
In the study that has developed the NPS pattern presented in section 3.1.1 [10], a pain-predictive weight-map of this pattern was constructed, so that it is possible to "apply" it to other images. This is done by the means of dot products, such as in the following equation:

$$\hat{Y} = \boldsymbol{w_{NPS}}.\boldsymbol{x} \tag{5.1}$$

By applying the pain-predictive NPS weight-map $\boldsymbol{w_{NPS}}$ to some data matrix $\boldsymbol{x}$, we find a prediction $\hat{Y}$, predictive of the pain rating. In the case of the project, this weight maps is applied to our univariate/multivariate weight maps, and a single value per participant is computed [16]. This value reflects the strength of the signature response for that particular participant.
The exact same thing is done with the second signature under study, i.e. SIIPS. We thus find two distinct values for each subject, corresponding to the two signatures, and we can transform the high-dimensional space of the entire maps into a two-dimensional space, with dimensions corresponding to signatures. In such a graph, each subject will be represented as one point in the space. The clustering step is performed on these points.
It should be noted that this procedure leads to much more stable weights compared to the initial weights of the entire weight map, since they are based on predefined patterns. However, using the patterns makes the weights more constrained, while they are much more flexible in the initial weight map. This explains why we would find better results after applying the signature than by directly studying the behaviors of the entire weight maps.

In this specific case, we would want to discover clusters that reflect different types of responses to the two signatures. If we chose to investigate the presence of four clusters, we would want them to separate the individuals in the following way: one cluster would pool the participants that show small responses to both signatures, one would contain the participants with high responses for both, and the last two ones would show participants that have a high response for one signatures and a low one for the other. As it will be seen in the results, the type of clusters depends on the choice of distance metric.

5.1.2 Dimension reduction using PCA

Another idea that comes in mind is to simply reduce the dimensionality using a regular dimension reduction technique. The one chosen for this project is the Principal Component Analysis, or PCA. This technique has already been mentioned in section 4.7.2, explaining the PCR principle. It is a multivariate statistical tool, very useful for the analysis of observations described by dependent, and often inter-correlated, variables. This tool is able to extract important information in the data, and then reduce the dimensionality by expressing this information in the shape of new orthogonal variables, the principal components. Those new variables usually consist of a more coherent representation of the data [63, 64]. The mathematical development of this tool is not presented in this project. For a full review, see Ref. [64]: George H. Dunteman, *Principal Components Analysis*.

It was found that in this case the number of components which explain 95 % of the variance in the initial data is about 250. The first idea was to directly use this number for the complete analysis. However, there is nothing proving that it would lead to the best results. Thus, the optimal number of components to be used in the PCA algorithm is determined by initiating the method several times, each time with a different number of components. The choice is made by the one giving the highest cluster quality value presented in a following section. The tested numbers are ranged between 2 and 300 components.
Then, the clustering algorithm is applied on the scores obtained with PCA, i.e. the representations of the data points in the principal component space.

5.1.3 Canonical networks

The use of canonical networks might also help reducing the dimensionality. In this project, the networks parcellation used is the one developed by Yeo and al. from the Buckner Lab [70]. The idea is that distributed systems of brain areas are responsible for the processing of complex behaviors, and those areas can form either anatomical connectivity, involving regions with a local hierarchical relations, or large-scale connectivity patterns which involve other circuits with no clear hierarchical connection. In this study, 1 000 subjects and their functional connectivity MRI data helped to determine seven networks of functionally connected regions across the cerebral cortex. Those networks can be visualized in Figure 5.1.
In the context of this pain stratification project, the seven-network parcellation was used in order

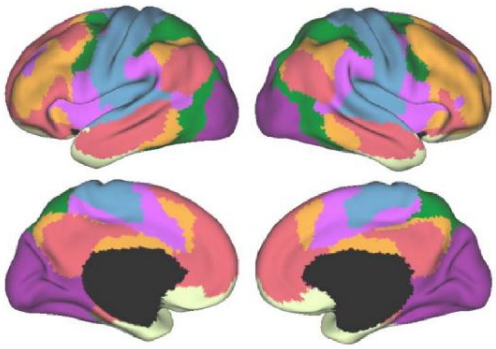


Figure 5.1: Seven network cortical parcellation [70]

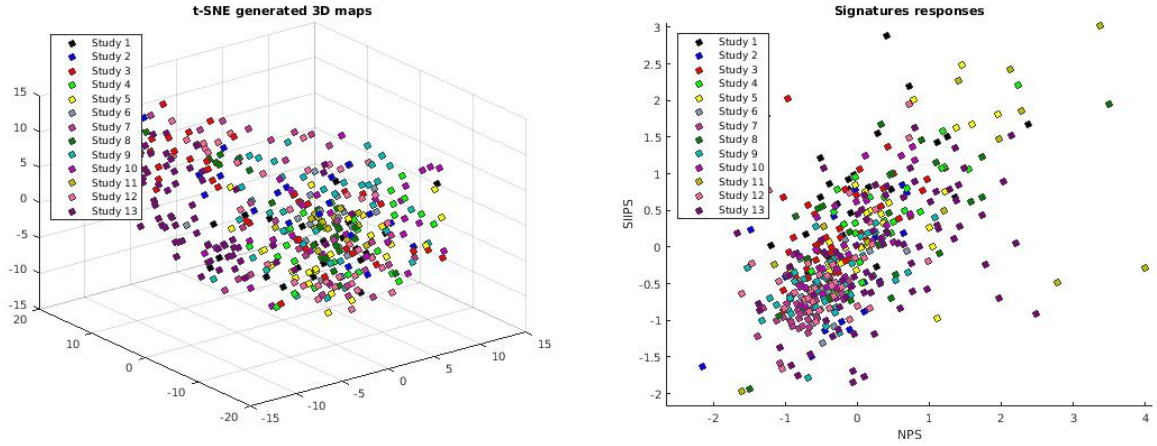
to average the neural activation related to pain sensation within those networks, and reduce the dimensionality of the data. Indeed, this procedure reduces the number of variables to seven, each variable corresponding to the average activation inside each network.

Except for the case of the signatures, because the resulting space is two-dimensional and the results can be visualized, the need for a visualization tool is crucial. Indeed, the only way to visualize the resulting clusters in the case of the entire maps, the PCA reduction and the canonical networks, is to transform the results space in a two or three-dimensional one. This is done with a method called "t-Distributed Stochastic Neighbor Embedding", which is a method based on the computation of similarity between points using stochastic tools [55]. The theory behind the technique is developed in Appendix A.

5.2 Cluster per study

Before considering applying the clustering algorithm on the different data, it was important to verify that the points, corresponding to the subjects, are "sufficiently" spread out around the space we are going to work with. Indeed, we don't want to find a cluster that would correspond to all the participants of one individual study of the dataset. This could mean that, for some reason, the subjects show a similar pattern of neural activation which is probably not only due to pain sensation like the rest of the 433 subjects. The spread out is shown for the whole weight maps (univariate), using the t-SNE method for three dimensions visualization, and after application of the signatures on the weight maps (multivariate), in Figures 5.2a and 5.2b, respectively. The two other cases (signatures on univariate maps and whole multivariate maps) appear in the Appendix C.

In these figures, each point color corresponds to one study.



(a) Univariate maps (visualized in 3D using t-SNE) (b) Multivariate maps (signatures responses)

Figure 5.2: Cluster by study

It can be seen that in the case of the signatures responses (and in the two cases presented in the Appendix), the points are rather well spread out, no cluster is formed with the points originating from the same study. However, in the case of the entire univariate maps, it appears that some points from the same study seem to be grouped in a kind of cloud that could form a cluster in the following study. However, the number of clusters investigated (4) is smaller than the number of studies present in the dataset (13), there is less chance than the clusters found are driven by the studies. One way to prove that would be to conduct an error analysis in the shape of a regression of study differences, but this surpasses the content of the project.

5.3 Clustering algorithm

As already said, the clustering algorithm used in this project is k-means, implemented in the Matlab function `kmeans.m`. This algorithm is used to separate the dataset into several, k to be precise, groups, based on some evaluation criteria. The groups are formed such as the points in a particular subset share similar properties, in our case similar brain activation due to pain, while data in different subsets show different behaviors. Two properties have to be met: high cohesive properties and low coupling property [35]. The number k of clusters must be fixed beforehand.

The k-means' way of working is based on the minimization of the total squared error. The goal is to maximize the intra-cluster similarity while minimizing the inter-cluster similarity [35]. The distance metric is the Euclidean distance, although the `kmeans.m` function from Matlab is equipped with several distance metrics in the `Distance` parameter. The separation into clusters is based on the following equation:

$$\min_{C,\{m_k\}_1^K} \sum_{k=1}^{K} N_k \sum_{C(i)=k} ||x_i - m_k||^2 \tag{5.2}$$

Where C is the cluster, m_k is the cluster mean, N_K is the number of data points belonging to cluster k and x_i is the point on which we are performing the operation.

The algorithm starts by randomly assign each point to a cluster. Then, an iterative procedure is started: first, the cluster means $\{m_1, ..., m_K\}$ are computed, taking into account the current cluster assignments. Then, given the cluster means computed, each observation is assigned to the cluster with the closest cluster mean, with the following equation:

$$C(i) = \arg \min_{1 \leq k \leq K} ||x_i - m_k||^2 \tag{5.3}$$

The process is stopped when the points assignments do not change [33].
In this work, the algorithm is applied to the participants, i.e. the rows of the weight matrix. The clusters then corresponds to subgroups of subjects with similar unique brain activation to pain.

Note that to allow the results to be reproducible, the seeds, i.e. the starting points, used to start the algorithm must be controlled for. This is done using the `rng ('default')` command from Matlab, which generates random seeds but keeps them in memory for future use. This way, we find the same clusters each time the algorithm is launched. Here, the seeds are randomly chosen, but sometimes it is important to choose them manually, for example when the number of clusters to be found is quite high.

Also, it is important to keep in mind that the k-means algorithm converges to a local, and not global, minimal distance. Thus, very different seeds points might lead to very different clusters, such as in the example from Figure 5.3. In this figure, it is clear that the clusters wished to be found are the ones presented in the graph on the left. But it is possible to obtain results such as the graph on the right when the starting points are incorrectly chosen [33]. The solution to this problem is to restart the algorithm several times, and this is done with the `'Replicates'` parameters in `kmeans.m` function from Matlab, which corresponds to the number of times the clustering is repeated, each time using new initial cluster centroid positions. In this work, in both cases, the number of replicates is set to 5.

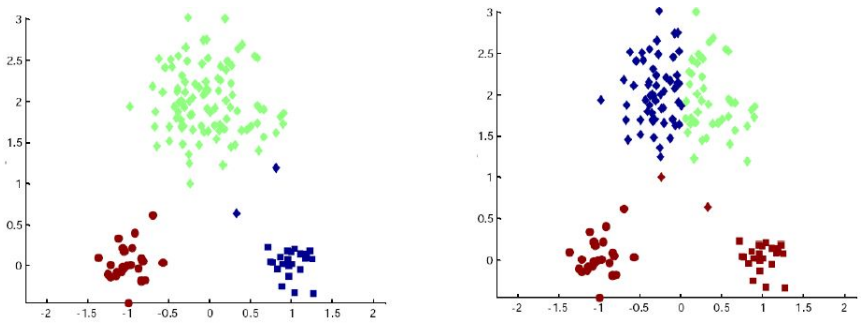


Figure 5.3: Different seeds producing different clusters [33]

The different distance metric parameters linked to the `kmeans.m` function lead to different types of clusters. The choice must be made between a number of metric types, depending on the type of data under study and the shape of the clusters wished to be discovered. In this project, several distance metrics have been tested, such as the squared Euclidean, the Mahalanobis, the correlation, the cityblock and the cosine distances. The choice of distance measure depends a lot on the type of data you want to cluster, and changing the distance parameter may have a great impact on the performance of the algorithm [35] and on the results obtained. The Mahalanobis distance is the distance between two points in multivariate space, and it is based on the variance-covariance matrix between data points [36, 37, 38, 39]. This definition implies that this metric would constitute the right choice for the clustering algorithm, at least for the multivariate maps, but it was put aside for two reasons: first because other metrics gave better results than this one. For example, in the case of the signature patterns, the clusters found with the Mahalanobis distance did not correspond to the four clusters types wished to be found, mentioned at the end of section 3.1.2.3. Secondly, the `kmeans.m` function is not equipped with this parameter. The k-medoids algorithm, which is a variant of the k-means one, had to be used instead, but it did not work on every type of data: the `kmedoids.m` function from Matlab does not support very large matrices, which is the case of the entire weight maps. It did not seem convenient to use different algorithm with different metrics, so the Mahalanobis distance was not an option.
The correlation distance is based on the Pearson's correlation coefficient, it measures how well scattered points from two measurements can be fitted with a straight line. The cityblock distance, or Manhattan distance, is the sum of distance along each dimension [58, 59]. Those two metrics were also tested but as their results did not show interesting behaviors, they are not presented in this report.
The two metrics presented here are the squared Euclidean distance and the cosine distance, which are further explained in the following sections. The results obtained with them are quite different, and it is interesting to compare them to decide which one leads to the most relevant clusters.

5.3.1 Euclidean distance

Theoretically, the Euclidean distance between two points a and b, with k dimensions, is written as follows:

$$d(x,y) = \sqrt{\sum_{j=1}^{k}(x_j - y_j)2} \tag{5.4}$$

One advantage of this technique is that it is not affected by the introduction of new points in the dataset, possibly outliers. In contrary to other distance metrics, the Euclidean distance preserved more information about the data by taking into account the magnitude of the measurements. However, one disadvantage to that is that the distances might be affected by differences in scale among

the dimensions from which the distances are computed [35, 58, 59, 33]. When using this distance metric in the k-means algorithm, each centroid corresponds to the mean of the points belonging to that cluster.

5.3.2 Cosine distance

The cosine distance is based on the cosine similarity measure. In this metric, the points are treated as vectors, and the cosine similarity is a measure which calculates the cosine of the angle between two vectors. It measures the orientation and not the amplitude. The cosine distance is equal to one minus the cosine similarity, it represents the angular distance between two vectors while ignoring their scale. The cosine distance between two vectors x and y can be written as follows:

$$d(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{\boldsymbol{x}\boldsymbol{y}'}{\sqrt{(\boldsymbol{x}\boldsymbol{x}')(\boldsymbol{y}\boldsymbol{y}')}} \tag{5.5}$$

When using this metric with the k-means algorithm, the centroids correspond to the mean of the observations present in their cluster, after normalizing those points to unit Euclidean length [58, 59, 60].

5.4 Objectives

The objectives of this project are of two kind: first, we want to examine the quality of the clusters found. Three measurements of cluster quality are used in this project, as explained in the following section. The second objective has a more neuroscientific aspect, we want to detect the neural activation differences between the clusters, and identify which brain regions are more or less involved in each subtype. This is mainly done by the means of voxelwise thresholded t-tests.

5.4.1 Cluster quality

The cluster quality is evaluated by the means of three methods: the cluster reproducibility/reliability, evaluating the subjects affiliation to clusters based on two distinct datasets, the cluster consistency, based on the Silhouette value, and the clusters correlation, based on a measure of the spatial pattern correlation between subject's weight maps from the same cluster on one side, and from different clusters on the other side.

5.4.1.1 Cluster reproducibility

As a reminder, the k-means algorithm will always locate k clusters, k being specified by the user, even if they do not exist. This is why it is important to verify that those clusters actually mean something, and this is done by identifying whether the subjects belong to the same cluster with two separated datasets.

5.4.1.1.1 Data separation To test whether the clusters are reliable, or reproducible, all the data available for each individual participant (the total number of single trials per subject) is separated into two portions. This is done using odd-even runs, i.e. each group takes one out of two single trials. Indeed, it is a better way to separate the data than just cutting the subject's dataset into two portions and let the first half in the first group and the second half in the second group. By doing that, we could possibly take some side effects as activation, such as head movement, which is usually greater at the end of the study than at the beginning.
We thus end up with two distinct datasets, with approximately the same amount of data, for each of the 433 participants of the study. We apply the whole process on each of the two datasets separately.

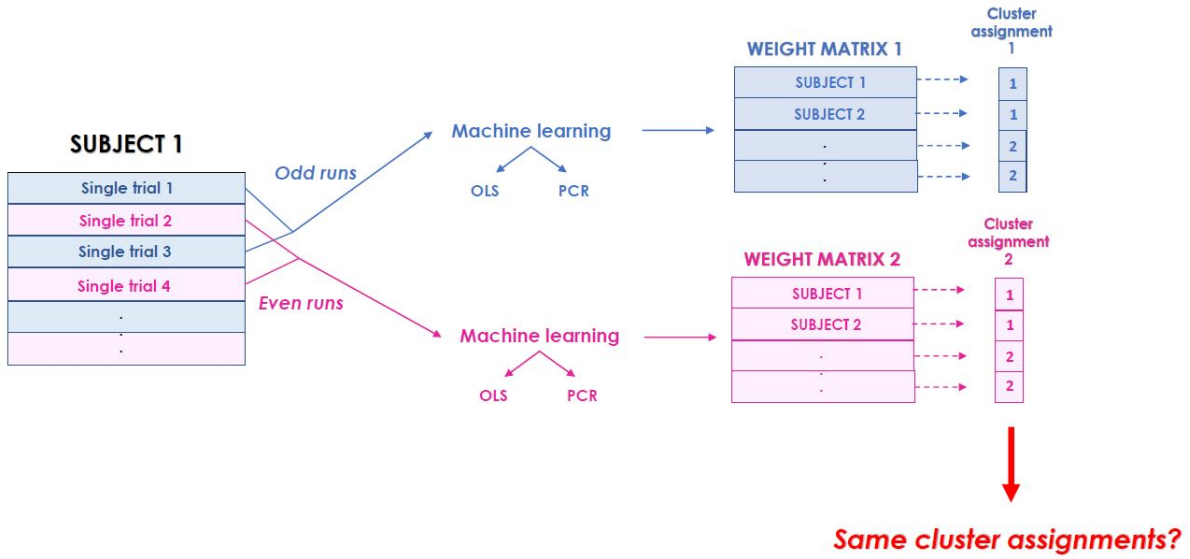


Figure 5.4: Scheme of the procedure for testing the reliability of the clusters

This creates two distinct weight matrices: they still have the same shape, i.e. ($nb\_participants \times nb\_voxels$), but now the regression step (OLS regression or PCR) has been performed on half of the data instead of all the single trials of the subject in question. After that, the clustering step can be performed on both matrices. Then, we can test whether the same participants form the same clusters in both datasets. For example, if subject 1 belongs to cluster 1 based on one half of their single trials data (the odd runs), they should belong to the same cluster based on the other half of the data (the even runs). This would in turn reflect the reliability of the clusters. The procedure is described in Figure 5.4.

5.4.1.1.2 Cluster comparison The comparison of cluster assignments is not as simple as it seems. If you keep in mind that for the clustering algorithm presented here above the first step is to randomly assigned each point to a cluster, it is easy to understand that two clusters that would be approximately in the same location in space might not have the same cluster label (here, the clusters are simply labeled 1, 2, 3 and 4). For example, this effect appears in the two upper graphs in Figure 5.5: it can be seen that clusters 1 and 4 correspond to the same cluster in the two graphs, while clusters 2 and 3 seem to be reversed. Here, the results are only presented to illustrate the effect mentioned in this section, but the actual results will be explained and discussed in the following chapter.

So, the first thing to do is to find how clusters labels correspond in both graphs. Comparing the plots could suggest one way to perform this task, but it would not be optimal because not automated. Another solution would be to compare the distances between the clusters centroids, and keep the ones with the smallest distance. For example, if we compare cluster 1's centroid in the first image with the four centroids in the second image, if the shortest distance is with cluster 2's centroid, we would consider that the points labeled as 1 in the first image would correspond to the points labeled as 2 in the second image. Once again, this solution is not optimal, because it would then be impossible to compare clusters performed in a two-dimensional space (such as the case of the signatures responses) with clusters in a higher dimensional space (such as the case of the whole weight maps). The final solution, which was implemented, consists of pairing the cluster labels that have the highest number of shared subjects.

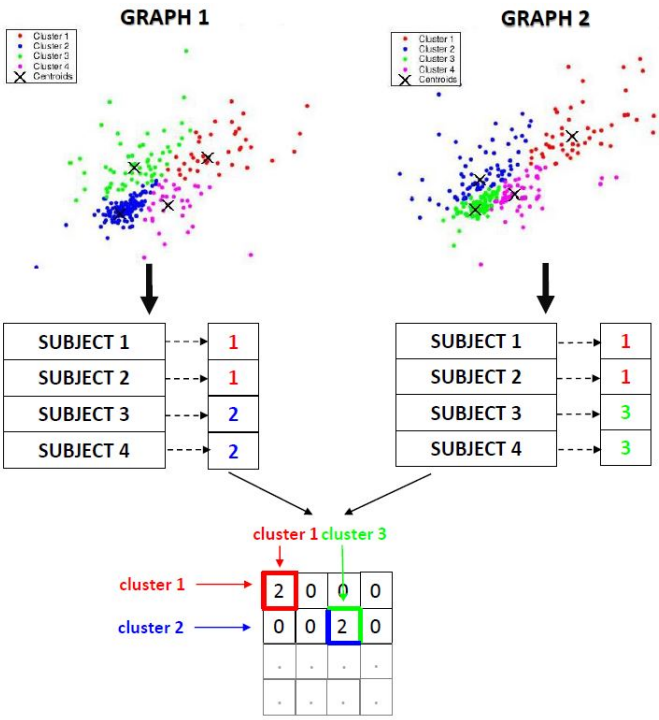


Figure 5.5: Procedure to find corresponding clusters

To determine the number of shared participants between the four clusters coming from each dataset, a confusion matrix is constructed. As a reminder, this tool is used to assess the accuracy of a classification problem, by comparing results from actual and predicted classifications and detecting correct and incorrect predictions [72]. In this case, the different classes of the confusion matrix correspond to the different clusters labels (from 1 to 4). In a perfectly theoretical situation, the clusters indexes would be the same in the two distinct datasets, meaning that all the subjects belong to the same cluster in both datasets. Considering that, we can imagine that the clusters indexes found in the first dataset constitute a kind of actual classifications, while the ones from the second dataset correspond to the predictions. The construction of the confusion matrix is conducted in the following way: the correct classifications correspond to the subjects that belong to the same cluster in both datasets, while the confusions correspond to the subjects belonging to other clusters. This matrix is then used to determine how the clusters correspond in the two distinct datasets, by keeping pairs of clusters corresponding to the highest correct classification rate. An example is shown in Figure 5.5. In the table in this figure, it can be seen that the first row, corresponding to cluster 1 in the first dataset, shows a highest number of shared participants with cluster 1 in the second dataset. Thus, we keep this pair of corresponding cluster to continue the assessment of cluster quality. However, for the second row, we see that cluster 2 from the first dataset would be associated with cluster 3 in the second dataset, as their number of shared participants is higher than for the other clusters.

However, it might happen that the situation occurs for two clusters, i.e. one cluster from the first dataset is associated with two different clusters in the second one. This situation is avoided by choosing the one pair defined by the highest number of shared subjects. The process is iterated until one cluster in the first dataset corresponds to one and only one cluster in the second dataset.

At the end of this process, the clusters labels are modified in the second datasets so that corresponding clusters are defined by corresponding labels. This step is illustrated in Figure 5.6.

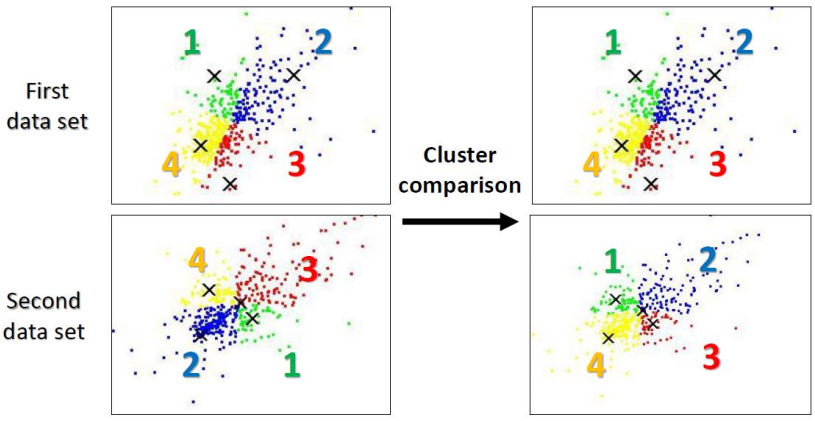


Figure 5.6: Rearrangement of clusters labels after clusters comparison

5.4.1.1.3 Global reliability The cluster reliability is evaluated by the percentage of correctly clustered subjects, i.e. subjects belonging to the same cluster in both datasets. This percentage is simply equal to the sum of correctly clustered participants in each cluster, divided by the total number of participants. In the following section, another reliability assessment technique is presented, which is a slightly different extension of the one described here. This is why the term "global" is used here, as opposition to the "balanced" one in the next section.

5.4.1.1.4 Balanced reliability After running some tests, it was noticed that this global reliability could, in some cases, reach extremely high values (around 95%). One has to stay vigilant, it should be kept in mind that the data used in the project comes from fMRI measures, which are known to be highly unstable. In addition, those results were found when performing the clustering step on the whole multivariate/univariate maps, although they were believed to lead to the worst results due to the high dimension of the data compared to the signatures patterns. See section Chapters 6 and 7 for the actual results. It seemed obvious that something was inaccurate with this reliability measure.
Indeed, by looking closely to the results obtained, it was discovered that in some cases, such as the case of the clustering on whole weight maps, almost the entirety of the subjects belonged to the same cluster. Then, the chance for one subject to belong to another cluster is much decreased, and that is what explains the high reliability values. To take this effect into account, a variant of the previous reliability metric was developed: the balanced reliability. Now, the reliability is computed for each cluster, by computing the ratio between the number of subjects of cluster i belonging to the right corresponding cluster j and the total number of subjects in cluster i. At the end, the balanced reliability equals the mean of these k individual cluster reliability values. We will see in the following results that this metric is sometimes higher, sometimes lower than the previous global reliability.

To test whether this balanced reliability is trustworthy, it was important to determine if these results were unlikely to happen by chance. The technique used to realize this task is permutation tests, also called randomization tests. This is a statistical tool which is based on the construction of sampling distributions, in the same spirit as the bootstrapping method. It consists of permuting, without replacement, the observed data in order to construct what is called a "permutation distribution", in the shape of a histogram. This procedure is employed to verify that the outcomes did not appear by chance, but instead they are observed independently on the observation labelling. If we consider the null hypothesis that the outcomes are observed whatever the condition, we can compute a t-statistic for each permutation, and create a permutation distribution of the statistic

under the null hypothesis. Thus, the comparison between the observed effect and the distribution of values obtained after permutation allows to evaluate the significance of this effect, corresponding to a p-value. It should be noted that the theory behind permutation tests has been developed many years ago, but only the recent progress of computational efficiency have allowed them to constitute a useful tool. When working with large datasets, computing all the permutations is not feasible, although the theory would require it. In fact, the tests conducted here (and in many situations) are "approximate permutation tests", consisting in a large number of random permutations of the data. An example of the distribution obtained using permutation tests is shown in Figure 5.7, where the blue line corresponds to the observed data. At first glance, the null hypothesis can be rejected, as the observed data seems to be quite extreme comparing to most of the outcomes obtained after permutations. A p-value is computed in order to verify that, by counting the number of permuted outcomes associated with larger value than the one actually observed [73, 74, 75].

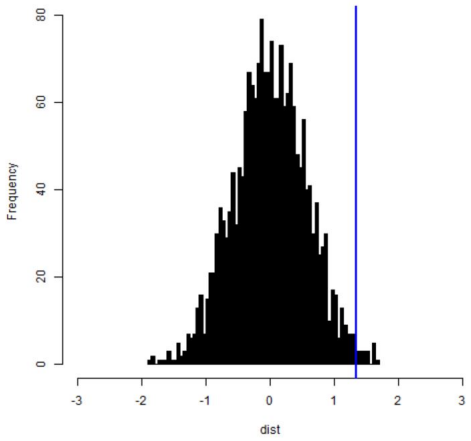


Figure 5.7: Example of the results of permutation tests [74]

Here, the permutations are performed on the cluster labels. We keep the cluster labels of the first dataset as they are, and we permute the ones of the second dataset. Due to the quite high number of participants (433), all the permutations cannot be tested. The solution is to perform a number of random permutation of the cluster labels. Here, this number is equal to 100 000. For each permutation, we compute the balanced accuracy between the unchanged cluster labels from dataset 1 with those from dataset 2 that were permuted. The results are visualized by the means of a histogram: the horizontal axis represent the different balanced reliability values obtained with different frequencies, defined by their lengths on the vertical axis. The actual observed balanced reliability, computed with the true labels (before any permutation), is also represented on the graph, as a red cross. A p-value is associated with it. Our hope is that the observed outcome lies on the extremity of the histogram, and is linked with a rather small p-value.

Due to quite high computation times, the balanced reliability investigation using permutation tests is applied only on the two data types leading to the best results.

It is important that the test for reproducibility should not be entirely based on the comparison of clusters. Indeed, in the specific case where the signature patterns are applied to the univariate/multivariate maps, our interest is in the difference between subjects on how they correlate to both signatures. The following section presents an alternative way of quantifying the reproducibility of the participants' behaviors.

5.4.1.1.5 Alternative: compare signatures responses Another way of assessing the reproducibility of the different behaviors in the participants is to directly compare their signatures responses amplitude. Discovering whether some subjects correlate more for one signature, e.g. NPS, than the other, e.g. SIIPS, constitutes one interesting outcome of this study. In the same spirit as earlier, we want to make sure that participants show the same behavior based on two distinct datasets.
In order to be able to compare the two signatures responses for each individual, it is important to rescale those values so that they can be comparable. Indeed, when we apply the signatures onto the weight maps, one signature usually shows higher signal amplitude than the other one. This is done by z-scoring, i.e. standardizing within subjects, the signatures responses. Thus, their values are transformed so that they have zero mean and unit standard deviation, as explained in section 4.6.
The idea here is to directly compare the signatures responses for each subject based on the two distinct datasets. In Figure 5.8, the two blue axis constitute a different representation of the data. The first direction, i.e. the long axis, represents a measure of the strength of the signal, for both signatures responses. However, the second direction, i.e. the sort axis, is a measure of the scores for each signature. Measuring the position of each individual point in this new representation would give help determining if participants reliably track one signature more than the other. Indeed, if we look at the position of the data point in the second axis, a positive value means that this participant shows higher correlation with the SIIPS signature than for the NPS one. A point with a negative value means the opposite.

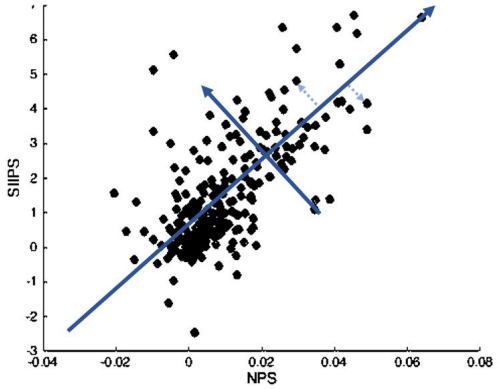


Figure 5.8: Measurement of signatures responses

In the results, the way for measuring this distance is to compute, for each subject individually, the difference between the NPS and the SIIPS responses values for the first dataset (the odd runs) and investigate its spatial correlation with the same difference for the second dataset (the even runs):

$$corr((zscore(NPS_1) - zscore(SIIPS_1)), (zscore(NPS_2) - zscore(SIIPS_2))) \quad (5.6)$$

5.4.1.1.6 Wrong cluster assignments At this point, with the graphs showing the different clusters and with the reliability computation, we can find which participants are not "well clustered", i.e. that do not belong to the same cluster in the two datasets. Indeed, since we were able to find and count the subjects that belong to the same clusters to quantify the reliability of the clusters, it was not difficult to find the ones that do not, and to visualize them on the graphs. We would expect to see them on the boundaries between the clusters, because from one clustering configuration to the other, the boundaries are rarely at the exact same location in the space.

5.4.1.2 Cluster consistency

The cluster consistency is directly assessed using one metric called the Silhouette value. It is a measure that combines the cohesion of the clusters, i.e. how similar is an observation to its own cluster, and their separation, i.e. how different they are. It helps to provide information both graphically and numerically. The Silhouette value, comprised between -1 and 1, for observation i assigned to a cluster A can be written as follows:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \tag{5.7}$$

Where $a(i)$ represents the average dissimilarity of observation i to all other objects of cluster A, and $b(i) = \min_{C \neq A} d(i, C)$ is the smallest distance between observation i and a cluster $C \neq A$, where $d(i, C)$ is the average dissimilarity of observation i to all objects of cluster C. $a(i)$ is a measure of how well object i is assigned to its cluster. The similarities measure how much the objects resembles, while the dissimilarities measure how far two objects are from each other. The cluster C corresponding to this minimal distance is called the "neighbor" of observation i and can be seen as the "second best choice", in case where the affiliation to cluster A is compromised.
A Silhouette value $s(i)$ close to 1 indicates that the observation i is well matched in the cluster assignment. This happens when $a(i)$ is very small compared to $b(i)$, meaning that i is very close to its assigned cluster and very far from its closest neighboring cluster. If the Silhouette value tends to -1 or 0, it means that the point is closer to its neighbor than its assigned cluster, or that it lies at the boundary between those two clusters, respectively.
As already said, the Silhouette values are computed for each observation separately, and the results can be observed graphically. An example of the results that could be obtained is shown in Figure 5.9. In this figure, each cluster is represented by a "silhouette" of the observations it contains, sorted by their Silhouette values. The longer the silhouette, the more consistent the clusters. In the figure, it can be seen that cluster 1 has rather high Silhouette values among the points assigned to it, while the three other clusters show negative values, meaning that some points should not be assigned to this cluster. For example, by observing this graph, we understand that cluster 4 is not consistent, since it involves more wrongly assigned points than right ones.

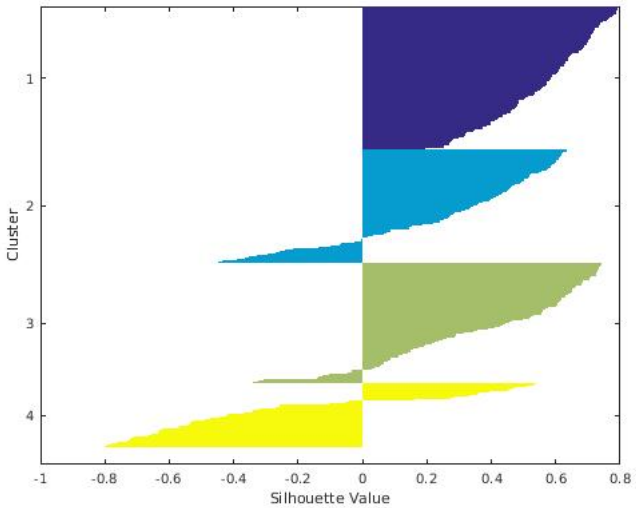


Figure 5.9: Example of Silhouette plot

This constitutes the graphical results of the use of Silhouette values, from which we can already make some conclusions about the clustering previously performed. A numerical value can be also computed, as each individual value does not mean very much on its own. Indeed, a mean of the Silhouette values for each object gives an idea of the overall clustering results. Obviously, the higher this mean value, the better.

For example, when investigating which distance metric was the most appropriate, the Silhouette plots associated with each metric constituted one criteria for the final choice. However, to give an idea of the cluster quality, representing the cluster quality in terms of consistency (as opposition to reproducibility), the mean Silhouette value was used.

5.4.1.3 Cluster correlation

When the k clusters are determined, one way to verify that they correctly separate the dataset into k groups of different brain activation is to use spatial pattern correlation measurements. In this project, Pearson correlation was used. It is a parametric measure which is used to compare pairs of variables in a population, and determine their strength and direction to a linear relationship. It produces a correlation coefficient, r, comprised between -1 and 1. If $r = -1$, it means that the two measurements have perfectly negative linear relationship, while $r = 1$ means that they have a perfectly positive one. A correlation coefficient equal to 0 means that no relationship exists in the two measurements [71].
Here, we want to make sure that corresponding clusters in the two distinct datasets have a high spatial correlation, i.e. a correlation coefficient close to 1, while having a low correlation, i.e. close to 0, with the other, non-corresponding, clusters. In other words, the within-cluster correlation must be high, and the between-cluster correlation must be low. Those two measurements are represented in Figure 5.10.

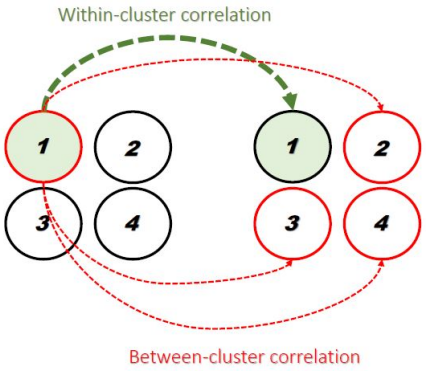


Figure 5.10: Within- and between-cluster spatial correlations

In this figure, and in the actual method, the weight maps used to compute the correlation coefficients for each cluster c correspond to the average of the weight maps from all the participants belonging to cluster c. In some way, the average weight maps would correspond to the cluster's centroid.
The spatial correlation is computed with the function `corr.m` from Matlab.

The spatial correlation is computed between entire maps, or more precisely, between clusters' average weight maps. It means that the correlation is assess voxel by voxel.
The accuracy of within-cluster and between-cluster correlations are investigated using permutation tests, described in section 5.4.1.1.4. In this case, as in the case of the balanced reliability, the indexes

are permuted several times. Then, the two correlations are computed based on those new indexes. The results consist of two individual histograms: the first one corresponds to the various within-cluster correlation values obtained with the permutations of the clusters labels, with a red-cross and a p-value associated with the actual observed one. As in the case of the reliability-based permutation test, the red-cross should lie at the right extremity of the histogram, with a rather small p-value. The second plotted histogram represents the outcomes obtained for the between-clusters correlations. The difference here is that we want this correlation to be quite small. Thus, we want to find the actual observed between-cluster correlation at the left extremity of the histogram, towards the smallest value. A p-value is also associated with this observed outcome and should be small as well. In this case, the number of randon permutations is equal to 1 000, resulting from a trade-off between computation times and method's accuracy.

The high computation times are also the reason why the correlation evaluation using permutation tests is applied only on the two data types leading to the best results.

5.4.2 Neuroscientific results

It is important that the neurological aspect of the clustering results is also investigated within this study. Indeed, the k clusters found in each situation described in section 5.1 should be analyzed to see whether the clusters are actually associated with unique and dissimilar brain activation. This step is performed using a specific use of the t-test, a voxelwise one sample t-test. The results of this t-test are then thresholded to visualize the brain regions that show higher or lower brain activation, specific for each cluster. This thresholded maps can be compared to each other, to see the differences in activation between the different clusters, but also compared the corresponding clusters in the two datasets, to see if this activation is rather constant.

Initially, a t-test, or Student's t-test, consists of a statistical tool used to compare the means of two groups. The idea is to determine whether or not this difference in groups' averages reflects a "real" difference in the population, i.e. the difference could not happen by chance. It can be of two types: either an independent t-test, used in the case where the two groups are independent of each other, or a paired t-test in the opposite situation. The test is based on a t-statistic value, which is computed with the following formula:

$$t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \tag{5.8}$$

Where $\bar{X_1}$ is the sample mean from the population with a size n, μ is the population mean and σ is the population standard deviation. The purpose of the t-test is to determine whether or not to reject the null hypothesis, which states that there is no difference between the true mean μ and the comparison value $\bar{x}$.
In the case of a one sample t-test, which is the case here, the null hypothesis states that the data comes from a normal distribution with zero mean and unknown standard deviation. In the case of a voxelwise t-test, a voxel is declared active when its test statistic reaches a sufficiently extreme values, compared to the statistic's distribution under the null hypothesis [40, 67, 76, 77]. In this situation, the weight maps, reflecting the activation at the voxel scale, is compared with the case where no activation is detected. For each cluster found, the t-test is applied to the group of weight maps (multivariate or univariate) corresponding to the group of subjects belonging to that cluster. Doing that, it is possible to detect which brain regions, more precisely which brain voxels, are more likely to be subject to pain-related activation in each individual cluster, and to visualize them. The visualization is preceded by a thresholding step in order to control for a certain significance level α, chosen by the user. If the voxel's test statistics determined by the t-test surpasses this threshold,

it is considered as active. However, it might lead to a rather high number of type I errors, i.e. false positives, voxels declared as active when they are in fact inactive. Thus, the correction is done with False Discovery Rate (FDR). This tool is used when performing multiple tests (here a test is performed for each voxel), and it is defined by the proportion of false positives among those tests for which the null hypothesis is rejected. It can be written as follows:

$$FDR = \frac{V_{ia}}{V_{ia} + V_{aa}} = \frac{V_{ia}}{D_a} \tag{5.9}$$

Where V_{ia} is the number of false positives and $D_a = V_{ia} + V_{aa}$ denotes the number of voxels declared active (true and false positives). When we use a procedure controlling the FDR, a threshold q, in the form of a percentage comprised between 0 and 1, is specified. It is used to ensure that, on average, the FDR does not surpass q. The controlling procedure is based on the following equation:

$$E(FDR) \leq \frac{T_i}{V} q \leq q \tag{5.10}$$

Where $E(FDR)$ is the expected value, V is the number of voxels to be tested and T_i is the number of truly inactive voxels [77]. For a deeper description of the procedure associate with FDR, see Ref. [77]: Christopher R. Genovese, Nicole A. Lazar, and Thomas Nichols, *Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate*, 2001.

The brain activation patterns are visualized using montage plots, showing positive and negative activation amplitudes with a color bar, displayed on several brain slices previously drawn.

5.5 Determination of the most accurate number of clusters

In the beginning, all the tests were conducted in the idea of discovering a number of 4 clusters. In the case of the signatures patterns, this chosen number makes sense (see types of clusters searched in section 3.1.2.3). But when we are working with the entire maps, there is no way to know which number of clusters would be more relevant. Although it is still interesting to perform the tests with a 4-clusters search, in order to compare the results with the one discovered using the signatures patterns, it is important to investigate if another number of clusters could show more reliable results.

In this section, two procedures are presented: the first one consists of building the dendrogram that would results of a hierarchical clustering of the data. It is applied on both the signatures patterns and the entire maps, in order to visualize which number seems the more relevant. The second procedure consists of comparing on one side the reliability results (balanced accuracy) for different number of clusters, and on the other side the mean Silhouette values for the same numbers of clusters. Here, the tested numbers go from 2 to 15 clusters.

5.5.1 Based on dendrograms

Dendrograms constitute the tool used for the visualization of hierarchical clustering, which is another unsupervised machine learning technique based on agglomerative clustering. Just as k-means clustering, this tool is used to detect groupings of objects that best represent measured relations of similarity [65].
The algorithm starts by assigning each object to its own cluster. The dissimilarities between the objects can be determined, in the form of a distance matrix. Then, in an iterative manner, the two most similar clusters are joined and replaced by only one cluster. The distance matrix is then updated and redefines the dissimilarities between this new cluster and the other objects/clusters.

This stops when all the objects belong to the same cluster, but the graphical visualization, i.e. the dendrogram presented here under, allows to determine which number of clusters seems to be the most appropriate. The computed distance between the different clusters can be of three kinds [66, 33]:

- **Single linkage**: the distance corresponds to the smallest distance between clusters (i.e. the closest points belonging to two separated clusters)
- **Complete linkage**: the distance corresponds to the largest distance between clusters (i.e. the farthest points)
- **Average linkage**: the distance corresponds to the average distance from all the points

In this project, complete linkage is used. It was chosen because the visualization of the results allows to understand that the points are very close to each other, we do not see well defined clusters in the data points alone. Thus, choosing a distance computed between the two farthest points is more likely to highlight several clusters among the data than the distance between the closest points.
The distance metric used is also a parameter that must be chosen by the user, the choices are approximately the same as for the k-means clustering. The results will be presented for the Euclidean and the cosine distances, consistently with the distance metrics investigated with the other algorithm.
The dendrogram is the visualization tool linked with this type of clustering, and its shape is shown in Figure 5.11. The horizontal axis represents the objects, or clusters of objects. For an easier representation, in the case where the number of objects is quite large, the object numbers already correspond to clusters. The vertical axis stands for the distance between the objects/clusters. The objects/clusters that are joined are combined by a horizontal line, and the length of the vertical line represents the distance between them [33]. For example, in this figure, we see that object 1 and object 4 are quite close to each other, since the vertical line between each of them and their junction is quite short.

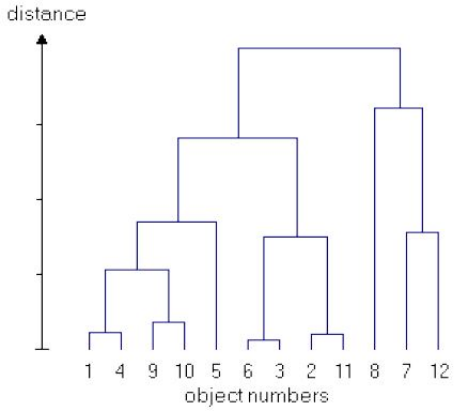


Figure 5.11: Example of dendrogram [33]

In this project, this tool is used to have an idea on which number of clusters seems to be the most appropriate in the data available.

5.5.2 Based on reliability measures

To select the best number of clusters based on reliability measures, the search for clusters is started k times, each time with a different number of clusters ranging from 2 to 15. This is done in parallel

with the two distinct datasets, and for each k the balanced accuracy is computed and saved. At the end of the process, a graph of the balanced accuracy as a function of the number of clusters is returned. The best number of clusters would correspond to the highest point on this graph.

5.5.3 Based on Silhouette values

The procedure for this search of the best number is identical to the previous one, with the exception that the measurement considered in this case is the mean Silhouette value of all the points for each number k of clusters. Since the two distinct datasets will return two distinct values, we take their mean as the measurements in the resulting graph.

Chapter 6

Stratification using multivariate maps

Contents

In this chapter, we present the results obtained by performing the clustering on the multivariate maps, on the entire maps and after each of the dimensionality reduction techniques described in section 5.1. In the first two sections, i.e. the clustering on the entire maps and the clustering on the signatures responses, the two distance metrics, Euclidean and cosine, are investigated. The results obtained in those sections will allow to choose one of these metrics, and to keep it for the rest of the work.

In each section, the different methods for assessing cluster quality are applied and the results are presented. After analyzing each data types individually, the second objective, i.e. the visualization of the brain regions activated in the different clusters, is conducted for the two techniques that lead to the best results. This was decided for the sake of clarity for the reader, and because the best techniques are more likely to give more meaningful results.

6.1 Entire maps

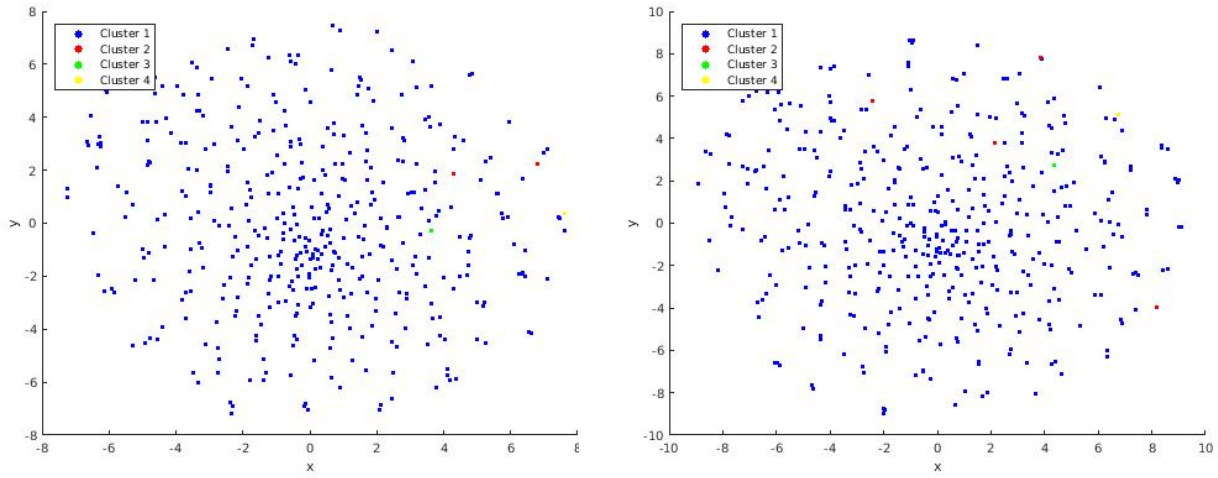
Due to the high instability linked to the great dimensionality (corresponding to the number of voxels, 352 328 in this case) of this type of data, we expect to find quite poor results. It seemed interesting to test it anyway and investigate whether or not the dimensionality reduction techniques can increase the various cluster quality metrics.
The graphical and numerical results for the multivariate entire maps appear in the following sections. The first section is dedicated to the Euclidean distance case, and the second one to the cosine distance. The results obtained here will already help to bias towards one of the distance metrics.

6.1.1 Euclidean distance

In order to present consistent results, the k-means algorithm and the Silhouette plot and values are computed based on the Euclidean distance.

6.1.1.1 Graphical results

The clusters obtained are displayed in Figures 6.1a and 6.1b. We can already notice that, in both cases, the majority of the points belong to the same cluster, i.e. cluster 1 on the graphs. It means that the data is not easily separable into a number of subgroups, at least with this distance metric. The points not belonging to cluster 1 could constitute a kind of outliers, but the dimensionality of the data does not allow to detect and remove them in an accurate manner. In this project, the outliers are never removed. Indeed, instead of thinking of suppressing them from the dataset, they should bring extra attention, those people represent neural activation that stands out from the average. Thus, there is no reason for suppressing them in this kind of work. In addition, it should be kept in mind that the previous analysis of the data already took the precaution of removing outliers.



(a) First dataset

(b) Second dataset

Figure 6.1: Clusters in the entire multivariate maps (Euclidean distance)

6.1.1.2 Cluster quality

As a reminder, three metrics are used to evaluate the cluster quality: the cluster reproducibility representing the percentage of the subjects that belong to the same cluster based on two distinct datasets, the cluster consistency based on the Silhouette plot and mean value, and the cluster correlation which tests the spatial correlation between corresponding clusters and non-corresponding clusters.

6.1.1.2.1 Cluster reproducibility Table 6.1 shows the number of shared subjects between the different clusters. The rows correspond to the cluster labels in the first dataset (Figure 6.1a) and the the columns correspond to the labels in the second one (Figure 6.1b). This table is used to find how the clusters correspond in the two datasets, as explained in section 5.4.1.1.2. When this is done, the number of participants that belong to the same cluster are counted, and this leads to the global and balanced accuracies values presented here under.

As previously seen in the graphs, the great majority of the participants belong to the first cluster. In total, the dataset is composed of 433 participants, and in this cluster they are 426, so more than 98%. With this table, we expect to find very high values for the reliability measure. Indeed,

	1	**2**	**3**	**4**
1	426	2	0	0
2	0	2	0	0
3	0	0	1	0
4	0	0	0	1

Table 6.1: Number of shared subjects in all clusters - multivariate maps with Euclidean distance

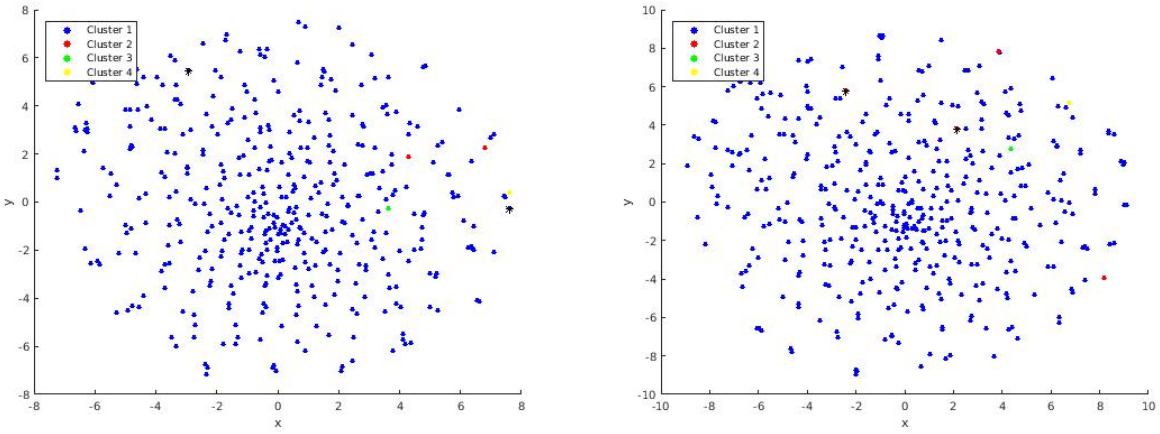
since more than 98% of the participants belong to the same cluster, there is very few chance to be "wrongly clustered".

- **Global accuracy**: 99,54 %

- **Balanced accuracy**: 99,88 %

We see that, as expected, the reliability reaches extremely high values. Although, such good results do not mean the method they came from is perfect. Indeed, there is not much information to be learned from clusters looking like this. This kind of result already brings reluctance to use this distance metric. Note that in this case the balanced reliability is as high as the global one, because exceptionally the few points composing the three smaller clusters are also well clustered.

As the numerical value implies, the number of wrongly assigned subject is very small. In Figures 6.2a and 6.2b, the black crosses indicate the participants that do not belong in corresponding clusters in the two graphs. Here, there are only two. Once again, this is due to the fact that the majority of participants belongs in one cluster.



(a) First dataset

(b) Second dataset

Figure 6.2: Wrong assignments in the entire multivariate maps (Euclidean distance)

There is no utility in pursuing the analysis of this type of maps because the clusters are not meaningful. Even if the cluster consistency and correlation could generate good results, there is no point in studying such clusters. This section shows a negative point for the Euclidean distance use. Thus, the second distance metric, cosine distance, is investigated in the following section.

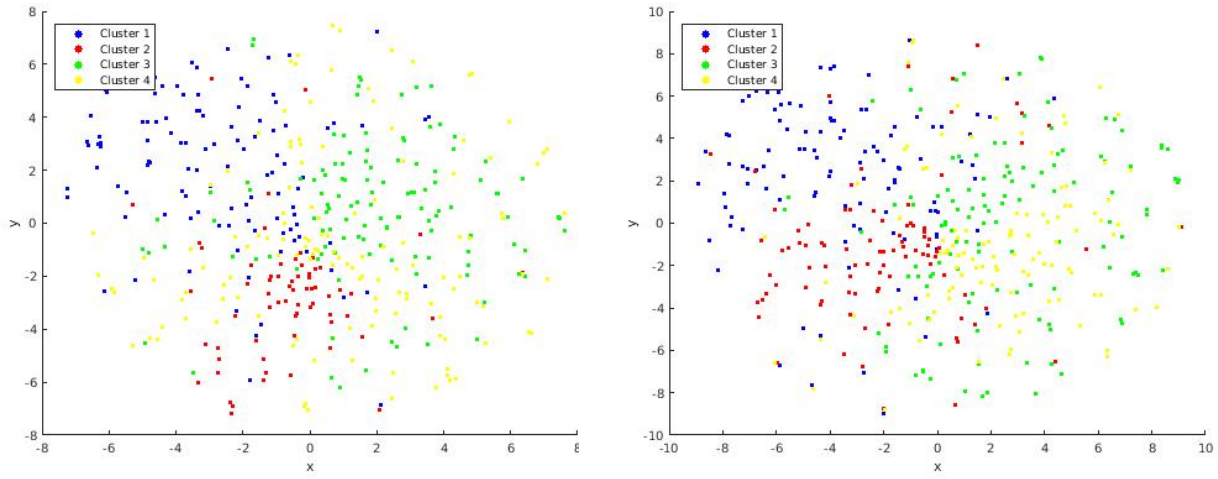
6.1.2 Cosine distance

After revising the case of the Euclidean distance, the same steps are performed on the entire multivariate maps with the k-means algorithm and the Silhouette plot and values based on the cosine distance.

6.1.2.1 Graphical results

Figures 6.3a and 6.3b show the clusters obtained with k-means based on the cosine distance. The simple visualization of those clusters implies that the clusters found are much more relevant than for the Euclidean distance. Indeed, each cluster seems to contain an approximately equal number of subjects. It should be kept in mind that the search for clusters is done directly on the maps, and then the t-SNE method is used only for visualization. The color code, i.e. the clusters labels, results from the k-means algorithm applied on the maps. It can be seen that the clusters are not

separated with clear boundaries in this case, but they are not completely spread out in the entire space, data points with equal label are usually close to each other.



(a) First dataset

(b) Second dataset

Figure 6.3: Clusters in the entire multivariate maps (cosine distance)

6.1.2.2 Cluster quality

The cluster quality is evaluated with the three metrics mentioned above.

6.1.2.2.1 Cluster reproducibility Table 6.2 presents the number of shared subjects between the four clusters in the two datasets, just as in the previous section. This table, along with the two graphs displayed here above, show that the participants are more fairly distributed in the different clusters, contrary to the results of the Euclidean distance. However, we do not see a clear correspondence between the clusters, that will be interpreted by a high value of shared subjects in each row next to three small values: the high value would imply the cluster correspondence while the three small values would represent the participants that are mis-clustered. It is not the case here, meaning that the reproducibility of the clusters is not very good. Although, this conclusion brings no surprises, as the clustering is performed on the entire maps, along with their 352 328 dimensions. Indeed, we did not expect good results due to the instability of the weights. The dimension reduction techniques are supposed to bring some stability, and better results at the same time. As expected, the accuracy metrics are quite low, due to the high instability. Although they

	1	**2**	**3**	**4**
1	64	24	12	14
2	12	19	22	24
3	11	19	59	37
4	21	30	35	45

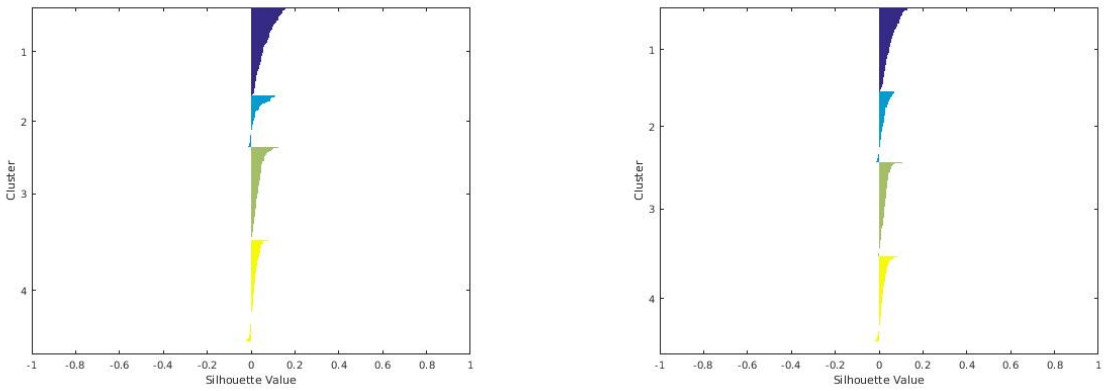
Table 6.2: Number of shared subjects in all clusters - multivariate maps with cosine distance

are much lower than for the Euclidean distance, they are more relevant, because they explain a behavior of clusters that makes more sense.

- **Global accuracy**: 41,90 %
- **Balanced accuracy**: 44,22 %

The graphs showing the participants that are mis-clustered are usually shown in order to analyze whether or not these subjects lie on the boundaries between the clusters. The clusters in the graphs from the two datasets, even if they look alike, rarely lie in the exact position in the space. This would imply that data points located very near the boundaries of the cluster in the first graph have more chance to belong to the neighboring cluster in the second graph. Showing that the mis-clustered points lie near the boundaries would support the clustering results. However, since in this case the clusters are not clearly delimited in space, showing the mis-clustered points is not useful. They appear in the Appendix C.

6.1.2.2.2 Cluster consistency The consistency is assessed with the Silhouette values, by the means of two tools: a visual description with the Silhouette plots in Figures 6.4a and 6.4b, and a numerical computation of the mean of the Silhouette values for each subject. This mean value is equal to 0,02 for both the datasets.
The plots and the mean value are very poor, which means that the clustering configuration is not right. Once again, this is probably due to the high dimensionality of the data.



(a) First dataset (b) Second dataset

Figure 6.4: Silhouette plots of the entire multivariate maps (cosine distance)

As stated in section 5.4.1.3, the correlation investigation is conducted only on the two data types leading to the best results. Here, working on entire maps means working with very high dimensional data, which is probably not the most appropriate method, and we expect to find better results with the dimensionality reduction techniques mentioned in section 5.1. The procedure is conducted in the following sections.

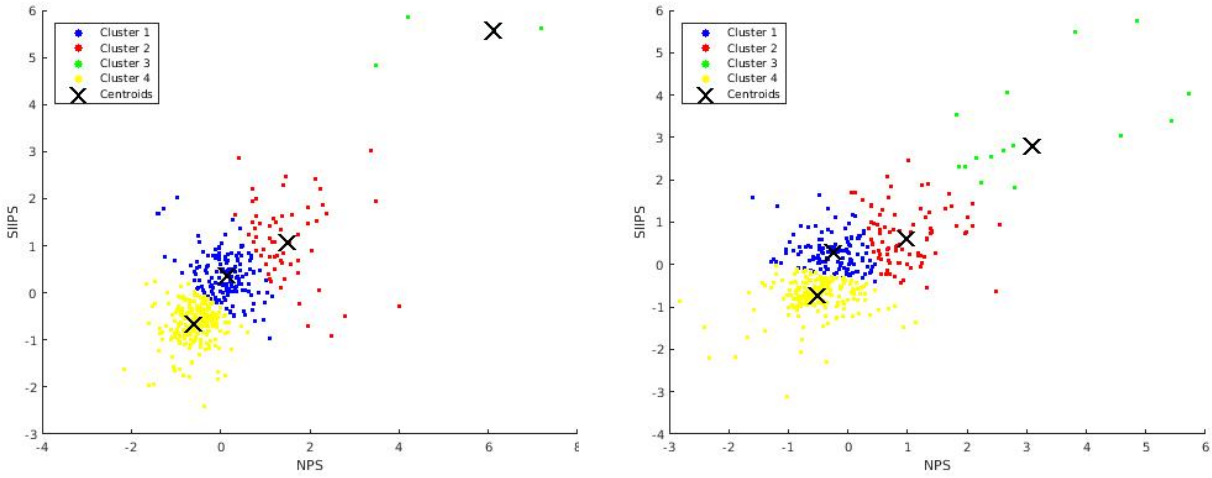
6.2 Signatures patterns

In this section, the first attempt to improve in some way the results by reducing the dimensionality of the data is presented. Here, the tool used is the application of two fixed patterns: the NPS and the SIIPS signatures. The theory behind those is explained in sections 3.1.1 and 3.1.2.

The two distance metrics, the Euclidean and the cosine distances, are tested with the signatures. At the end of this section, one of them is chosen for the rest of the work.

6.2.1 Euclidean distance

Figures 6.5a and 6.5b display the clusters obtained by applying the k-means algorithm, based on the Euclidean distance, on the signature responses on the multivariate maps.
Using Euclidean distance do not lead to false results, but they do not reflect the effect under study. Indeed, the clusters obtained show four distinct subgroups corresponding to four different levels of overall activation for both signatures: here, cluster 4 pools the subjects that have minimal activation amplitude, and clusters 1, 2 and 3 correspond to subjects with increasing amplitude, cluster 3 grouping the highest. Those clusters seem to measure the strength of the activation signal. This map might display true subgroups of subjects, but this effect is not what we are interested in for this project. What we are interested about is to detect different behaviors between the signatures, as explained in section 3.1.2.3.



(a) First dataset

(b) Second dataset

Figure 6.5: Clusters in the signatures responses on multivariate maps (Euclidean distance)

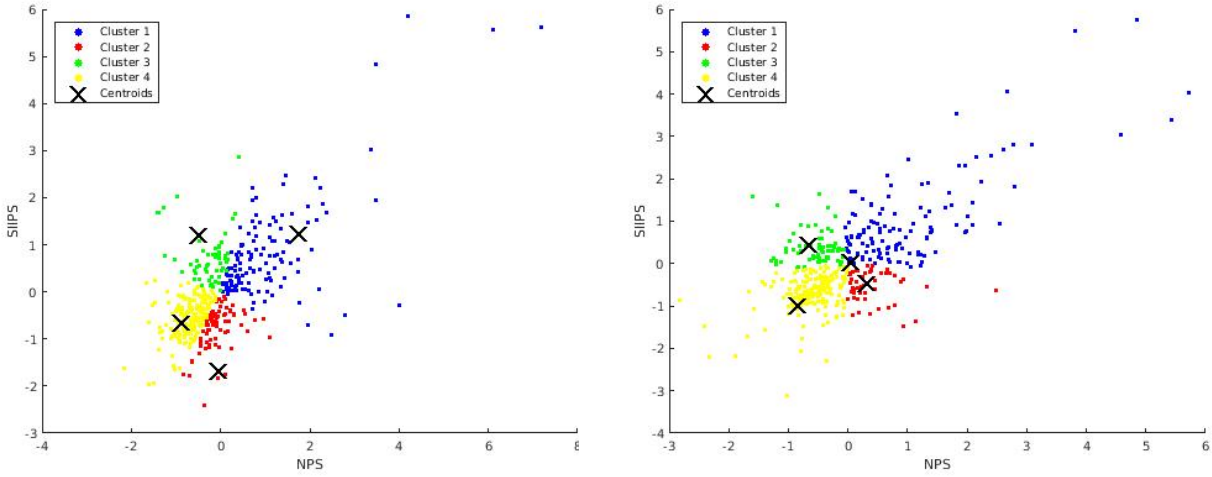
This outcome finishes to help making the decision about the distance metric to use for the rest of the project, which would be the cosine distance, since the Euclidean distance was not able to generate meaningful and useful results.

6.2.2 Cosine distance

In this case, the only way to detect the subgroups we are interested about is to use the cosine distance instead of the Euclidean one. Figures 6.6a and 6.6b exhibit the results obtained for this distance metric. This time, it is possible to detect the four clusters wished to be found. The cluster in the downer left quadrant corresponds to the subjects that have low activation for both signatures. On the downer right quadrant, we can see the subjects that have low SIIPS activation but high NPS activation. On the upper left quadrant, we see the contrary: the subjects that have low activation for NPS but high activation for SIIPS. And the last cluster, on the upper right quadrant, pools the subjects that have high activation for both signatures.

Although they seem slightly shifted, the correspondence between the clusters in both graphs is rather obvious, unlike in the case of the entire maps clustering.

Once again, these results support the choice of using cosine distance as the distance metric.



(a) First dataset

(b) Second dataset

Figure 6.6: Clusters in the signatures responses on multivariate maps (cosine distance)

6.2.2.1 Cluster quality

6.2.2.1.1 Cluster reproducibility The results are expected to be higher than for the whole multivariate maps, as the application of signatures brings a certain stability to the data. In the whole maps, the data are less constrained, more flexible, but at the same time they are highly unstable.

- **Global accuracy**: 52,08 %
- **Balanced accuracy**: 59,27 %

The accuracy outcomes are much higher than in the case of the entire maps. Therefore, the permutation tests for the investigation of the balanced reliability trustworthiness can be conducted. Figure 6.7 displays the results obtained. It can be seen that the observed reliability is quite high compared to the reliability values obtained after permutation of the cluster labels, and that it is associated with a quite small p-value. If we consider a significance level $\alpha = 5\%$, the null hypothesis can be rejected. This means that the balanced accuracy value found with the data did not appear by chance.

In the case of the clustering on the signature responses, it seems more appropriate to analyze the spreading of the participants that do not belong to the same cluster in both graphs. Indeed, in Figures 6.6a and 6.6b, we have seen that the boundaries between the clusters are well defined. But we have also seen that the boundaries in the second graph are slightly shifted comparing to the ones in the first graph. Therefore, we would expect to see the majority of the mis-clustered points in this shift area. The wrong assignments are shown in Figures 6.8a and 6.8b, as black crosses. It can be seen that the wrongly assigned subjects do not only lie on the boundaries, although a great

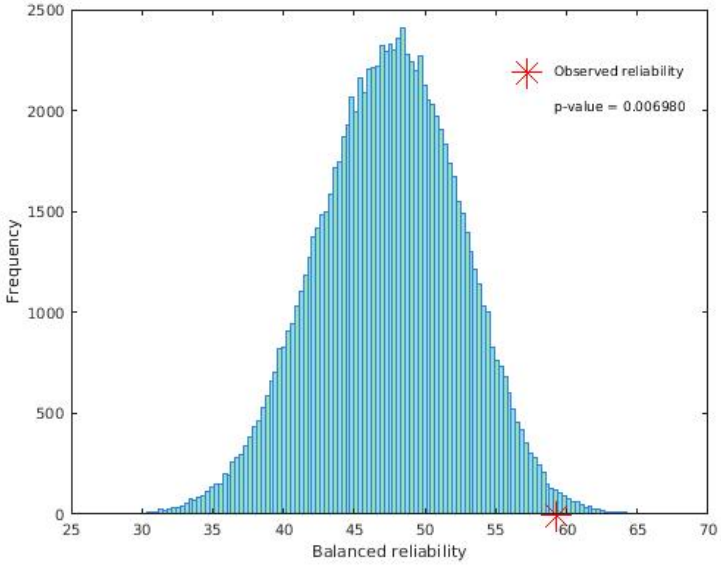
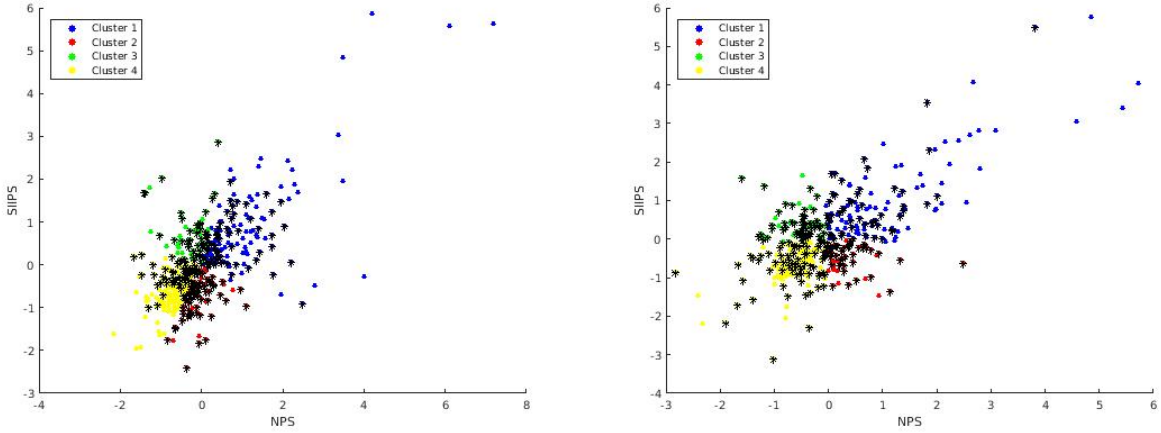


Figure 6.7: Permutation tests for cluster balanced reliability - Signatures patterns on multivariate maps

proportion of them do, but they seem kind of spread out in the entire space. The clusters corresponding to the low activation (cluster 4) or high activation (cluster 1) for both signatures seem to be less subject to this effect. The two other clusters contain a lot of mis-clustered subjects, which is why it seems interesting to apply the alternative method for evaluating the cluster reliability: based directly on the signatures responses coordinates in the 2D space.



(a) First dataset

(b) Second dataset

Figure 6.8: Wrong assignments in the signatures responses on entire multivariate maps (cosine distance)

This method was described in section 5.4.1.1.5, and can obviously be used only in the case of the signatures responses. As a reminder, this method is used to determine if the subjects track one signature more than the other, and it relies on the correlation between the difference in those

signatures, for each subject. Figure 6.9 displays this correlation, and shows that there is no obvious correlation between the measurements. Indeed, the Pearson correlation coefficient is equal to $r = 0,23$. This result is consistent with the fact that the majority of mis-clustered points appear in the two clusters corresponding to a higher activation of one signature than the other. The effect of interest is not so well captured with multivariate maps.

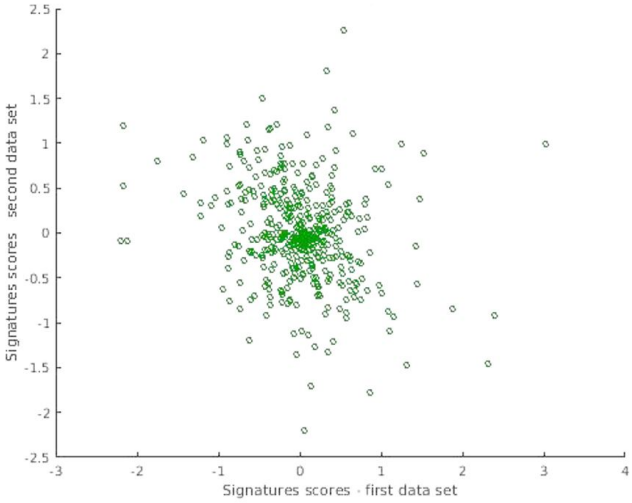
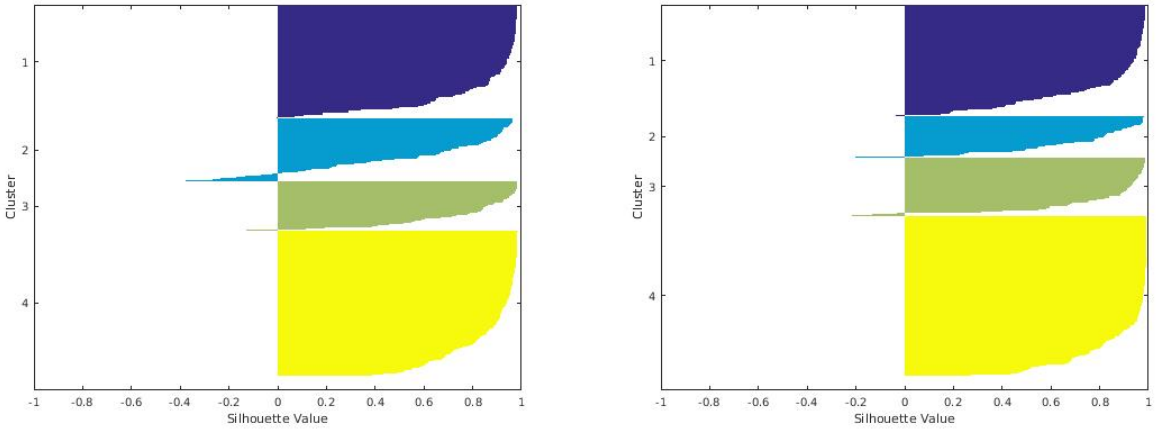


Figure 6.9: Correlation between signatures scores - Multivariate maps

6.2.2.1.2 Cluster consistency The cluster consistency is once again evaluated by the means of the Silhouette values. Figures 6.10a and 6.10b display the Silhouette plots for the two distinct datasets. It can be seen that the values are much higher than the Silhouette plots of the entire multivariate maps shown in Figures 6.4a and 6.4b. Besides, there are very few negative value, meaning that there are very few points that should not belong to their cluster, based on the Silhouette metric. The mean Silhouette values for the two datasets are equal to 0,68 and 0,73.

In addition to prove that the clustering results are rather accurate, the Silhouette plot is an additional verification that the cosine distance leads to better results than the Euclidean distance. Indeed, Figure 6.11 displays the Silhouette plot based on the Euclidean distance for the clustering of the signatures responses on the multivariate maps. The plot corresponds to the second dataset, which was more obvious. The first dataset plot appears in the Appendix C. The mean Silhouette values for the Euclidean distance were 0,57 and 0,48.

6.2.2.1.3 Cluster correlation The two histograms resulting from the permutation tests to evaluate the accuracy of the within- and between-cluster correlations appear in Figures 6.12a and 6.12b. In these figures, it can be seen that the within-cluster correlation, expected to be better than chance, reaches a quite high value compared to the permutations results. Considering once again a significance threshold at $\alpha = 5\%$, the p-value implies that the results are accurate for the observed within-cluster correlation. For the between-cluster correlation, the same conclusion about the p-value can be made, except that in this case the correlation value is rather small compared to the permutations ones. This is what was expected, thus we can conclude that the observed correlations are rather accurate in this situation.
It should be noticed that the number of random permutation will lead to more trustworthy results if it was higher than 1000, but the computation time did not allow that in this project.



(a) First dataset (b) Second dataset

Figure 6.10: Silhouette plots of the signatures responses on multivariate maps (cosine distance)

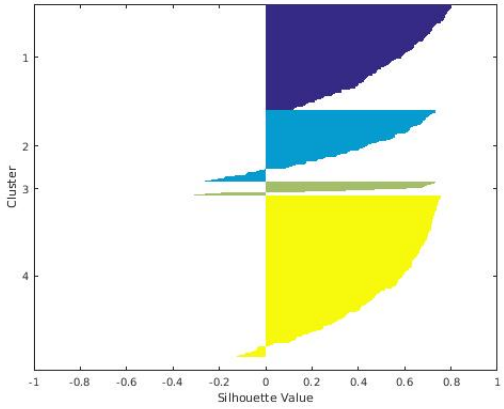
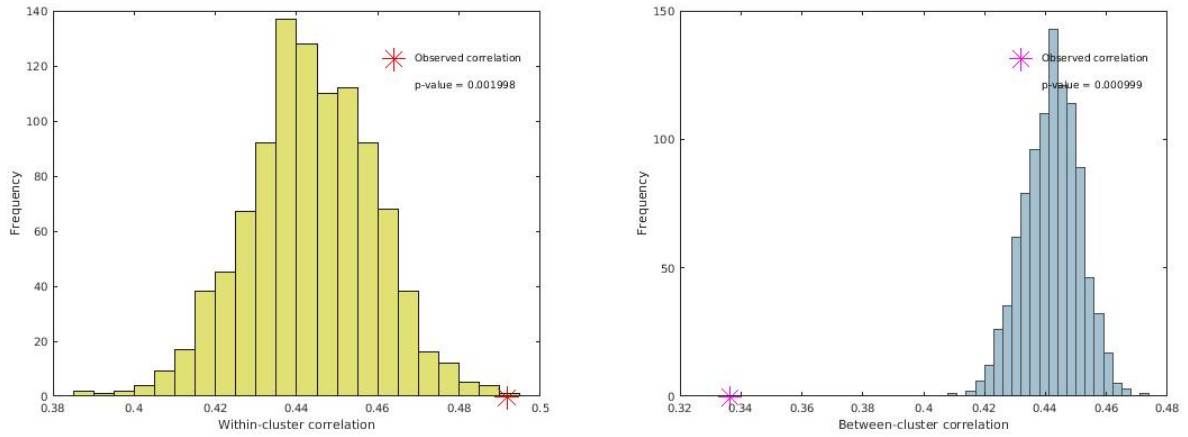


Figure 6.11: Silhouette plot of the signatures responses on multivariate maps (Euclidean distance)



(a) Within-cluster correlation

(b) Between-cluster correlation

Figure 6.12: Permutation tests for correlation - signatures patterns on multivariate maps

6.3 PCA reduction

The second dimension reduction technique which is tested is the Principal Component Analysis that reduces the dimensionality to a number of components chosen to lead to the highest cluster reproducibility. In this section, we do not use a prefixed pattern such as the signatures, but the reduction of dimensions should bring some stability and better results than for the entire maps.

6.3.1 Optimal number of components

The very first step consists of investigating which number of components to use in the PCA analysis. This number will in turn correspond to the new dimensionality of the data. Figure 6.13, representing the balanced accuracy as a function on the number of components, was constructed in the idea of finding the number of components corresponding to a maximum in the curve. This corresponds to a number of 30 components, but the curve is highly unstable and is characterized by various local maxima. We cannot see one obvious good solution in this graph. Therefore, in the case of the multivariate maps we will use 30 components in the PCA analysis, but the investigation of the optimal number must be applied to the univariate maps as well. It would be interesting to base the optimal number of components decision on another, more accurate, metric.

6.3.2 Cosine distance

As said earlier, the only distance metric that will be tested from now on is the cosine distance. The Euclidean distance results were computed but as they gave similar results than those presented here above, there is no utility in studying them in this project.
Once again, both the k-means algorithm and the Silhouette values are based on the cosine distance.

6.3.2.1 Graphical results

Figures 6.14a and 6.14b exhibit the four clusters obtained by performing the k-means algorithm on the PCA reduced maps, for each dataset. Although they still overlap quite a bit, the delimitations between them are more clear than the clustering configuration displayed in Figures 6.3a and 6.3b in

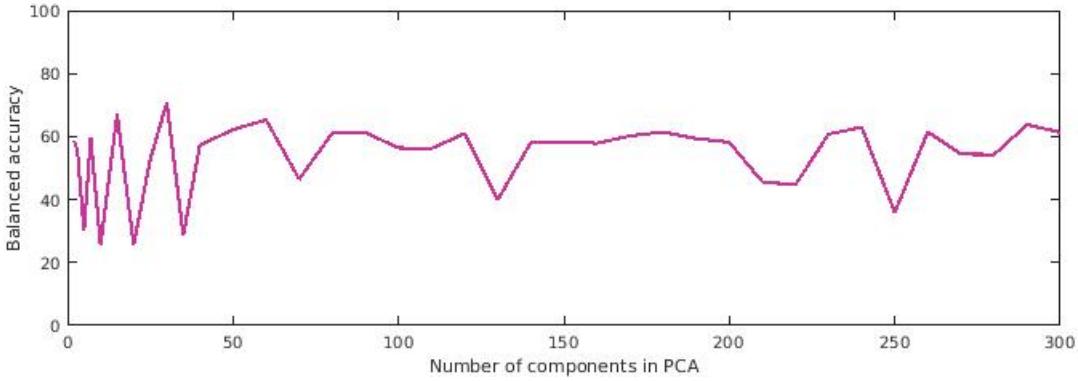
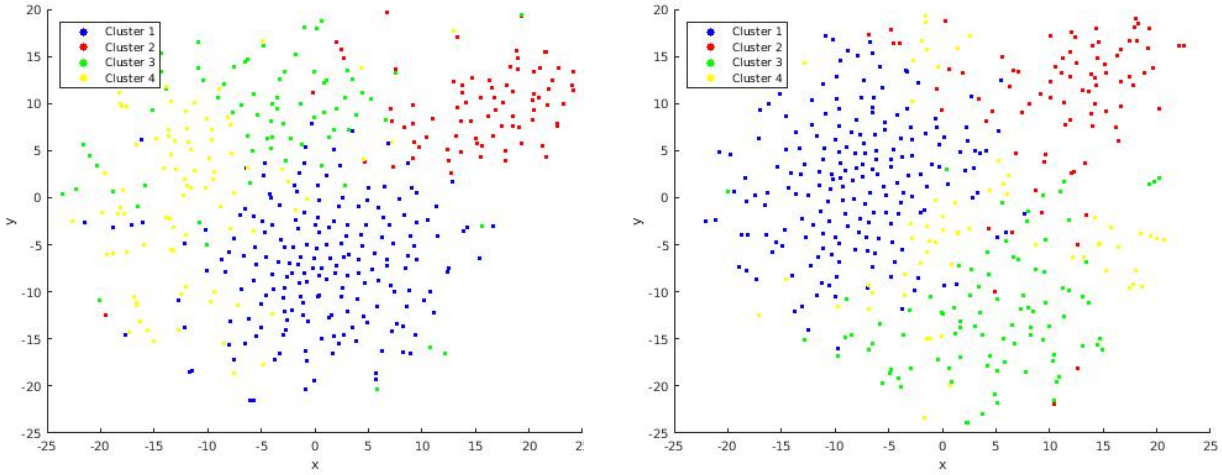


Figure 6.13: Balanced reliability as a function of the number of components in PCA - Multivariate maps

the case of the entire maps. It can be noticed that although they are not located in the approximate same place on the space, corresponding clusters seem to have quite similar shape and size, the correspondence is rather obvious here.



(a) First dataset

(b) Second dataset

Figure 6.14: Clusters in the PCA reduced multivariate maps (cosine distance)

6.3.2.2 Cluster quality

With the results of the previous section, we expect to find an increase in cluster quality compared to the first section about the clustering in the entire multivariate maps.

6.3.2.2.1 Cluster reproducibility For the global and balanced accuracies, we find the following results:

- **Global accuracy**: 55,09%
- **Balanced accuracy**: 70,78%

As expected, the two accuracies values are increased: they were equal to 41,90 and 44,22 for the entire maps). They also surpasses the accuracies values found for the signatures responses, which were equal to 52,08 and 59,67. It is clear that the high dimensionality of the data is responsible for the confusion inside the clusters discovery. Reducing the dimensionality using a simple dimension reduction technique as PCA allows to increase the accuracy of more than 20%.

Figure 6.15 exposes the histogram resulting from the permutation tests investigating the reliability of the balanced accuracy found here above. Once again, the observed value is displayed as a red cross, and it can be seen that it is quite extreme and defined by a rather good p-value. Thus, the null hypothesis is also rejected in this case and we can state that the results are not due to chance.

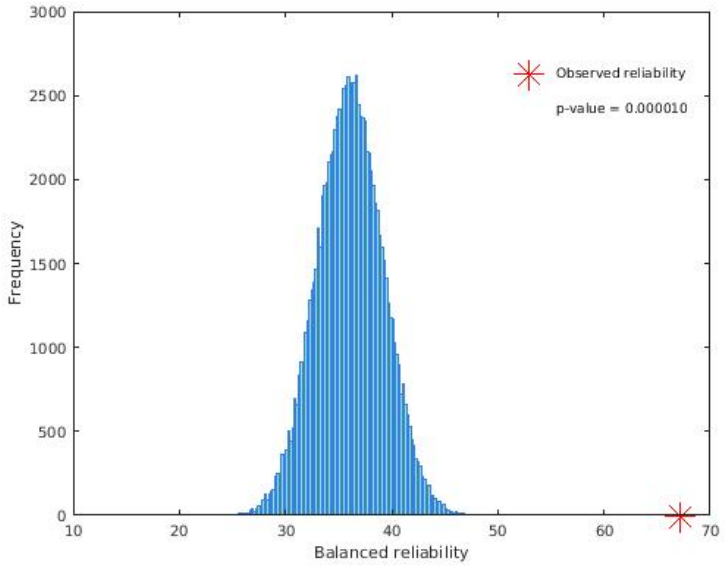
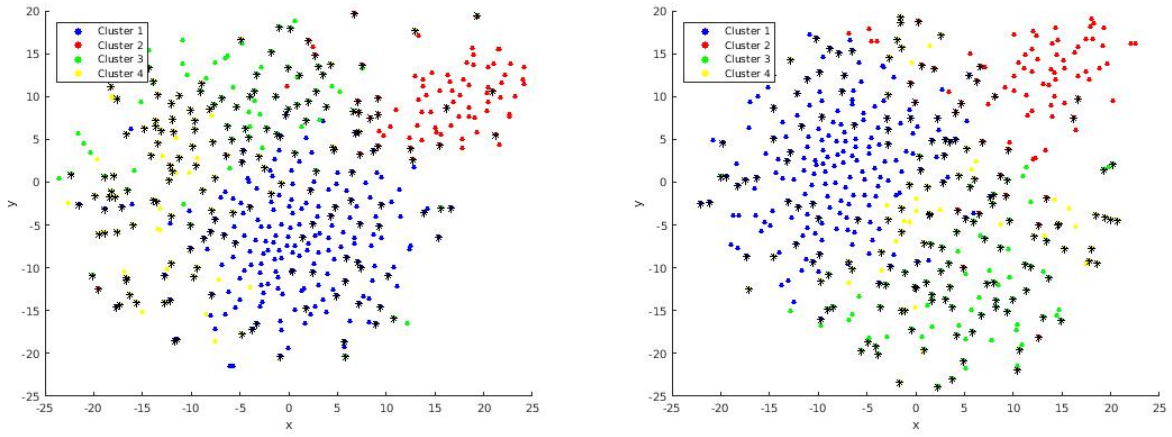


Figure 6.15: Permutation tests for cluster balanced reliability - PCA reduced multivariate maps

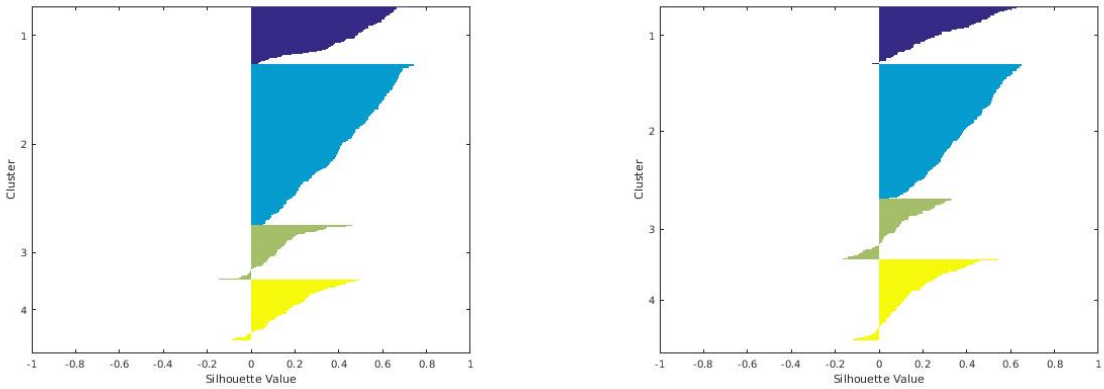
Since the boundaries are more clearly defined, the visualization of the mis-clustered points could be interesting. Figures 6.16a and 6.16b allow that. It can be seen that, even though they are not well delimited, a lot of data points that correspond to the mis-clustered subjects lie on the boundaries between the clusters. Indeed, the centers of each cluster, except for cluster 4 that is much more spread out in the space than the other ones, contains very few mis-clustered objects.

6.3.2.2.2 Cluster consistency The Silhouette values, presented in both ways, also manifest better results than for the clustering on the entire maps. It can be seen in Figures 6.17a and 6.17b that the Silhouette plots are slightly longer than for the other case. Indeed, the mean value is equal to 0,22, where it was equal to 0,02 for the entire maps. Nevertheless, it is still smaller than for the signatures responses. But an explanation would be that the Silhouette value depends on the dimensionality of the data (here it is equal to 30). Indeed, when we compare those plots with the ones found with the signatures responses in Figures 6.10a and 6.10b, from which the data lie in a two-dimensional space, the theory seems to be confirmed. In the following case, where the use of canonical networks is investigated, the number of dimensions will be reduced to a number of 7. If the Silhouette values are still significantly lower than for the signatures responses, an additional confirmation emerges.



(a) First dataset (b) Second dataset

Figure 6.16: Wrong assignments in the PCA reduced multivariate maps (cosine distance)

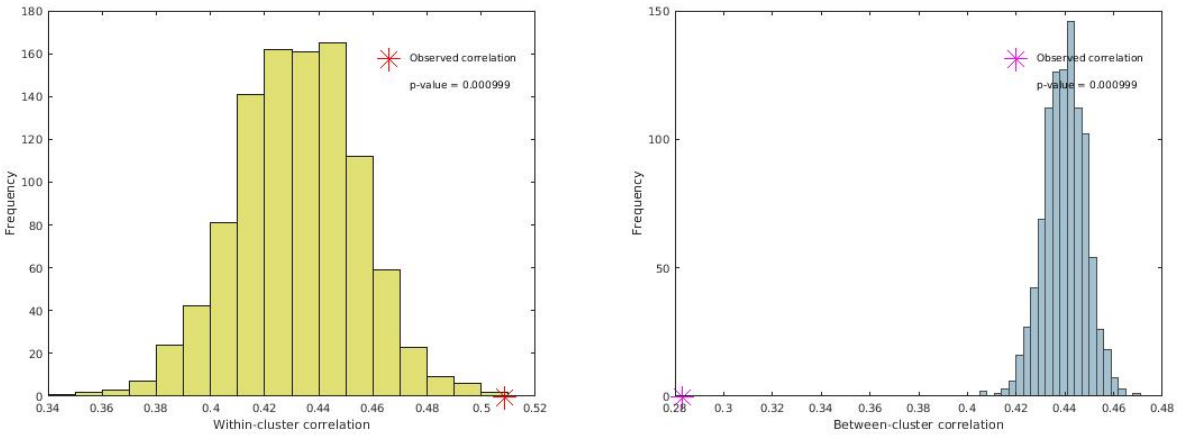


(a) First dataset (b) Second dataset

Figure 6.17: Silhouette plots of the PCA reduced multivariate maps (cosine distance)

6.3.2.2.3 Cluster correlation The within- and between-cluster correlation is investigated in the same way described in section 5.4.1.3 and applied in the signatures responses case. Figures 6.18a and 6.18b exhibit the associated results. The same conclusions as for the signatures patterns can be drawn: the observed within-cluster correlation is high compared to permutations results and linked with a rather good p-value, and the observed between-cluster one is quite small and also defined by a good p-value.

At the end of this section, it was shown that the use of PCA to reduce the dimensionality in the data is a rather good solution to improve the quality of the clusters. For now, both the signatures responses and these PCA reduced maps have exhibited quite interesting outcomes, although they seem to display different behaviors. A final temptation for data dimensionality reduction, the use of canonical networks, is investigated in the next section.



(a) Within-cluster correlation (b) Between-cluster correlation

Figure 6.18: Permutation tests for correlation - PCA reduced multivariate maps

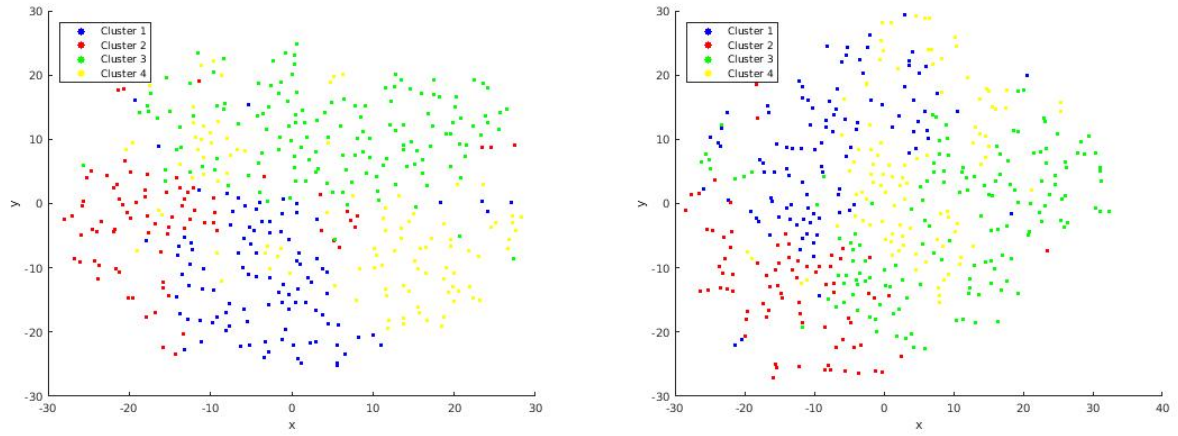
6.4 Canonical networks

The last attempt for dimensionality reduction is the one explained in section 5.1.3, reducing the activation into seven canonical networks corresponding to the Buckner map. The goal is to determine whether or not this technique leads to more accurate results that the two previous ones.

6.4.1 Cosine distance

6.4.1.1 Graphical results

The clusters discovered in this section are displayed in Figures 6.19a and 6.19b. Once again, the clusters are rather well defined, even though they are still overlapping. The simple visualization of the graphs does not allow to make conclusions about the accuracy of the results, the next sections will help doing that.



(a) First dataset (b) Second dataset

Figure 6.19: Clusters in the canonical networks on multivariate maps (cosine distance)

6.4.1.2 Cluster quality

The idea is to compute the different cluster quality metrics in order to compare the results with the PCA reduced maps, since the clusters kind of look alike. The results based on the signatures responses are rather unique and show understandable cluster types, they constitute a kind of separate case. However, for the two other dimension reduction techniques, the visualization with t-SNE is required and there is no way of figuring out the behavior of each cluster based on their representation in a two-dimensional space. Therefore, it is interesting to compare their results in order to choose the method that leads to the best outcomes.

6.4.1.2.1 Cluster reproducibility The two accuracies are computed and we find the following values:

- **Global accuracy**: 31,94 %
- **Balanced accuracy**: 37,63 %

Since the reliability results are quite poor comparing to the previous techniques, the canonical network reduction solution does not seem appropriate for this kind of study. It seems that averaging the activation into a number of cortical networks is not suitable for the search of different and unique subtypes of neural activation, probably because those subtypes are characterized by brain effects which are smaller than the predefined cortical regions. Thus, the analysis is not further described, but the wrong assignments and cluster consistency results can be found in the Appendix C.

Many different results from many different methods have been presented in this chapter. In order to have an organized overview of the different methods results, Table 6.3 was constructed.

	Global reliability	**Balanced reliability**	**Consistency**	**Within-cluster correlation**
Entire	41,90 %	44,22 %	0,02	Not determined
Signatures	52,08 %	59,27 %	0,70	0,49
PCA	55,09 %	70,78 %	0,22	0,51
Networks	31,94 %	37,63 %	0,23	Not determined

Table 6.3: Recapitulative table of the results associated with each data type - Multivariate maps

6.5 Neuroscientific results

In order not to display too many non-meaningful results, the analysis of the neuroscientific results is conducted only on the two methods that gave the best results, i.e. the signatures responses and the PCA reduced maps. The method applied here is described in section 5.4.2. As a reminder, the purpose of this section is to determine whether or not the various clusters actually correspond to different and unique brain activation patterns, and if those patterns are similar between corresponding clusters.

In the different results presented in the following sections, the disposition of the figure is the following: on top of the figure, the montage plot corresponding to the first dataset, i.e. the odd runs, is displayed. In the middle, the montage plot corresponding to the second dataset, i.e. the even runs, is shown. It is then easier to directly visually compare the two plots, to see if they look alike. When they do, a third montage plot is constructed, which results from the search for clusters applied to

the entire dataset (i.e. the entirety of the single-trials are used, for each subject individually, to determine the clusters). Logically, this third montage plot should resemble the two previous ones, and it constitutes the most accurate one, on which the potential further studies should be conducted. To test the reliability of the third montage plot, or more precisely of the third way of searching clusters in the data, we could have used a specific statistical tool called the Spearman-Brown reliability coefficient. It creates a numerical estimation of the change in reliability when the number of items constituting the test is increased. It is useful in the case of a split-half method, to detect whether or not there is an increase in reliability from the test based on half of the items to the one based on all of them [80]. This tool is perfectly adapted to our situation, but has not been implement due to lack of time.

6.5.1 Signatures responses

Figures 6.20, 6.21, 6.22 and 6.23 display the montage plots described above, for each of the four clusters determined. In this case, the significance level q is fixed at 0.05 using False Discovery Rate (FDR). We see that cluster 1 is characterized by a rather high overall activation, at least compared to the three other clusters. If we compare these montage plots with the visualization of the clusters obtained in the signatures responses from multivariate maps, shown in Figure 6.14, we understand that this cluster is the one corresponding to the high responses for both signatures. Thus, it seems logical that it is the one showing the highest activation pattern. Cluster 2 corresponds to the high response for NPS and low for SIIPS. We see that only some regions are activated. Cluster 3 is the one showing high response for SIIPS and low for NPS, showing a little bit more and different activation within the brain. Cluster 4 is the one corresponding to the low activation for both signatures, which is consistent with the results obtained as almost no activation is shown in the montage plot for this cluster.

It can be seen that although clusters 2 and 4 resemble a bit, the four clusters seem to be associated with four distinct activation patterns. In addition, the corresponding clusters from the two distinct datasets always show similar activation, also similar to the one found with the entire dataset.

As the majority of the activation displayed lies on rather small brain regions, it would be difficult to properly name those regions for someone with almost nonexistent background in the neurological field. The specific section requires the collaboration of an expert in the neuroscience field. Here, the only characteristic highlighted is that the patterns look dissimilar in different clusters.

This section is the part of the work that should be much deeper studied in the future. It reflects the finality of the entire stratification procedure, and the simple visualization of montage plots is not sufficient to draw appropriate conclusions. In addition, a neurological expert's opinion would bring useful information on exactly which brain regions are activated. Various ideas on how to look further into the results presented here are discussed in the last chapter.

6.5.2 PCA reduced maps

The same steps are applied to the clusters found in PCA reduced multivariate maps, and similar figures are obtained. Figures 6.24, 6.25, 6.26 and 6.27 display the results obtained for cluster 1, 2, 3 and 4, respectively. Just as in the previous section, we see that each cluster is characterized with rather unique and different activation pattern, as well as showing quite high similarity between corresponding clusters. We see in these figures that some clusters show negative activation in some portions of the brain, which was nonexistent or negligible in the signatures responses.

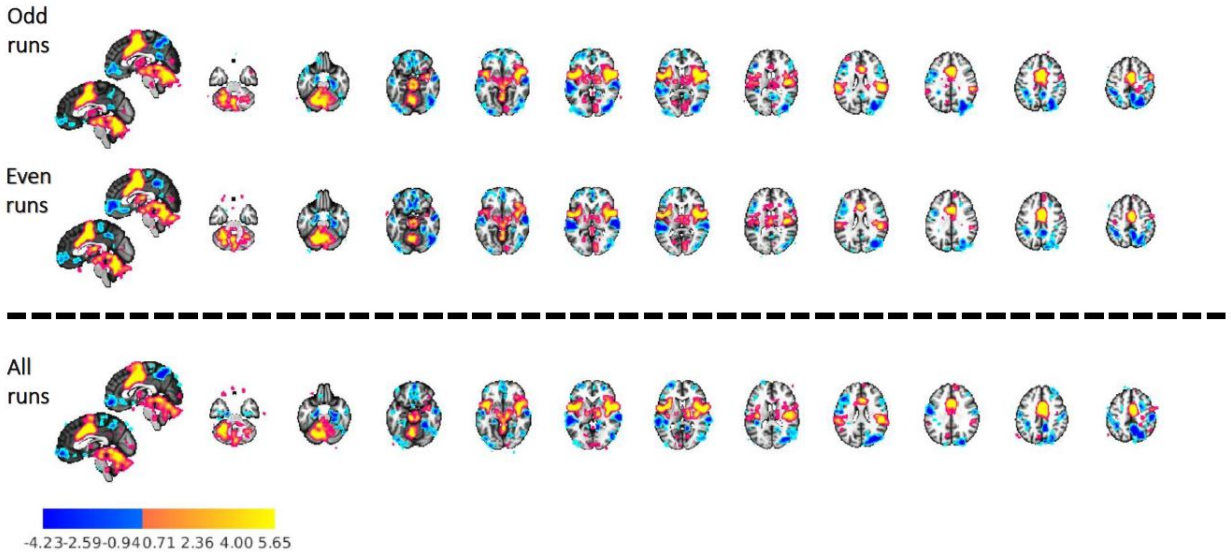


Figure 6.20: Cluster 1 - Signatures responses on multivariate maps

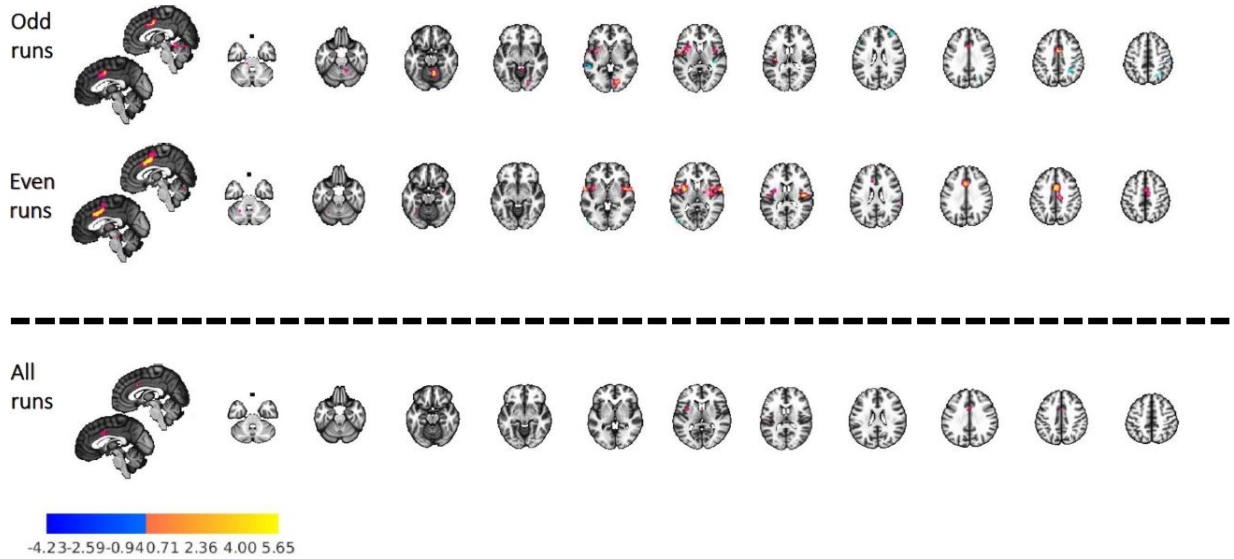


Figure 6.21: Cluster 2 - Signatures responses on multivariate maps

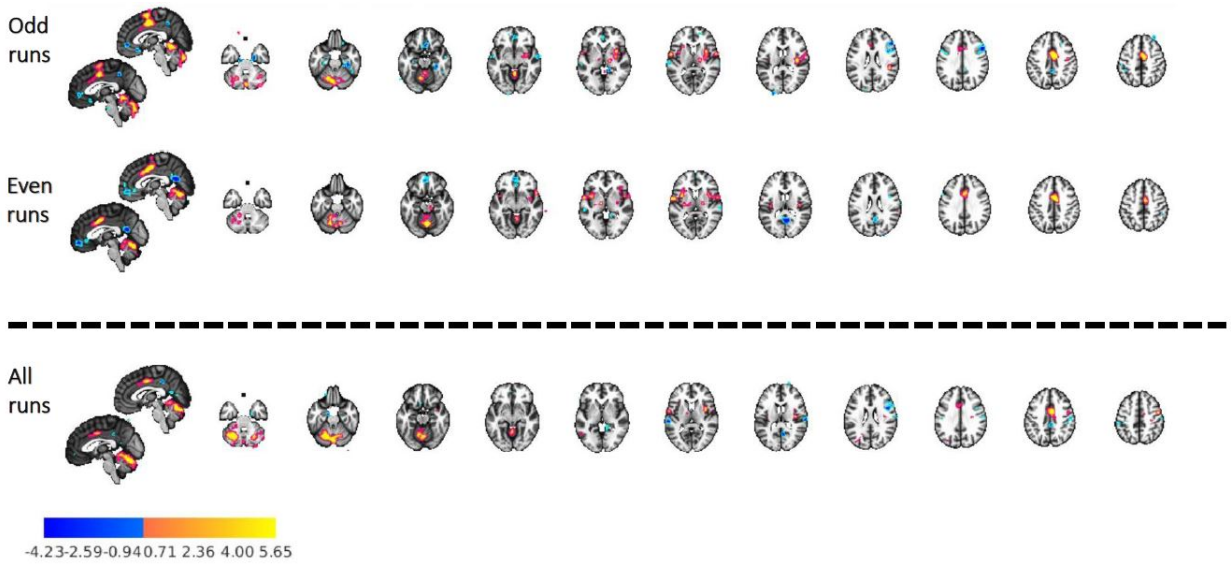


Figure 6.22: Cluster 3 - Signatures responses on multivariate maps

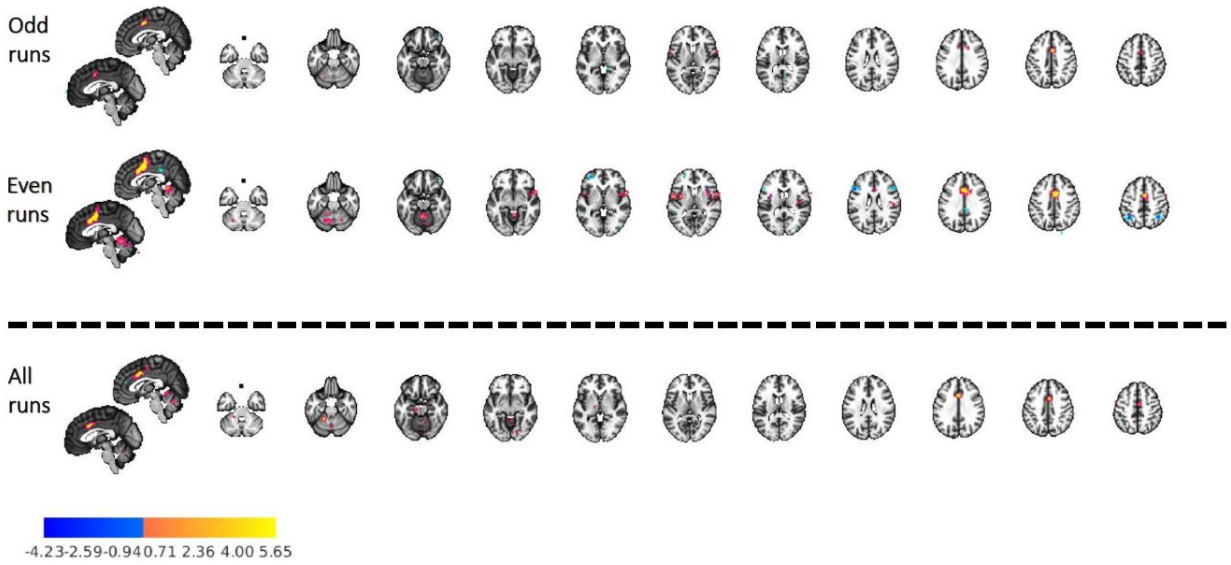


Figure 6.23: Cluster 4 - Signatures responses on multivariate maps

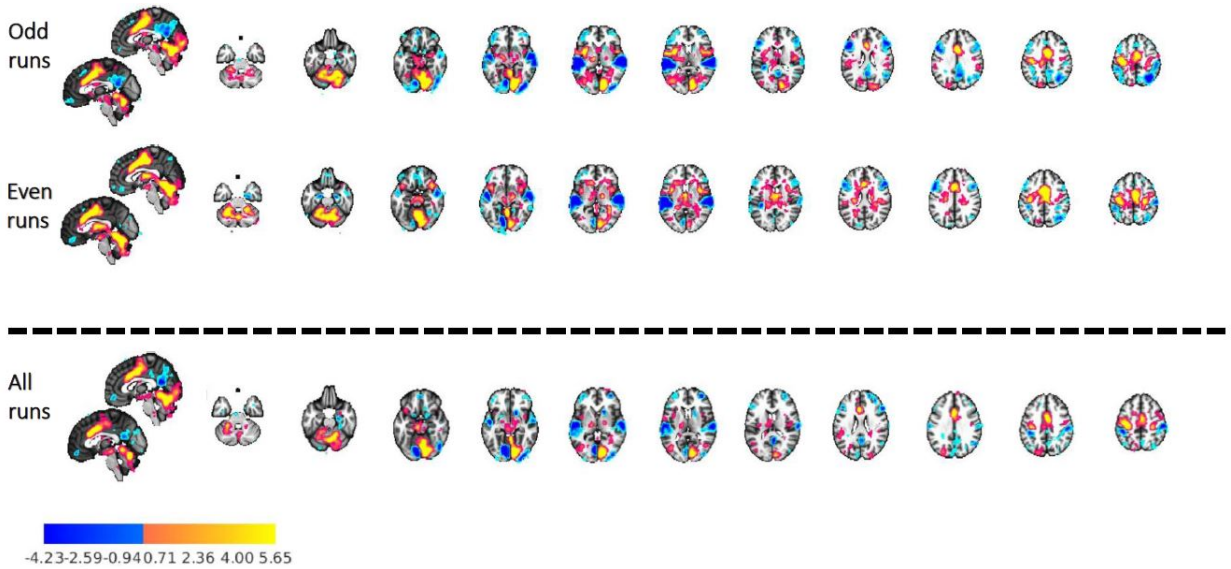


Figure 6.24: Cluster 1 - PCA reduced multivariate maps

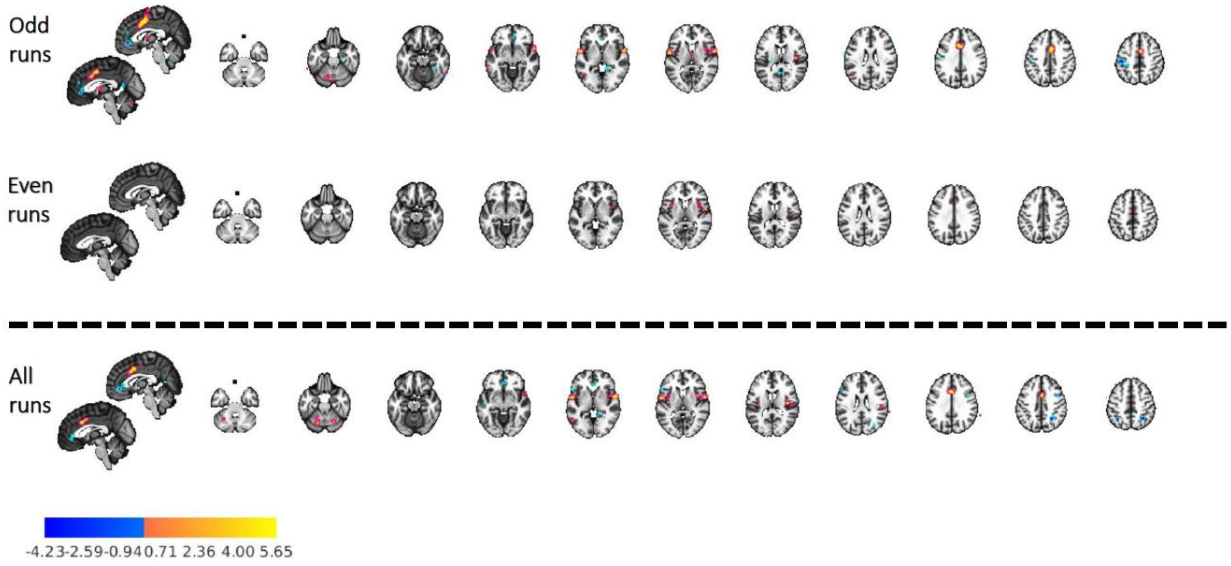


Figure 6.25: Cluster 2 - PCA reduced multivariate maps

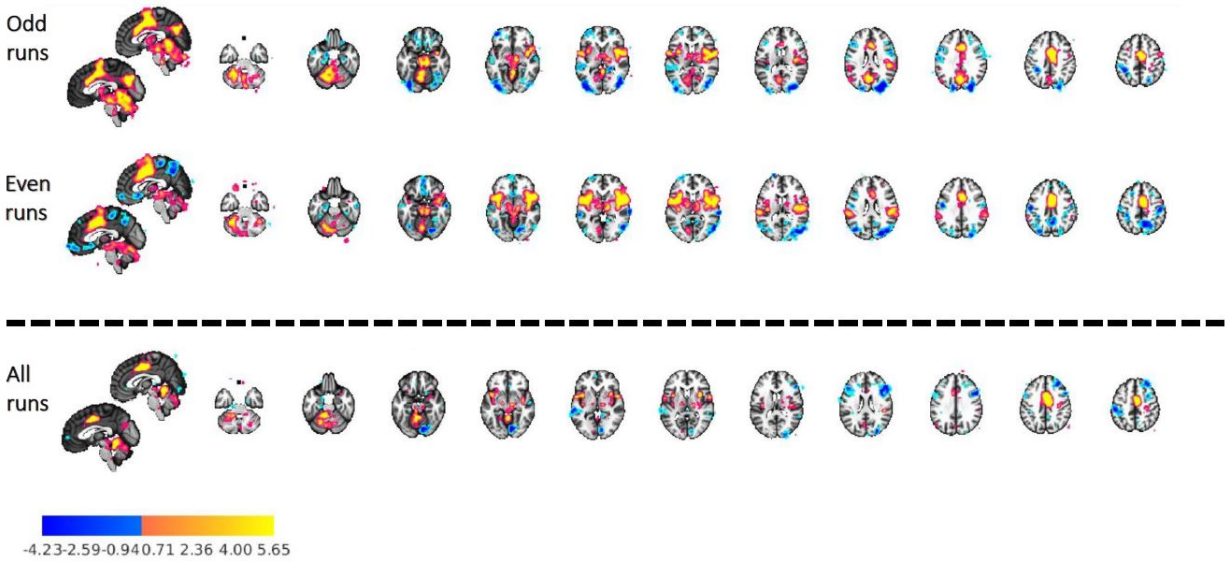


Figure 6.26: Cluster 3 - PCA reduced multivariate maps

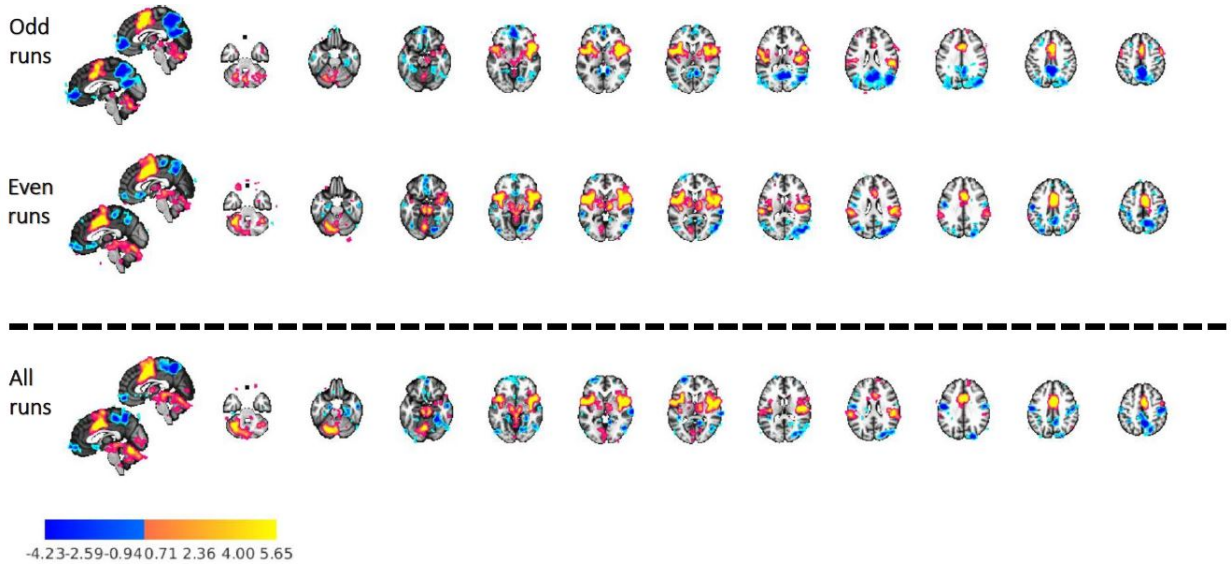


Figure 6.27: Cluster 4 - PCA reduced multivariate maps

6.6 Determination of the optimal number of clusters

The last step in the analysis of multivariate maps is to determine which number of clusters in the most appropriate given the data type. Indeed, all the steps were conducted with the idea of discovering a number of four clusters, mainly because the different behavior wished to be discovered in the case of the signatures responses was divided into four subtypes. However, we have no proof that this number of clusters is the one giving the most stable results.

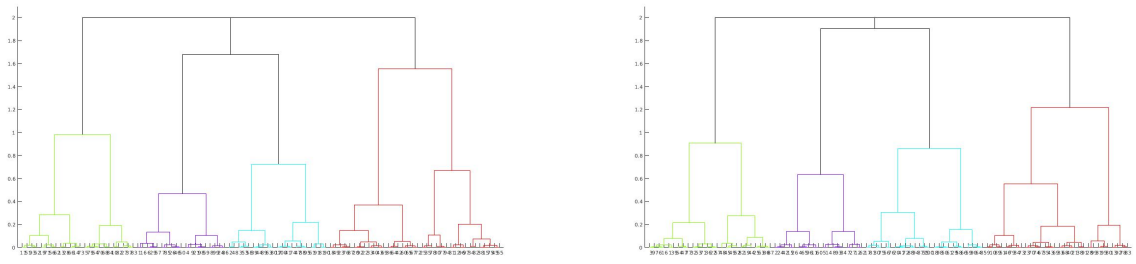
In this section, we determine the optimal number of clusters by the means of three techniques: first with hierarchical clustering and its visualization in the shape of a dendrogram, then with the computation of the balanced accuracy for various numbers of clusters, and finally with the computation of the mean Silhouette value for various numbers of clusters.

6.6.1 Based on dendrograms

The theory linked to the hierarchical clustering and the dendrograms appears in section 5.5.1.
In this technique, as well as in the two following ones, the search for the optimal number of clusters is conducted on the signatures responses and the PCA reduced maps, since they constitute the data types that led to the most accurate results.

6.6.1.1 Signatures patterns

The dendrograms corresponding to the signatures responses (for both datasets) appear in Figures 6.28a and 6.28b. In these graphs, the number of objects in the dataset under study, i.e. the participants, is reduced to a number of 100 for visualization. It means that the object numbers on the horizontal axis represent either one object alone, or a number of objects that have already been joined into one cluster. With these figures, it can be understood that dividing the data into four clusters, corresponding to the four color codes, is quite accurate. However, dendrograms cannot show clearly which division is optimal given the dataset under study. This task is performed with the two other techniques.



(a) First dataset

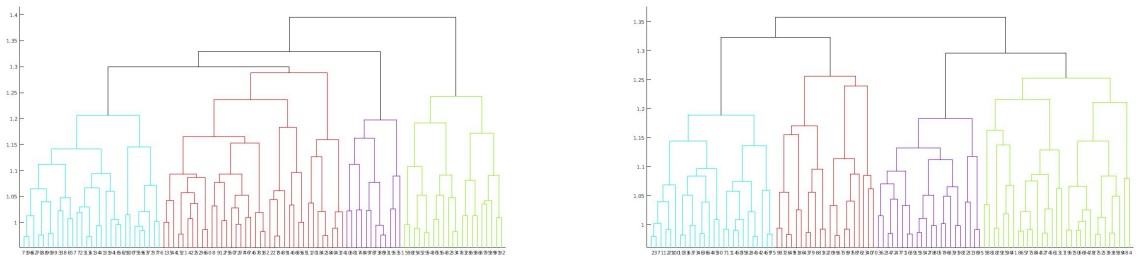
(b) Second dataset

Figure 6.28: Dendrograms - signatures responses on multivariate maps (cosine distance)

6.6.1.2 PCA reduced maps

The same conclusions can be made in the case of the hierarchical clustering on the PCA reduced maps, as it can be seen in Figures 6.29a and 6.29b. Although, the division into four clusters is not as striking as in the previous case, but this is another consequence of the high dimensionality of the

data.



(a) First dataset (b) Second dataset

Figure 6.29: Dendrograms - PCA reduced multivariate maps (cosine distance)

One other conclusion that can be drawn from the construction of dendrograms is that it supports the choice of the cosine distance as the distance metric. Indeed, Figure 6.30 displays the dendrogram obtained when using Euclidean distance, and it can be seen that there are no obvious optimal division within the data, unlike the case of the cosine distance.

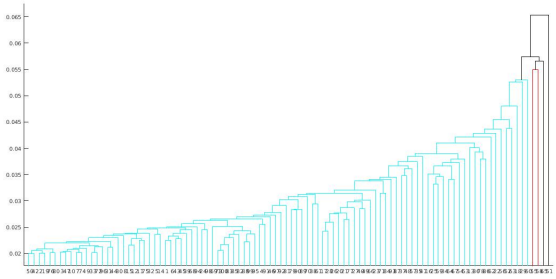


Figure 6.30: Dendrogram - PCA reduced multivariate maps (Euclidean distance)

6.6.2 Based on reliability measures

To generate the graph from Figure 6.31, the entire method for the discovering of clusters and their comparison is initiated with numbers of clusters going from 2 to 15. In each case, the balanced accuracy described in section 5.4.1.1.4 is computed and the following graph is constructed. The conclusion that can be drawn from this figure is that the cluster reproducibility decreases with the number of clusters. It seems logical when we take into account the closeness of the data points as it has been shown in a lot of previous figures. Indeed, most of these figures do not display any "natural" clusters and their detection is forced by the k-means algorithm. Thus, a smaller division of these data points is always more appropriate.
However, in this figure, the balanced accuracy linked with a number of four clusters is quite acceptable and this choice of division used in the entire work seems rather appropriate.

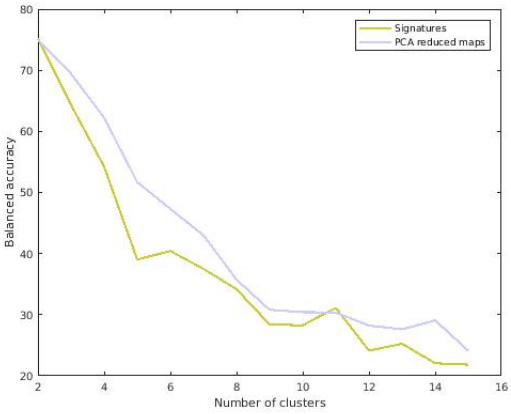
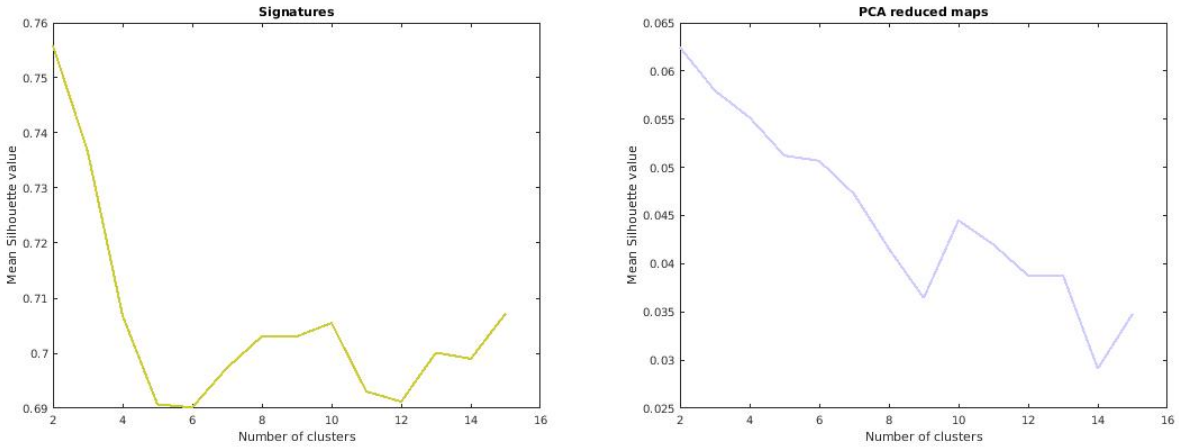


Figure 6.31: Balanced accuracy as a function of the number of clusters

6.6.3 Based on Silhouette values

Figures 6.32a and 6.32 imply the same conclusions as in the previous section.



(a) Signatures responses

(b) PCA reduced maps

Figure 6.32: Mean Silhouette value as a function of the number of clusters

Chapter 7

Stratification using univariate maps

Contents

For the analysis of the univariate maps, only the two data types leading to the best results are investigated: the signature responses and the PCA reduced univariate maps. Actually, the two other data types, i.e. the entire univariate maps and the canonical networks, were studied, but they gave similar results than for the multivariate maps. Therefore, in order no to overload this report, the results are not shown.

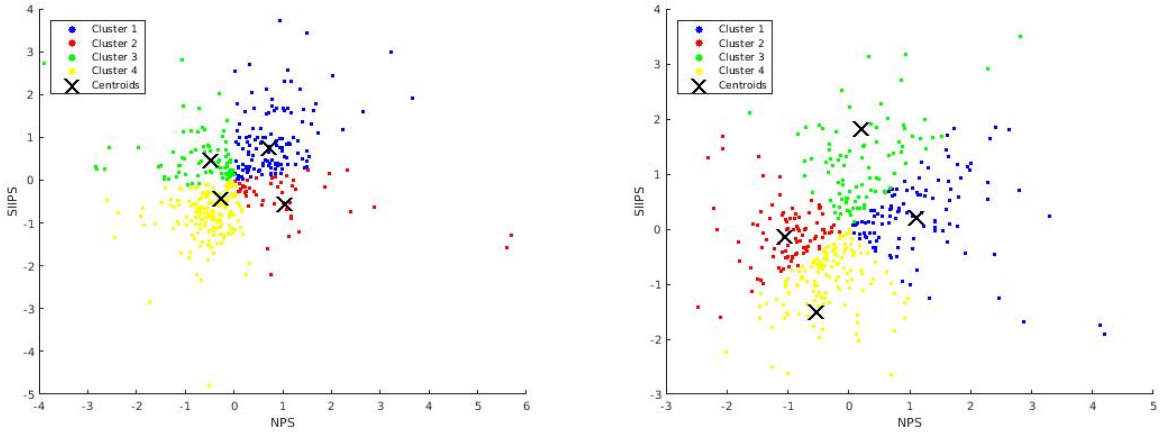
Also, there is no investigation of the difference in distance metric. In this chapter, the distance used is the cosine distance. Once again, the research with Euclidean distance has been conducted but it did not lead to interesting results, and so they are not presented here.
For the two data types under study, the different steps are exactly the same as explained in Chapter 5.

7.1 Signatures patterns

The first dimension reduction technique used is the application of the two signatures NPS ans SIIPS. It should be noted that the pre-developed signatures patterns are multivariate, and they are applied on univariate data. Keeping this in mind, we expect to find poorer results than in the case of the multivariate maps.

7.1.1 Cosine distance

Figures 7.1a and 7.1b display the clusters obtained in the signatures responses applied to univariate maps. While the first one seems to reliably separate the dataset into the clusters types previously mentioned in section 3.1.2.3, the second one is quite different. In addition to not being able to display the expected results, the correlation between corresponding clusters is unlikely to reach acceptable values. Actually, when looking closely at those graphs, we see that the correspondence between clusters determined by the process described in section 5.4.1.1.2 does not make any sense: for example cluster 2, the red one, appears in totally different locations in both graphs.



(a) First dataset

(b) Second dataset

Figure 7.1: Clusters in the signatures responses on univariate maps (cosine distance)

7.1.1.1 Cluster quality

7.1.1.1.1 Cluster reproducibility Due to the configuration issue described here above, we would expect to find quite poor accuracies for the cluster reproducibility:

- **Global accuracy**: 42,82 %
- **Balanced accuracy**: 38,21 %

Both the global and balanced accuracies are lower than in the case of the multivariate maps, which were equal to 52,08 and 59,27, respectively. The histogram constructed with permutation tests based on the balanced reliability is shown in Figure 7.2. We can see that the observed outcome seems to lie in the middle of all the permutations values, and it is linked with a very high p-value. In this case, the null hypothesis cannot be rejected and the results might have obtained by chance. It seems logical when we observe the clusters in Figures 7.1a and 7.1b to notice that they have very dissimilar shapes.

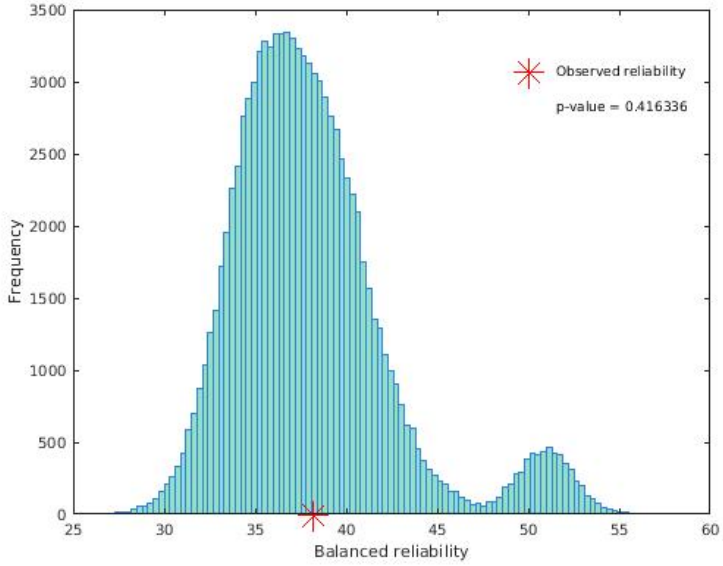


Figure 7.2: Permutation tests for cluster balanced reliability - Signatures patterns on univariate maps

However, if we execute the method for comparing the signatures responses directly, which procedure is described in section 5.4.1.1.5, we find a coefficient $r = 0,49$. As a reminder, the coefficient which was found in the case of the multivariate maps was $r = 0,23$. This correlation, computed between the difference in signatures from both datasets, is shown in Figure 7.3.

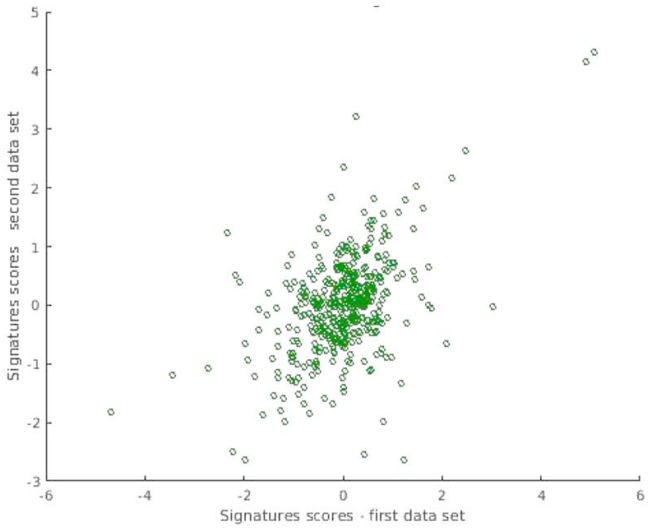


Figure 7.3: Correlation between signatures scores - Univariate maps

These last two results imply that univariate maps are not an adapted tool for the search of clusters in the signatures responses; using multivariate maps seems more appropriate for this task. However, it seems that univariate maps are able to determine more reliably how different subjects track one signature more than the other. Indeed, there was almost no correlation in the multivariate maps results. Since the principal purpose of this project is to investigate the presence of clusters in the data, the analysis of signatures responses in univariate maps is no further extended.

The wrong assignments, Silhouette plots and permutation tests histograms for within- and between-cluster correlations are shown in the Appendix C.

7.2 PCA reduction

As mentioned in section 6.3.1, the search for the optimal number of components in the PCA reduction must be performed for the univariate maps. Indeed, the results found for the multivariate maps were too unstable to make a decision that is likely to work in all cases. This is done in the next section.

7.2.1 Optimal number of components

Figure 7.4 shows how the balanced reliability varies with the number of components in the PCA method. The instability of the curve is as bad as the case of the multivariate maps, the decision for the optimal number is also difficult. Here, the number corresponding to the maximum is 220, but the computation time could not allow to test every number of components between 1 and 300. Thus, it is very likely that one of the untested number could lead to a better result than the ones present in this figure.

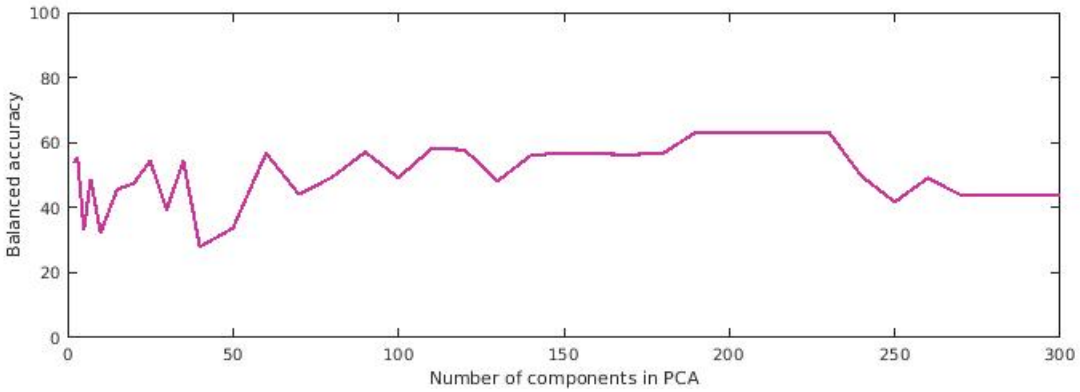


Figure 7.4: Balanced reliability as a function of the number of components in PCA - Univariate maps

7.2.2 Cosine distance

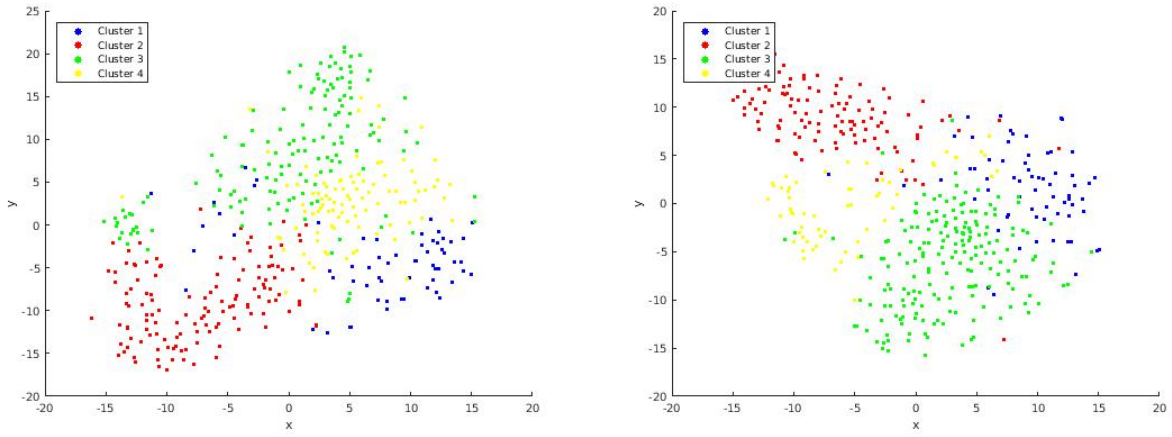
Once again, only the cosine distance is tested.

7.2.2.1 Graphical results

Figures 7.5a and 7.5b display the clusters found in the PCA reduced unviariate maps. In these figures, the four clusters seem rather well delimited from each other. Thus, it is interesting to continue the analysis in order to make a decision on which type of maps (multivariate or univariate) is the most appropriate in this case. We have already concluded that the use of signatures patterns is more suited for multivariate maps.

7.2.2.2 Cluster quality

7.2.2.2.1 Cluster reproducibility Just as all the previous sections, we start by assessing the cluster quality with two measures of cluster reproducibility:



(a) First dataset (b) Second dataset

Figure 7.5: Clusters in the PCA reduced univariate maps (cosine distance)

- **Global accuracy**: 58,10 %
- **Balanced accuracy**: 63,25 %

The global accuracy is slightly higher than for the multivariate maps (55,09) while the balanced accuracy is slightly lower (70,78). The histogram obtained with permutation tests based on balanced reliability is shown in Figure 7.6. In this case, the p-value is quite high (around 15%), which means that the null hypothesis cannot be rejected in this case.

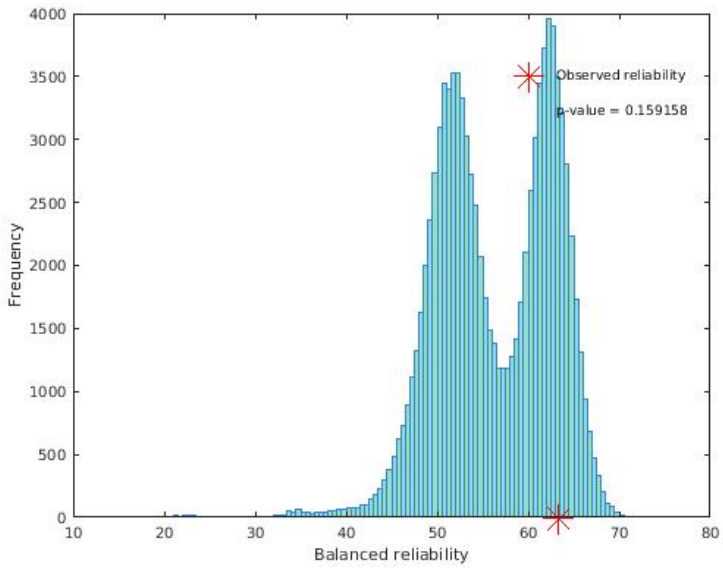
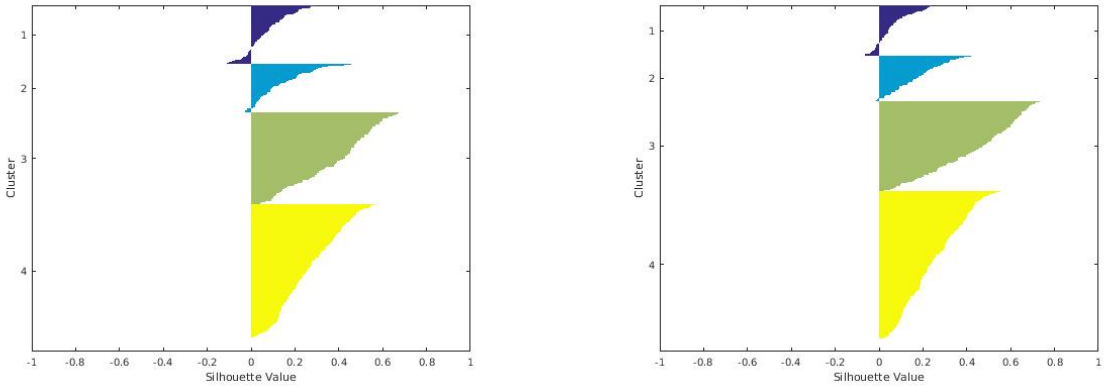


Figure 7.6: Permutation tests for cluster balanced reliability - PCA reduced univariate maps

Even if the balanced reliability is not trustworthy, the analysis is pursued, but with some precautions in the conclusions following the results.

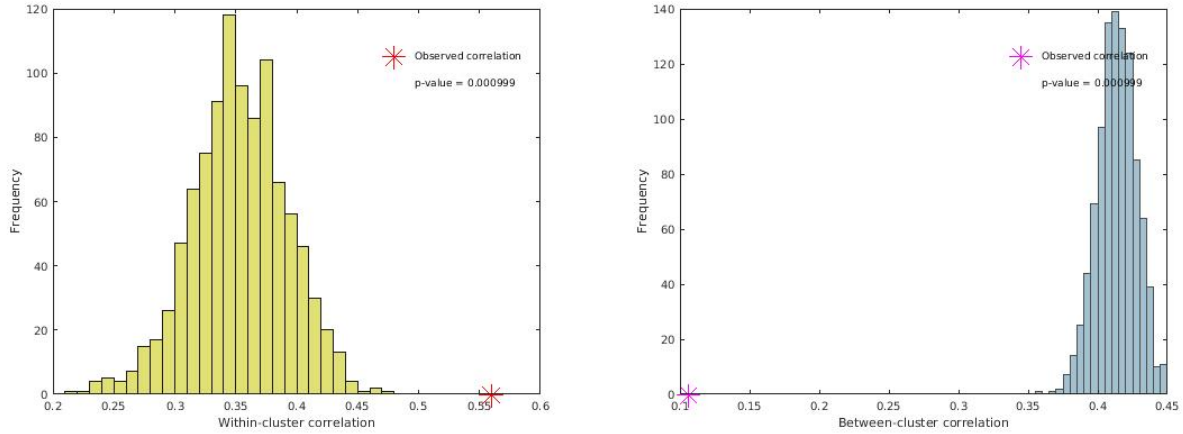
7.2.2.2.2 Cluster consistency In this section, we investigate the cluster consistency based on the Silhouette values. Figures 7.7a and 7.7b display the Silhouette values for each data point in each cluster, and it can be seen that the bars are not very long, but they seem acceptable. The mean value is equal to 0,20 in this case, which is similar to the one found in PCA reduced multivariate maps.



(a) First dataset (b) Second dataset

Figure 7.7: Silhouette plots of the PCA reduced univariate maps (cosine distance)

7.2.2.2.3 Cluster correlation Figures 7.8a and 7.8b exhibit the histograms resulting from the permutation tests studying the within- and between-cluster correlations. In contrary to the histogram obtained for the balanced reliability, here the p-values are very small and the observed correlations are rather extreme compared to the permutations results.



(a) Within-cluster correlation (b) Between-cluster correlation

Figure 7.8: Permutation tests for correlation - PCA reduced univariate maps

A similar recapitulative table is constructed for the univariate maps stratification, in Table 7.1.

	Global reliability	Balanced reliability	Consistency	Within-cluster correlation
Signatures	42,82 %	38,21 %	Not determined	Not determined
PCA	58,10 %	63,25 %	0,20	0,57

Table 7.1: Recapitulative table of the results associated with each data type - Univariate maps

7.3 Neuroscientific results

In the case of the univariate maps, only the results obtained with PCA reduced maps are investigated in this section. Indeed, we have shown that the clusters found in the signatures responses are not worth being more deeply analyzed.

7.3.1 PCA reduced maps

Figures 7.9, 7.10, 7.11 and 7.12 expose the montage plots obtained for each cluster discovered in the PCA reduced univariate maps. Once again, it can be seen that corresponding clusters are visually similar, and different clusters seem to display different neural activation pattern. In this case, the q rate for FDR is set to 0.000005. The four distinct activation patterns show very different behaviors in this case, much more than for the multivariate maps. Indeed, we see that cluster 3 is characterized with a high overall positive activation and cluster 4 with a large amount of negative activation widespread in the brain. Deeper studies should determine whether these activation patterns define some interesting different behaviors in the cluster or if they simply result from noise specific to the univariate way of proceed.

7.4 Determination of the most accurate number of clusters

There is nothing that can prove that the most accurate number of clusters is the same as for the multivariate maps. As a reminder, in that case, it was shown that although this number is not the one leading to the best results, the search for four clusters in the data is rather accurate. Indeed, considering the proximity of the data points, it seems logical that the less clusters discovered, the more accurate.
In this section, we investigate whether or not we can draw similar conclusions with univariate maps. Just as in the case of the multivariate maps, the optimal number is examined first with hierarchical clustering displayed as a dendrogram, and then with cluster quality measurements, i.e. balanced reliability and Silhouette value.

7.4.1 Based on dendrograms

As said here above, the first attempt to determine this optimal number is by using hierarchical clustering. We present the results only for the PCA reduced maps.

7.4.1.1 PCA reduced maps

Figures 7.13a and 7.13b, presenting those dendrograms, allow to imagine how the four clusters can be constructed, as it was the case for the multivariate maps. Thus, the same conclusions can be made in this case.

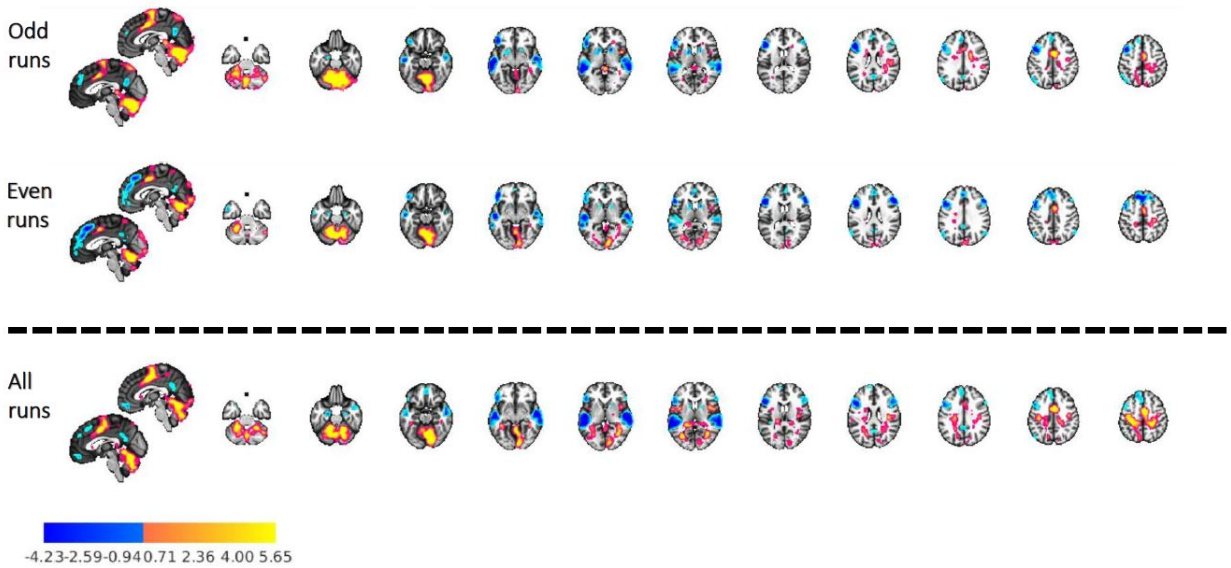


Figure 7.9: Cluster 2 - PCA reduced univariate maps

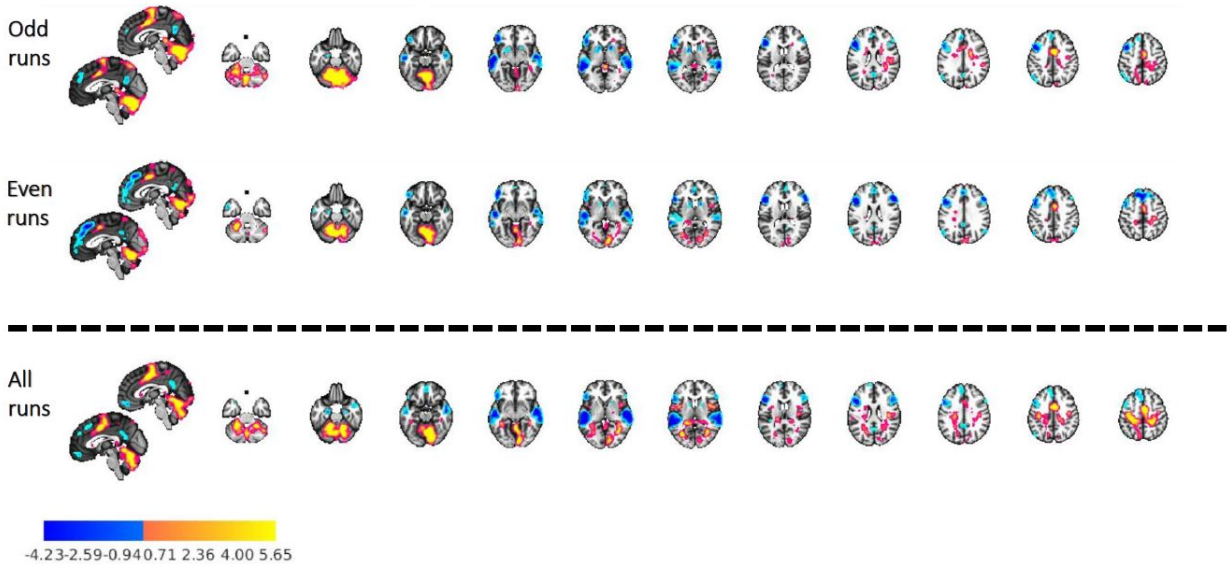


Figure 7.10: Cluster 2 - PCA reduced univariate maps

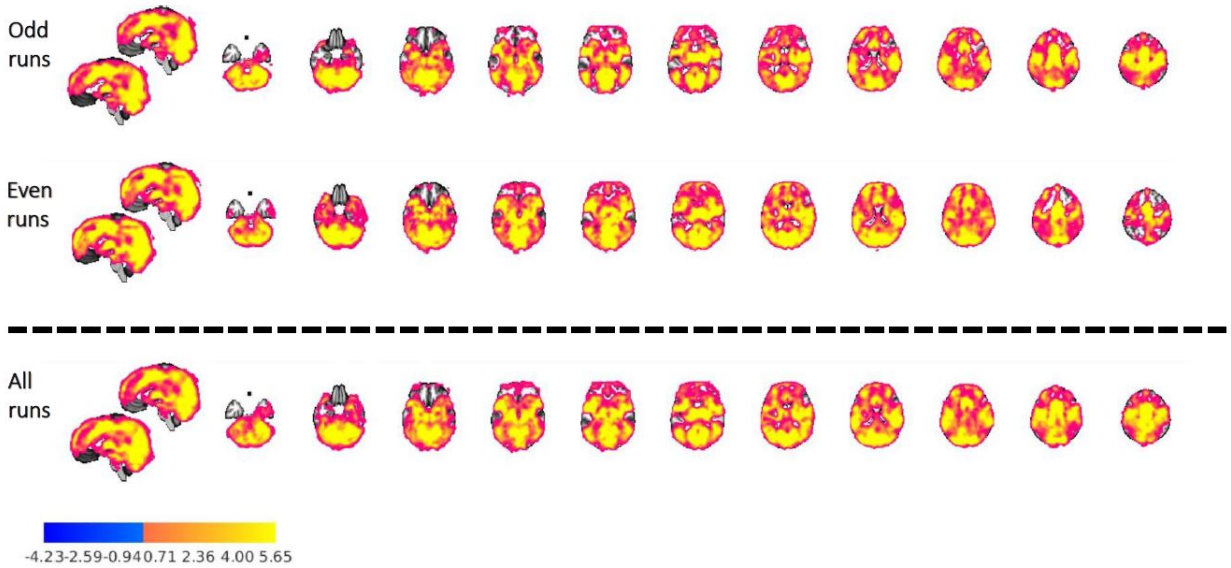


Figure 7.11: Cluster 3 - PCA reduced univariate maps

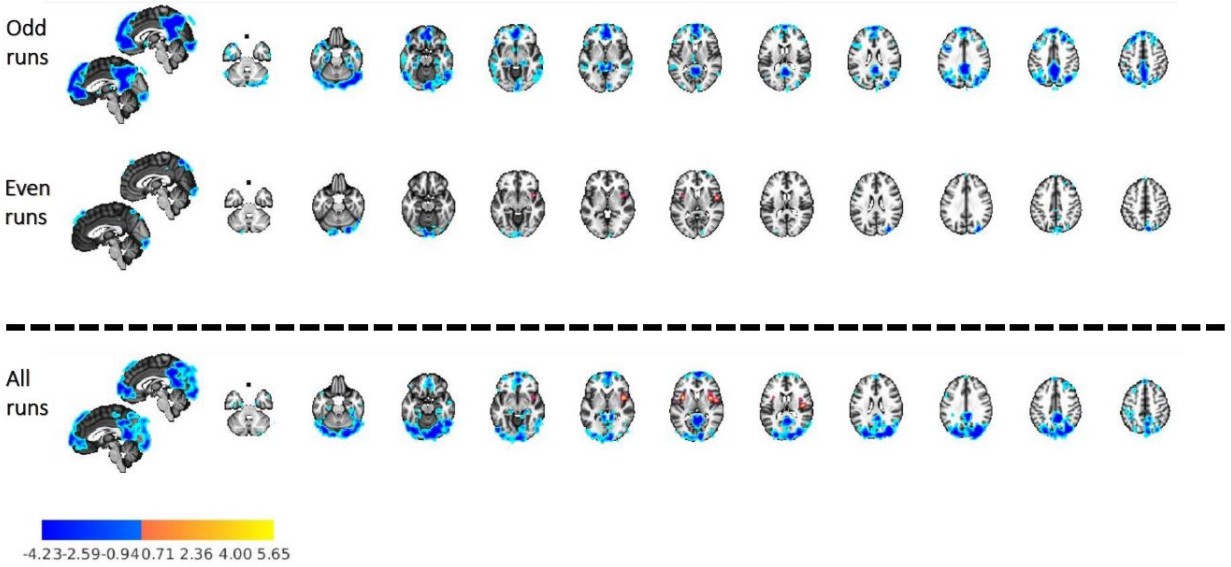
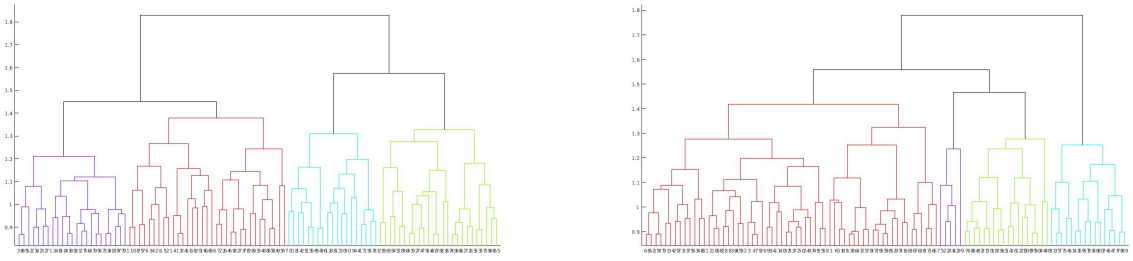


Figure 7.12: Cluster 4 - PCA reduced univariate maps



(a) First dataset (b) Second dataset

Figure 7.13: Dendrograms - PCA reduced univariate maps (cosine distance)

7.4.2 Based on reliability measures

Here, the results were computed in the same way of the multivariate maps, thus the case of the signatures responses is investigated. The balanced reliability measures imply the same conclusion as for the multivariate maps: in Figure 7.14, we can see that for both data types, the reliability decreases with the number of clusters, but the value linked with number four is still quite acceptable.

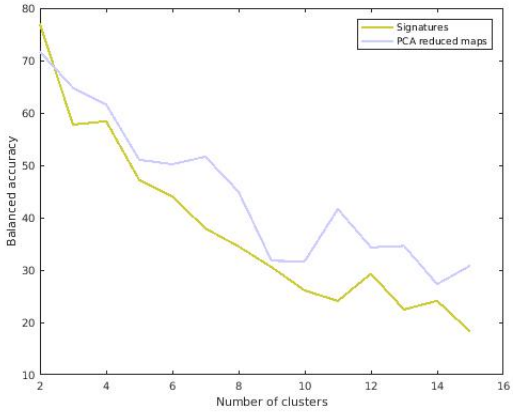
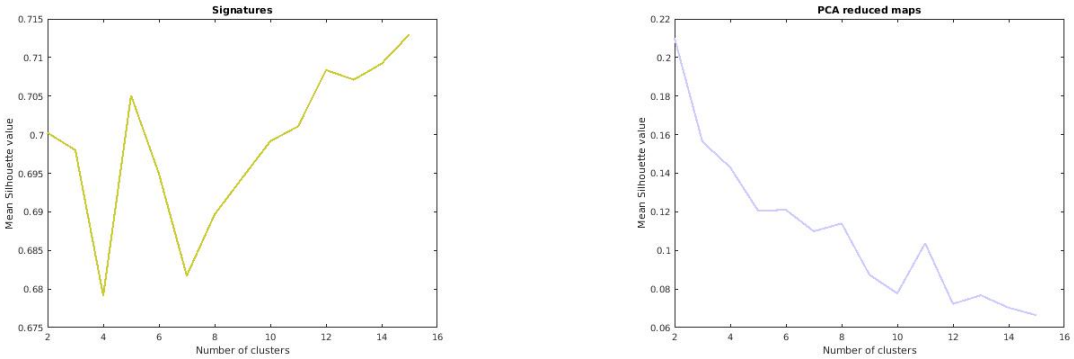


Figure 7.14: Balanced accuracy as a function of the number of clusters

7.4.3 Based on Silhouette values

In the case of the Silhouette value, there is one small difference from before: in Figure 7.15a, for the signatures responses, we see that the mean Silhouette value is unstable and finally reached its maximum for a number of 16 clusters. This seems wrong if we take into account the total number of participants, i.e. 433, the discovery of 16 clusters does not seem appropriate. However, it is consistent with the quite bad results we have shown for this data type on univariate maps.
In the case of the PCA reduced maps, displayed in Figure 7.15, we have similar results than the previous cases. With all that, we consider once more that the discovery of four clusters is a good solution in these data.



(a) Signatures responses (b) PCA reduced maps

Figure 7.15: Mean Silhouette value as a function of the number of clusters

7.5 Compare all clusters

At the end of the analysis of the various techniques and data types, it seemed interesting to investigate whether or not the clusters found correlate with one another, and see if the same subjects form the same clusters in two different data types. For example, it would be interesting to compare the clusters found for the signatures responses from multivariate maps and univariate maps, or to compare clusters coming from the same type of maps but with different dimension reduction techniques (e.g. investigate the correlation between the clusters found in the signatures responses and the ones found in the PCA reduced maps, both in the multivariate maps). The results appear in Table 7.2. The correlation between all clusters is rather small, and never exceeds 50%. This is probably due to the fact that using univariate approach is not appropriate in this case. Also, when observing the neuroscientific results in the multivariate case, the four unique brain activation patterns found in the signatures responses were highly dissimilar from the ones found in the PCA reduced maps. It seems that both of the techniques are adapted, but they do not reflect the same behaviors to pain sensation.

	Multivariate	**Univariate**	**Multivariate V.S. Univariate**
Signatures	59,27 %	38,21 %	41,86 %
PCA	70,78 %	63,25 %	45,08 %
Signatures V.S. PCA	42,63 %	46,56 %	-

Table 7.2: Balanced accuracy between all clusters

Chapter 8

Conclusion and perspectives

The goal of the finality of this whole project is to bring some additional information about the neural basis of pain to any study that could be conducted in the future, in the field of pain or even any other field that could think of these findings as useful.

At the end of this project, the search for clusters in the data have been conducted for different ways of expressing the data, and the results showed that some of those expressions lead to quite accurate outcomes, while others do not. Two main types of weight maps have been tested, based on a multivariate approach on one side and on an univariate one on the other. For each approach, four ways of expressing the data were examined: the entire maps, the signatures responses, the maps reduced with PCA and the maps reduced to seven canonical networks. The first and last ones gave quite poor results, so they were put aside almost immediately. In the case of the entire maps, it is believed that the poor results are due to the very high dimensionality of the data. On the other hand, for the reduction into seven canonical networks, it is probably due to the fact that the differences in brain activation in the different clusters are not contained inside those networks, but presumably delimit more precise dissimilarities. The two remaining data types allowed to define four reliable clusters with unique brain activation patterns. However, those patterns are rather distinct in the two situations. Since the clustering performed on signatures responses is the result of pre-defined patterns, it seems logical that the activation maps obtained reflect the different behavior of these signatures. Indeed, the four clusters correspond to the four following situations: subjects having high activation for both signatures, subjects having high activation for NPS and low activation for SIIPS, subjects having high activation for SIIPS and low activation for NPS, and subject having low activation for both signatures. The visualization of the brain patterns seems consistent with the cluster's characteristics. In the case of the PCA reduced multivariate maps, performed with 30 components after investigation of the optimal number based on reliability measurements, the different clusters seem to be associated with different patterns, which are also distinct from the signatures responses patterns. The behavioral explanation of each activation pattern is less clear than for the signatures responses, implying that further analysis should be added in this case.

The same steps were conducted on the univariate weight maps in order to compare with the multivariate approach, and finally make a decision on which type is the most appropriate. In the signatures responses, the clusters localization in space was inconsistent with what was expected, in addition to be unreproducible. Thus, the multivariate approach seems more accurate in this case. However, another way of assessing the signatures tracking for each individual was tested, based on the signatures responses coordinates in the two-dimensional space. This method was only interesting for the univariate weight maps, it was shown that the participants track more reliably one signatures more than the other. This constitutes an interesting outcome, but as the main goal of

this project is the stratification into several clusters, it was not deeper investigated. In the case of the PCA reduced univariate maps, the different measures of cluster quality were not as good as in the multivariate case. The visualization of activation patterns was still conducted, and showed even more dissimilar behaviors: for example, we find a cluster with high overall positive activation, and one with high overall and exclusive negative activation. This could establish interesting conclusions, but the poor cluster quality results imply to stay vigilant with this type of data.
All in all, multivariate approach seems to compose the most appropriate choice in the search of clusters in single-trials data related to pain.

As said in section 6.5.1, the portion of the work which needs to be much deeper investigated is the neuroscientific results. Indeed, the simple visualization of the brain patterns as they are presented in that section is not sufficient to draw accurate conclusions. In addition, the results presented in this report seem to be subject to some noise. One idea to construct more accurate activation maps, which was not implemented due to lack of time, is the searchlight analysis. It is a multivariate pattern analysis (MVPA) tool developed specifically for fMRI studies, used to identify locally informative regions in the brain. It was developed in the idea of pursuing two main goals: first being able to identify small brain regions, which is often a requirement in fMRI studies, but also taking into account the spatial structure of the BOLD signal, i.e. the fact that adjacent voxels usually have similar activation timecourses. The result of searchlight analysis is a map constructed by the measurements of information, i.e. activation, in small spherical subsets centered in each voxel, as an imaginary actual searchlight [78].

On these activation maps, numerical tools might be used to detect actual differences across the cluster related activation patterns. The idea is to employ a One-way ANOVA technique, which is an analysis of variance tool. This technique is able to detect a statistically significant difference between the means of one or more independent groups [79]. In this project, it will be used to assess the difference in brain activation across the four distinct clusters discovered in the data. The collaboration of a neuroscience expert is also crucial for an accurate description of the different activation patterns established.

Some other methods might be added in several sections of the work. For example, at the beginning of the process, there might exist a way to improve the pain-predictive weight maps directly. One possible solution is to use a tool called Group-Regularized Individual Prediction, or GRIP method [22], which was partly developed by researchers from the CAN Lab. This tool was elaborated in the purpose of overcoming the difficulties related to the predictions making at the individual's level.
The idea behind the development of this method is that brain representations are idiographic, i.e. each individual has a unique brain with specific characteristics and patterns. However, usually, there is information which is shared across individuals, at a larger scale. Knowing that, Lindquist and al. [22] have developed a method where both notions are taken into account: in their paper, they constructed individual weight maps that result from a weighted contribution of both the idiographic and the group-level maps. The weights were derived with two different methods: an Empirical Bayes approach based on within- and between-subjects variances, and a cross-validated approach based on prediction accuracy. In this work, just as the work presented in the paper, the group-level pattern would correspond to the NPS signature described in section 3.1.1.
A further description of the method appears in the Appendix B. Since the NPS pattern is multivariate, this tool would be exclusively used in the case of multivariate maps. The multivariate maps previously constructed correspond to the idiographic maps described in the method. The process of implementing this method was started but due to lack of time and low priority, it was finally put aside. Besides, two opposite outcomes were expected: either the addition of this step would have

helped to stabilize the weights and led to more accurate results, or it would have "forced" the weight maps to resemble a little bit more to the NPS pattern and led to biased results. That is why it did not seem indispensable.

In future studies, the outcomes of this project might be used in several ways. First of all, a separated dataset consisting in the same type of data could be used in order to verify that the findings generalize to new subjects. Indeed, in this project, it was chosen to apply the process on the entirety of the dataset, comprising the 433 participants from the 13 studies, and there was no separation into train and validation sets. The only way to prove that the clusters found in the data are reliable is to verify that similar clusters appear in a totally different dataset, composed of entirely different participants. By doing that, we could actually claim that there exist four different and unique neurological biotypes associated to pain.

To pursue the study even further, psychological measures such as age, sexe, etc. linked to each participant individually could establish a solution for individual prediction. It would be interesting to investigate whether each individual could be assigned to the right cluster based on their psychological scores. Each cluster would be characterized by a psychological profile which will be used to determine the cluster assignment for new unseen subjects.

The different steps applied in this process are easily reproducible for any condition, provided that a certain number of single-trials maps is available for individual subjects, and that each of them is associated with some continuous measurement which can be predicted. The search for clusters reflecting different neurological basis related to a specific medical condition could be applied in various situations different than pain.

Bibliography

[1] Martin A. Lindquist, Tor D. Wager, *Principles of fMRI*, Youtube, August 24, 2015.

[2] Christophe Phillips, *Introduction à la statistique médicale*, ULiège, Year 2016-2017.

[3] D. Purves, G.J. Augustine, D. Fitzpatrick, W.C. Hall, A.-S. LaMantia, L.E. White, *Neurosciences - 5th edition*, Neurosciences et cognition, De Boeck Supérieur, 2015.

[4] Katja Wiech, Markus Ploner and Irene Tracey, *Neurocognitive aspects of pain perception*, Cell Press, 1364-6613/, 2008 Elsevier Ltd. doi:10.1016/j.tics.2008.05.005, July 5, 2008.

[5] Jerry L. Prince, Jonathan M. Links, *Medical imaging signals and systems*, Copyright 2015, 2006 by Pearson Education, Inc., publishing as Prentice Hall.

[6] José M.Soares, Ricardo Magalhães, Pedro S. Moreira, Alexandre Sousa, Edward Ganz, Adriana Sampaio, Victor Alves, Paulo Marques and Nuno Sousa, *A Hitchhiker's Guide to Functional Magnetic Resonance Imaging*, Brain Imaging Methods, a section of the journal Frontiers in Neuroscience, November 10, 2016

[7] John C. Gore, *Principles and practice of functional MRI of the human brain*, Vanderbilt University Institute of Imaging Science, Nashville, Tennessee, USA, PERSPECTIVE SERIES Medical imaging; Joy Hirsch, Series Editor, J. Clin. Invest. 112:4–9 (2003). doi:10.1172/JCI200319010, 2003.

[8] Kerstin Preuschoff, *Physiological Basis of the BOLD Signal*, Social and Neural systems Lab, University of Zurich, `http://www.fil.ion.ucl.ac.uk/spm/course/slides10-zurich/Kerstin_BOLD.pdf`, consulted on April 13, 2018

[9] Thomas T.Liu, *Noise contributions to the fMRI signal: An overview*, NeuroImage 143(2016)141–151, September 6, 2016.

[10] Tor D. Wager, Ph.D., Lauren Y. Atlas, Ph.D., Martin A. Lindquist, Ph.D., Mathieu Roy, Ph.D., Choong-Wan Woo, M.A., and Ethan Kross, Ph.D., *An fMRI-Based Neurologic Signature of Physical Pain*, The New England journal of medicine, N Engl J Med 2013;368:1388-97. DOI: 10.1056/NEJMoa1204471, April 11, 2013.

[11] DRUGBANK, *Remifentanil*, Drug created on June 13, 2005, Last updated on April 17, 2018, `https://www.drugbank.ca/drugs/DB00899`, consulted on April 17, 2018.

[12] Choong-Wan Woo, Liane Schmidt, Anjali Krishnan, Marieke Jepma, Mathieu Roy, Martin A. Lindquist, Lauren Y. Atlas and Tor D. Wager, *Quantifying cerebral contributions to pain beyond nociception*, Nature Communications, DOI: 10.1038/ncomms14211, February 14, 2017.

[13] James V. Haxby, *Multivariate pattern analysis of fMRI: The early beginnings*, NeuroImage, doi:10.1016/j.neuroimage.2012.03.016, March 9, 2012.

[14] Atlas LY, Lindquist MA, Bolger N, Wager TD. *Brain mediators of the effects of noxious heat on pain*, Pain. 2014;155(8):1632-1648. doi:10.1016/j.pain.2014.05.015., May 17, 2014.

[15] Choong-Wan Woo, Mathieu Roy, Jason T. Buhle, Tor D. Wager, *Distinct Brain Systems Mediate the Effects of Nociceptive Input and Self-Regulation on Pain*, PLoS Biology, 13(1):e1002036. doi:10.1371/journal.pbio.1002036, January 6, 2015.

[16] Anjali Krishnan, Choong-Wan Woo, Luke J Chang, Luka Ruzic, Xiaosi Gu, Marina López-Solà, Philip L Jackson, Jesús Pujol, Jin Fan, and Tor D Wager, *Somatic and vicarious pain are represented by dissociable multivariate brain patterns*, eLIFE, 2016;5:e15166. DOI: 10.7554/eLife.15166, June 14, 2016.

[17] Roy M, Shohamy D, Daw N, Jepma M, Wimmer GE, Wager TD, *Representation of aversive prediction errors in the human periaqueductal gray*, Nat Neurosci.; 17(11): 1607–1612. doi:10.1038/nn.3832, November 2014.

[18] Lauren Y. Atlas, Niall Bolger, Martin A. Lindquist, and Tor D. Wager, *Brain mediators of predictive cue effects on perceived pain*, J Neurosci.; 30(39): 12964–12977. doi:10.1523/JNEUROSCI.0057-10.2010., September 29, 2010.

[19] Liane Schmidt, Daphna Shohamy and Tor D. Wager. (In Prep.) *Challenging the effects of perceived control: effects of certainty and agency on pain.*

[20] L. Koban, M. Jepma, Tor D. Wager, (In prep.). *Distinct brain mediators of social influence and conditioned cue effects on pain.*

[21] Cyril R. Pernet, *Misconceptions in the use of the General Linear Model applied to functional MRI: a tutorial for junior neuro-imagers*, Frontiers in Neuroscience, January 2014, Volume 8, Article 1.

[22] Martin A. Lindquist, Anjali Krishnan, Marina López-Solà, Marieke Jepma, Choong-WanWoob, Leonie Koban, Mathieu Roy, Lauren Y. Atlas, Liane Schmidt, Luke J. Chang, Elizabeth A. Reynolds Losin, Hedwig Eisenbarth, Yoni K. Ashar, Elizabeth Delk, Tor D.Wager, *Group-regularized individual prediction: theory and application to pain*, NeuroImage, `http://dx.doi.org/10.1016/j.neuroimage.2015.10.074`, 2015.

[23] Links - Gatton College of Business & Economics - University of Kentucky, *Handout - Normalizing variable*, BA 762, Research methods, `http://www.analytictech.com/ba762/handouts/normalization.htm`, consulted on March 15, 2018.

[24] Sandro Saitta, *Standardization vs. Normalization*, DataMiningBlog.com, Published on July 10, 2007 in data preprocessing, normalization, scaling, standardization, `http://www.dataminingblog.com/standardization-vs-normalization/`, consulted on March 19, 2018.

[25] Andrew T. Drysdale, Logan Grosenick, Jonathan Downar, Katharine Dunlop, Farrokh Mansouri, Yue Meng, Robert N. Fetcho, Benjamin Zebley, Desmond J. Oathes, Amit Etkin, Alan F. Schatzberg, Keith Sudheimer, Jennifer Keller, Helen S. Mayberg, Faith M. Gunning, George S. Alexopoulos, Michael D. Fox, Alvaro Pascual-Leone, Henning U. Voss, B.J. Casey, Marc J. Dubin and Conor Liston, *Resting-state connectivity biomarkers define neurophysiological subtypes of depression*, Nature Medicine, doi:10.1038/nm.4246, December 5, 2016.

[26] Ismael Conejero, Emilie Olié, Raffaella Calati, Déborah Ducasse, Philippe Courtet, *Psychological Pain, Depression, and Suicide: Recent Evidences and Future Directions*, Mood Disorders (E Baca-Garcia, Section Editor), Current Psychiatry Reports (2018) 20:33, `https://doi.org/10.1007/s11920-018-0893-z`, April 05, 2018.

[27] Megan H. Lee, Christopher D. Smyser, and Joshua S. Shimony, *Resting state fMRI: A review of methods and clinical applications*, Washington University School of Medicine; St. Louis, MO, AJNR Am J Neuroradiol. 2013 Oct; 34(10): 1866–1872, doi: 10.3174/ajnr.A3263, August 30, 2012.

[28] Marti J. Anderson, *Permutation tests for univariate or multivariate analysis of variance and regression*, Published on the NRC Research Press Web site on March 7, 2001.

[29] Stefan Haufe, Frank Meinecke, Kai Görgend, Sven Dähne, John-Dylan Haynes, Benjamin Blankertz, Felix Bießmann, *On the interpretation of weight vectors of linear models in multivariate neuroimaging*, NeuroImage 87 (2014) 96–110, Published by Elsevier Inc., November 15, 2013.

[30] John-Dylan Haynes, *A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives*, Neuron Primer, Neuron 87, Elsevier Inc, July 15, 2015.

[31] Jessica Schrouff, *Interpreting predictive models in terms of anatomically labelled regions*, Slides: 'Pattern Recognition for NeuroImaging (PR4NI)', Standford University, CA, USA, `http://ohbm.i4adev.com/files/PatternRecognition4NeuroImaging_Schrouff_Jessica.pdf`, consulted on May 3, 2018.

[32] Andrew F. Hayes, Li Cai, *Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation*, Behavior Research Methods, 39 (4), 709-722, 2007.

[33] Pierre Geurts, *Introduction to machine learning: Unsupervised learning*, ULiège, Year 2017-2018.

[34] Louis Wehenkel, *Introduction to machine learning: Applied inductive learning*, ULiège, Year 2017-2018.

[35] Dibya Jyoti Bora, Dr. Anil Kumar Gupta, *Effect of Different Distance Measures on the Performance of K-Means Algorithm: An Experimental Study in Matlab*, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2501-2506, 2014.

[36] Igor Melnykov, Volodymyr Melnykov, *On K-means algorithm with the use of Mahalanobis distances*, Statistics and Probability Letters 84 (2014) 88–95, Elsevier B.V, September 27, 2013.

[37] Statistics How To, *Mahalanobis Distance: Simple Definition, Examples*, Copyright © 2018 Statistics How To, `http://www.statisticshowto.com/mahalanobis-distance/`, consulted on May 8, 2018.

[38] Gabriel Martos, Alberto Muñoz, and Javier González, *On the Generalization of the Mahalanobis Distance*, J. Ruiz-Shulcloper and G. Sanniti di Baja (Eds.): CIARP 2013, Part I, LNCS 8258, pp. 125–132, 2013. c Springer-Verlag Berlin Heidelberg, 2013

[39] R. De Maesschalck, D. Jouan-Rimbaud, D.L. Massart, *Tutorial: The Mahalanobis distance*, Chemometrics and Intelligent Laboratory Systems 50 2000. 1–18, Elsevier Science B.V, 2000.

[40] Benson Mwangi, Tian Siva Tian, and Jair C. Soares, *A review of feature reduction techniques in neuroimaging*, NIH Public Access, Published in final edited form as: Neuroinformatics. 12(2): 229–244. doi:10.1007/s12021-013-9204-3, April 2014.

[41] Jason Brownlee, *Linear Regression for Machine Learning*, Machine Learning Mastery website, `https://machinelearningmastery.com/linear-regression-for-machine-learning/`, March 25, 2016, consulted on May 9, 2018.

[42] Saahil Sharma, *Machine Learning 102: Linear Regression, Ordinary Least Squares (OLS), Correlation and Analysis of Variance(ANOVA)*, MEDIUM Website, July 5, 2017, `https://medium.com/@saahil1292/machine-learning-102-linear-regression-ordinary-least-squares-ols-correlation-and-analysi` consulted on May 9, 2018.

[43] R.X. Liu, J. Kuang, Q. Gong, X.L. Hou, *Principal component regression analysis with SPSS*, Computer Methods and Programs in Biomedicine 71 (2003) 141/147, 2002 Elsevier Science Ireland Ltd, April 11, 2002.

[44] NCSS Statistical Software, *Chapter 340: Principal Components Regression*, NCSS, LLC., `https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Principal_Components_Regression.pdf`, consulted on May 10, 2018.

[45] PennState - Eberly College of Science, STAT 897D - Applied Data Mining and Statistical Learning,*Lesson 7: Dimension Reduction Methods*, `https://newonlinecourses.science.psu.edu/stat857/node/208/`, consulted on May 10, 2018.

[46] Payam Refaeiladeh, Lei Tang, Huan Liu, *Cross-validation*, Comp. by: BVijayalakshmiGalleys0000875816 53, November 6, 2008.

[47] Christophe Leys, Olivier Klein, Yves Dominicy, Christophe Ley, *Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance*, Journal of Experimental Social Psychology 74 (2018) 150–156, 2017 Elsevier Inc., September 26, 2017.

[48] Robert G. Garrett, *The chi-square plot: a tool for multivariate outlier recognition*, Journal o/Geochemical Exploration, 32 (1989) 319-341, Elsevier Science Publishers B.V., Amsterdam, October 20, 1988.

[49] Unknown author, *Finding Multivariate Outlier*, Applied Multivariate Statistics – Spring 2012, ETH - Swiss Federal Institute Technology Zurich, `https://stat.ethz.ch/education/semesters/ss2012/ams/slides/v2.2.pdf`, consulted on May 12, 2018.

[50] SPM Website, *SPM*, By members & collaborators of the Wellcome Trust Centre for Neuroimaging, London UK, `http://www.fil.ion.ucl.ac.uk/spm/`, consulted on May 12, 2018.

[51] Frederik Maes, *H06W8a: Medical image analysis - Class 3: Image registration*, KULeuven, Year 2017-2018.

[52] Ronald Sladky, Karl J. Friston, Jasmin Tröstl, Ross Cunnington, Ewald Moser, Christian Windischberger, *Slice-timing effects and their correction in functional MRI*, NeuroImage 58 (2011) 588–594, 2011 Elsevier Inc., July 2, 2011.

[53] L. Freire, A. Roche, and J.-F. Mangin, *What is the Best Similarity Measure for Motion Correction in fMRI Time Series?*, IEEE Transaction on Medical Imaging, Vol. 21, NO. 5, May 2002.

[54] JW Hardin, JM Hilbe, *Generalized linear models and extensions - Second edition*, A Stata Press Publication, StataCorp LP, College Station, Texas, First edition 2001, Second edition 2007, ISBN-10: 5-9718-014-9.

[55] Laurens Van der Maaten, Geoffrey Hinton, *Visualizing Data using t-SNE*, Journal of Machine Learning Research 9 (2008) 2579-2605, Published on August 11, 2008.

[56] John R. Hershey and Peder A. Olsen, *Approximating the Kullback-Leibler Divergence between Gaussian Mixture Models*, 1-4244-0728-1/07/$20.00 2007 IEEE, 2007.

[57] Laurens Van der Maaten, Geoffrey Hinton, *User's Guide for t-SNE Software*, consulted on May 14, 2018.

[58] Unknown author, *Gene Expression Data Analysis Suite (GEDAS)*, site created and maintained by T.V. Prasad, Created March 2003, last updated Jan 2006, `http://gedas.bizhat.com/dist.htm`, consulted on May 19, 2018.

[59] Unknown author, *Distance Metrics*, Math.NET Numerics, Maintained by Christoph Rüegg, `https://numerics.mathdotnet.com/distance.html`, consulted on May 19, 2018.

[60] Christian S. Perone, *Machine Learning :: Cosine Similarity for Vector Space Models (Part III)*, Terra Incognita, `http://blog.christianperone.com/2013/09/machine-learning-cosine-similarity-for-vector-space-models-part-iii/`, Posted on September 12, 2013, consulted on May 19, 2018.

[61] Peter J. Rousseeuw, *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*, Journal of Computational and Applied Mathematics 20 (1987) 53-65, Elsevier Science Publishers B.V, November 27, 1986.

[62] Kapild Alwani, *Using Silhouette analysis for selecting the number of cluster for K-means clustering. (PART 2)*, Data Science Musing of Kapild, November 10, 2015, `https://kapilddatascience.wordpress.com/2015/11/10/using-silhouette-analysis-for-selecting-the-number-of-cluster-for-k-means-clustering/`, consulted on May 19, 2018.

[63] Herve Abdi and Lynne J. Williams, *Overview - Principal component analysis*, Volume 2, DOI: 10.1002/wics.101, July/August 2010.

[64] George H. Dunteman, *Principal Components Analysis*, Series: Quantitative application in the social science, SAGE University paper, Book number: 0-8039-3104-2, 1989.

[65] Stephen C. Johnson, *Hierarchical clustering schemes*, Psychometrika - vol. 32 no. 3, September 1967.

[66] F. Murtagh, *A Survey of Recent Advances in Hierarchical Clustering Algorithms*, The Computer Journal, vol. 26, no. 4, 1983.

[67] Tae Kyun Kim, *T test as a parametric statistic*, Korean Journal of Anesthesiology, Statistical Round, pISSN 2005-6419 eISSN 2005-7563, September 7, 2015.

[68] Statwing, *Statistical Significance (T-Test)*, Statwing Documentation - Statwing's approach to statistical testing, `http://docs.statwing.com/examples-and-definitions/t-test/statistical-significance/`, consulted on May 20, 2018.

[69] Stephanie from StatisticsHowTo Website, *T Test (Student's T-Test): Definition and Examples*, last modified on January 6th, 2018, `http://www.statisticshowto.com/probability-and-statistics/t-test/`, consulted on May 20, 2018.

[70] B. T. Thomas Yeo, Fenna M. Krienen, Jorge Sepulcre, Mert R. Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L. Roffman, Jordan W. Smoller, Lilla Zöllei, Jonathan R. Polimeni, Bruce Fischl, Hesheng Liu, and Randy L. Buckner, *The organization of the human cerebral cortex estimated by intrinsic functional connectivity*, J Neurophysiol 106: 1125–1165, doi:10.1152/jn.00338.2011, June 1, 2011.

[71] Unknown author, *SPSS Tutorials: Pearson Correlation*, Kent State University, `https://libguides.library.kent.edu/SPSS/PearsonCorr`, Last Updated: May 25, 2018, consulted on May 26, 2018.

[72] Virendra from Learning and Experience Website, *What is a Confusion Matrix in Machine Learning?*, March 30, 2018, `http://learning.maxtech4u.com/what-is-a-confusion-matrix-in-machine-learning/`, consulted on June 4, 2018.

[73] Thomas E. Nichols and Andrew P. Holmes, *Nonparametric Permutation Tests For Functional Neuroimaging: A Primer with Examples*, Human Brain Mapping 15:1–25(2001), 2001.

[74] Thomas J. Leeper, *Permutation tests*, R course tutorials, November 14, 2013, `https://thomasleeper.com/Rcourse/Tutorials/permutationtests.html`, consulted on June 4, 2018.

[75] Tian Ge, BT Thomas Yeo, Anderson M Winkler, *A brief overview of permutation testing with examples*, Organization for Human Brain Mapping, March 19, 2018, `https://www.ohbmbrainmappingblog.com/blog/a-brief-overview-of-permutation-testing-with-examples`, consulted on June 4, 2018.

[76] Statistics Solutions website, *One Sample T-Test*, `http://www.statisticssolutions.com/manova-analysis-one-sample-t-test/`, 2018, consulted on June 4, 2018.

[77] Christopher R. Genovese, Nicole A. Lazar and Thomas Nichols, *Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate*, NeuroImage 15, 870–878 (2002), February 23, 2001.

[78] Joset A. Etzel, Jeffrey M. Zacks, and Todd S. Braver, *Searchlight analysis: promise, pitfalls, and potential*, Neuroimage.; 78: 261–269. doi:10.1016/j.neuroimage.2013.03.041, September 2013.

[79] Unknown author, *One-way ANOVA in SPSS Statistics*, Laerd Statistics, `https://statistics.laerd.com/spss-tutorials/one-way-anova-using-spss-statistics.php`, 2018, consulted on June 6, 2018.

[80] G. F. Kuder and M. W. Richardson, *The theory of the estimation of test reliability*, Psychometrika, vol.2, no.3, September 1937.

[81] Image from Figure 2.2, `https://www.google.com/search?q=bold+signal&source=lnms&tbm=isch&sa=X&ved=2ahUKEwjZ6aDZ97faAhXI54MKHTS1AXMQ_AUoAXoECAAQAw&biw=1280&bih=566#imgrc=tfQDaDZNhkUHlM:`, consulted on April 13, 2018.

[82] Image from Figure 2.4, `https://www.google.com/search?q=bold+signal&source=lnms&tbm=isch&sa=X&ved=2ahUKEwi7oKX5n7jaAhVs5YMKHXbzDT0Q_AUoAXoECAAQAw&biw=1280&bih=615#imgrc=jvyI88e6oHJgxM:`, consulted on April 13, 2018.

[83] Image from Figure 2.5, `https://www.google.com/search?biw=1280&bih=615&tbm=isch&sa=1&ei=QpXTWqujMoiGjwTb9pWwCQ&q=predicted+fmri+signal&oq=predicted+fmri+signal&gs_l=psy-ab.3...63443.66641.0.66859.21.14.0.0.0.0.414.1776.0j1j5j0j1.7.0....0...1c.1.64.psy-ab..14.4.1068...0j0i67k1.0.sDLjb_QY4-s#imgrc=RJOfBMIUtNesVM:`, consulted on April 15, 2018.

[84] Image from Figure 2.7, https://www.google.com/search?q=drift+fmri+signal&source=lnms&tbm=isch&sa=X&ved=2ahUKEwjZw5_D77zaAhUJ74MKHXr-ANMQ_AUoAXoECAAQAw&biw=1280&bih=615#imgrc=xgLfbaW5uAOF2M:, consulted on April 15, 2018.

[85] Image from Figure 2.1, https://www.google.com/search?biw=1280&bih=615&tbm=isch&sa=1&ei=iXnWWrm-KoXcjwTA_pTwBg&q=physiology+of+pain+perception&oq=physiology+of+pain+perception&gs_l=psy-ab.3..0.592470.597428.0.598009.29.18.0.9.9.0.272.2020.0j10j3.13.0....0...1c.1.64.psy-ab..8.21.1853...0i67k1j0i5i30k1j0i24k1j0i30k1.0.GA3STOGxqak#imgrc=skj5d2ZthX-pLM:, consulted on April 17, 2018.

Appendix A

t-Distributed Stochastic Neighbor Embedding (t-SNE)

This technique allows to visualize high-dimensional data by giving each data point a location in a 2D or 3D space. They start by computing a matrix of pairwise similarities, and then use the actual t-SNE method for visualizing the resulting similarity data. It is able to show both the local and global structures, for example the presence of clusters at several scales [55].

Stochastic Neighbor Embedding

In stochastic neighbor embedding, we start by converting the high-dimensional Euclidean distances between data points into conditional probabilities with the following formula:

$$p_{j|i} = \frac{\frac{\exp(-||x_i - x_j||^2}{2\sigma_i^2}}{\sum_{i \neq k} \frac{\exp(-||x_i - x_j||^2}{2\sigma_i^2}} \tag{A.1}$$

Where $p_{j|i}$ is the conditional probability reflecting the similarity between data point x_i and data point x_j. The neighbors of x_i (such as x_j) are chosen based on their probability density under a Gaussian centered at x_i, with a variance σ_i. This conditional probability would be very high for nearby data points, and almost infinitesimal for widely separated ones.
We can compute the same metric for the low-dimensional representations y_i and y_j of data points x_i and x_j:

$$q_{j|i} = \frac{\exp(-||y_i - y_j||^2}{\sum_{i \neq k} \exp(-||y_i - y_j||^2} \tag{A.2}$$

In this case, the Gaussian variance is chosen to equal $\frac{1}{\sqrt{2}}$.
Since we are only interesting in modeling the pairwise similarities, we set these two values $p_{j|i}$ and $q_{j|i}$ to zero.
If y_i and y_j are correct low-dimensional representations of x_i and x_j, the conditional probabilities $p_{j|i}$ and $q_{j|i}$ should be equal. Thus, the idea is to find the low-dimensional data representation that minimizes the mismatch between $p_{j|i}$ and $q_{j|i}$. The metric chosen to perform this task is the Kullback-Leibler (KL) divergence, which is a tool widely used in statistics to measure the similarity between two density distributions [56]. The SNE method minimizes the sum of KL divergences over all data points using a gradient descent technique. The cost function C is written as:

$$C = \sum_i KL(P_i||Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}} \tag{A.3}$$

Where P_i represents the conditional probability distribution over all other data points given data point x_i and Q_i, the conditional probability distribution over all map points considering map point y_i. This cost function focuses on retaining the local structure in the initial dataset.
The Gaussian which is centered over each high-dimensional data point x_i cannot have the same variance σ_i for all points, since they are likely to vary in density in the dataset. For example, in more dense region, a smaller value for the variance is more appropriate than in sparser regions. A particular value for the variance will induce a probability distribution P_i over all the other data points. The entropy of this distribution increases with the variance. A binary search is initiated to find the value of σ_i which will produce a P_i with a fixed perplexity, defined by the user, which could be thought of as a smooth measure of the effective number of neighbors. The formula is the following:

$$Perp(P_i) = 2^{H(P_i)} \tag{A.4}$$

Where $H(P_i)$ is the Shannon entropy of P_i:

$$H(P_i) = \sum_j p_{j|i} \log_2 p_{j|i} \tag{A.5}$$

The gradient descent technique used for the minimization of the cost function shown in Eq. (A.3) is given by:

$$\frac{\delta C}{\delta y_i} = 2 \sum_j (p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j) \tag{A.6}$$

This gradient descent can be interpreted with a physical point of view: if we imagine a set of springs between map point y_i and all other map points y_j, the gradient corresponds to the resultant force between them. All springs exert a force in the direction $(y_i - y_j)$, and the spring between y_i and y_j either repel or attract the map point, if the distance between the two is too small or too large to represent the similarities between the two high-dimensional points x_i and x_j, respectively.
The initialization of the gradient is performed by sampling map points randomly from an isotropic Gaussian with small variance, centered around the origin [55].

t-distributed Stochastic Neighbor Embedding

The SNE method presented above shows different limitations: for example, the cost function is difficult to optimize, and this method is subject to a problem called "crowding problem", which will be explained shortly. The t-SNE is able to overcome these limitations, first by employing a different cost function, which is symmetrical and with simpler gradients, and used a Student-t distribution rather than a Gaussian to compute the similarity between two points in the low-dimensional space.

Symmetric SNE

Instead of computing the sum of the Kullback-Leibler divergences between the conditional probabilities $p_{j|i}$ and $q_{j|i}$, we can compute a single KL divergence between a joint probability distribution P in the high-dimensional space and a joint probability distribution Q in the low-dimensional space, which will be in turn minimized:

$$C = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{A.7}$$

Where, just as before, p_{ij} and q_{ij} are set to zero. It is said to be symmetrical because we have the following properties: $p_{ij} = p_{ji}$ and $q_{ij} = q_{ji}$, for every value of i and j. In this situation, pairwise

similarities in the low-dimensional space can be computed using the following formula:

$$q_{ij} = \frac{\exp(-||y_i - y_j||^2)}{\sum_{k \neq l}(-||y_k - y_l||^2)} \tag{A.8}$$

However, in order not to be affected by outliers, we define the joint probability p_{ij} in the high-dimensional space to be the symmetrized conditional probability:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n} \tag{A.9}$$

Where n is the number of data points. By doing that, we ensure that each point has a significant contribution to the cost function, which is not the case when a high-dimensional data point x_i is an outlier.
The main advantage of using a symmetric SNE is the gradient function, which is easier to compute [55]:

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j) \tag{A.10}$$

The crowding problem

The crowding problem occurs when the data points are distributed in a region on a high-dimensional manifold around a point i and we try to model the pairwise distance from i to the data points in a two-dimensional map. Imagine that we have 11 mutually equidistant data points in a ten-dimensional manifold. There is no way to represent these points correctly in a two-dimensional space. Indeed, small distances points can be accurately modeled in the 2D map, but the area available to accommodate moderately distant data points will not be large enough compared with the area dedicated to nearby points. Thus, most of the moderately distant points will be too far away in the 2D map. From the physical point of view explained earlier, it would mean that the attractive force between i and those points would be very small. The very large number of such forces collapses together the points in the center of the map and prevents gaps from forming between the natural clusters [55].

Mismatched tails can compensate for mismatched dimensionalities

The crowding problem mentioned above can be solved with the following method: in the high-dimensional space, the distances are converted into probabilities using a Gaussian distribution, while in the low-dimensional space, we choose a Student-t distribution with one degree of freedom. This type of distribution has much heavier tails than the Gaussian one when it converts distances into probabilities. Doing that, a moderate distance data point is more faithfully modeled in the low-dimensional space. The unwanted attractive forces are thus removed. The joint probabilities of the low-dimensional space q_{ij} are now defined as:

$$q_{ij} = \frac{(1 + ||y_i - y_j||^2)^{-1}}{\sum_{k \neq l}(1 + ||y_k - y_l||^2)^{-1}} \tag{A.11}$$

Then, the gradient of the KL divergence between the Gaussian based joined probability distribution P and the Student-t based joined probability distribution Q is given by [55]:

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 + ||y_i - y_j||^2)^{-1} \tag{A.12}$$

This section showed the theory behind the tool used for the visualization of the high-dimensional

data we work with in this project. The Matlab `tsne.m` function starts by reducing the number of components using a Principal Component Analysis (PCA) dimension reduction technique. The number of dimensions at the end of the PCA pre-process is determined by the user. In this projet, it has been set to 100. After that, the t-SNE method is applied and the user can choose to visualize the results in a two-dimensional or three-dimensional space. The perplexity is also a parameter that needs to be defined. Several values of the perplexity, advised to be in a range of [5 50], were tested in order to find the one that leads to the best results, which was 30 [57].
Note that the clustering was not applied directly on the low-dimensional maps obtained with this method, it was first applied on the weight maps (or after a dimension reduction technique, such as PCA), and then the results were visualized using t-SNE. In some cases, the technique might make clusters appear inside the data, but it was not the case in ours.

Appendix B

Group-Regularized Individual Predictions (GRIP)

A recurrent problem linked with fMRI studies, not only revising pain, is that the amount of data is usually too small and noisy to be able to generate correct results. With these limitations, one can easily understand that with the limiting number of individual images for each task, along with the noisy nature of fMRI measures, it is difficult to make predictions at the individual's level. However, the model might be improved by using the GRIP (Group-Regularized Individual Prediction) method described here.

The idea behind the development of this method is that brain representations are idiographic, i.e. each individual has a unique brain with specific characteristics and patterns. However, usually, there is information which is shared across individuals, at a larger scale. Knowing that, Lindquist and al. [22] have developed a method where both notions are taken into account: in their paper, they constructed individual weight maps that result from a weighted contribution of both the idiographic and the group-level maps.

Idiographic maps

The first step consists in finding individual weight maps. For each participant in each study, we apply machine learning techniques in order to predict the outcomes, i.e. the pain ratings, based on the single-trials data linked to the participant. Theoretically, if the single-trials data were of infinite number, the predictions will tend to perfect results, and the GRIP method would not need to be applied. However, the amount of single-trials estimates is quite small, usually we have between 50 and 90 images per patient. That is why we need to improve the predictions by the means of the GRIP method.

We have, for each participant, m observations corresponding to m trials of a certain stimulus applied. We denote these observations $(\boldsymbol{x}_j, y_j)$ for $j = 1, ..., m$ the trial number. The vector $\boldsymbol{x}_j$ is a vector of features (in this case: the parameter estimates for each pixel, which composes a summary of the brain response) and the y_j are the scalar outcome variables, i.e. the pain ratings.
With the help of standard machine learning methods, we can develop the idiographic brain weights $\hat{\boldsymbol{w}}_{\boldsymbol{i}}$ and the outcomes predictions that come along with it, with the following formula:

$$\hat{Y}_i = \hat{\boldsymbol{w}}_{\boldsymbol{i}}{}^T \boldsymbol{x}^* \tag{B.1}$$

Where $\hat{Y}_i$ are the pain ratings predicted by the individual maps and $\boldsymbol{x}^*$ are the brain features.
Both the predictions and the weight maps could be extracted from the `predict.m` function developed

by the CAN Lab team, in the same process than the multivariate maps.

Group-level maps

To design the group-level maps, we used the NPS signature pattern. Indeed, this signature was developed by studying common brain activation patterns among a quite large number of subjects (about 30), so that makes it perfect for the type of maps we need for the GRIP method.
Thus, in this case, the weight maps corresponding to the population-level study, noted $\hat{\boldsymbol{w_p}}$ are already determined and present in the lab tools. The predictions that come along with these maps can be found as follows:

$$\hat{Y_p} = \hat{\boldsymbol{w_p}}^T \boldsymbol{x^*} \tag{B.2}$$

Where $\hat{Y_p}$ are the pain ratings predicted from the population-level maps (from the NPS pattern) and $\boldsymbol{x^*}$ are the brain features.
This is done by applying the NPS signature to the data with the means of the `apply_nps.m` function developed by the Can Lab team. The predictions are found by the means of simple dot products between the weight maps and the observed features, i.e. what compose the `.dat` matrix, the parameter estimates. The population-level weight maps constitute what is important for the continuity of the method implementation. Obviously, they are identical for each subject in this study, since we use a previously defined pattern, i.e. the NPS signature.

Shrinkage factor

The shrinkage factor λ, also called the GRIP estimator, is a number comprised in the range $[0, 1]$ defined as follows:

$$\hat{\boldsymbol{w_G}} = \lambda \hat{\boldsymbol{w_i}} + (1 - \lambda)\hat{\boldsymbol{w_p}} \tag{B.3}$$

Where λ is the so-called shrinkage factor, specifically defined for each subject of the study.
In the paper presented [22], two manners of computing the shrinkage factor have been tested: one is an Empirical Bayes approach, which is based on the ratio of the within-subjects and between-subjects variances, and the other one is a cross-validated approach, based on the prediction accuracy obtained using the idiographic maps and population-level maps, respectively. The paper results showed that the first method gave every time better results.

In the Empirical Bayes (EB) approach, the within-subject and between-subjects variances are needed to compute the shrinkage factor. To find them, some hypotheses have to be made. We assume that for subject i on the j^{th} trials, where $j = 1, 2, ..., M$, the true pain report Y follows a normal distribution with subject-specific mean μ_i and variance $\sigma^2_{w,i}$, which can be written as follows:

$$Y_{ij} \sim \mathcal{N}(\mu_i, \sigma^2_{w,i}) \tag{B.4}$$

Another assumption is, in turn, made on the subject mean μ_i: it can also be modeled as a normal distribution with population mean μ and between-subjects variance σ^2_B. This statement is written as follows:

$$\mu_i \sim \mathcal{N}(\mu, \sigma^2_B) \tag{B.5}$$

If μ_i can be estimated by $\hat{Y_i}$ in Eq. (B.1) and μ can be estimated by $\hat{Y_p}$ in Eq. (B.2), the within-subject variance can be computed as the variance of the observed pain reports (true outcomes) and those predicted with the idiographic maps, and the total variance can be computed as the variance of the observed pain reports to those predicted by the population-level maps. These are denoted

$\hat{\sigma}^2_{w,i}$ and $\hat{s}^2 = \hat{\sigma}^2_{w,i} + \hat{\sigma}^2_B$, respectively. Thus, the between-subject variance, $\hat{\sigma}^2_B$, can be estimated with the two computed values with the following formula:

$$\hat{\sigma}^2_B = max(0, \hat{s}^2 - \hat{\sigma}^2_{w,i}) \tag{B.6}$$

This step is perform in order to ensure non-negative estimates.
The shrinkage factor, for each subject $i = 1, 2, ..., m$, can be computed with the following formula:

$$\lambda_i = \frac{\hat{\sigma}^2_B}{\hat{\sigma}^2_{w,i} + \hat{\sigma}^2_B} \tag{B.7}$$

And this is the factor used in Eq. (B.3). The behavior of the GRIP method is easily interpretable by studying Eq. (B.7): if the between-subject variance is large relative to the within-subject variance, then the shrinkage factor would be close to 1 and the idiographic map has a higher contribution. On the other hand, if the within-subject variance is larger, the shrinkage factor would be close to 0 and in that case the population-level map dominates [22].
A framework of the GRIP method appears in Figure B.1.

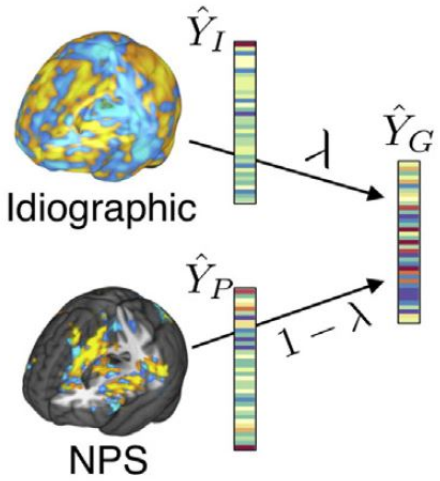


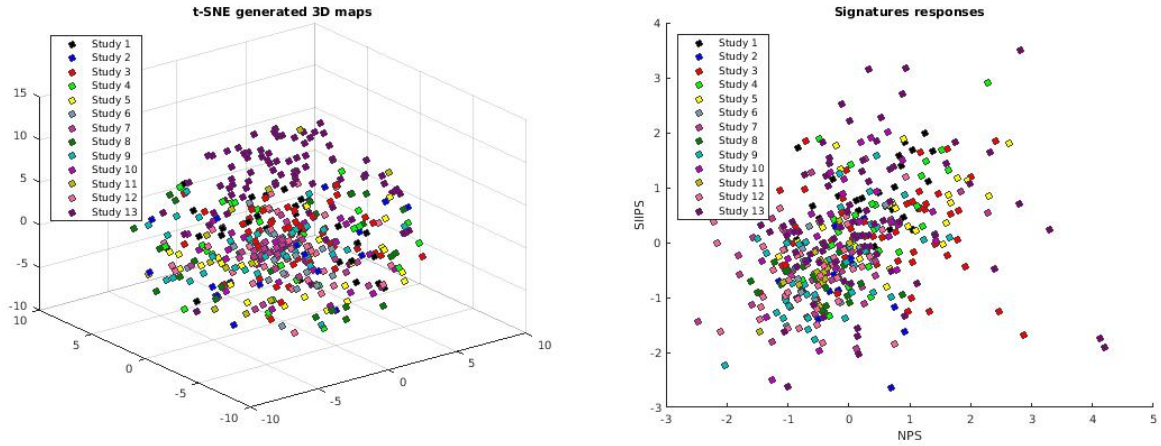
Figure B.1: GRIP method framework [22]

This whole process would give rise to individual pain-predictive weight maps, on which the clustering can be performed. The final matrix would have a size of 433×352328, where each subject is represented as a row, and each column represents a weighted contribution of each type of maps (idiographic and group-level), which are themselves a linear combination of the initial parameter estimates for each voxel. The clustering would be applied directly on those weight maps.
The GRIP method is supposed to make the analysis easier by stabilizing the weights inside the weight maps, which are usually not steady due to the noisy nature of fMRI measurements. However, since, in this work, we are looking at individual differences, it might not be helpful to apply this method, because we impose the weight maps to be a little bit more similar to the NPS signature and we could miss some other activation patterns..

Appendix C

Additional results

Cluster by study



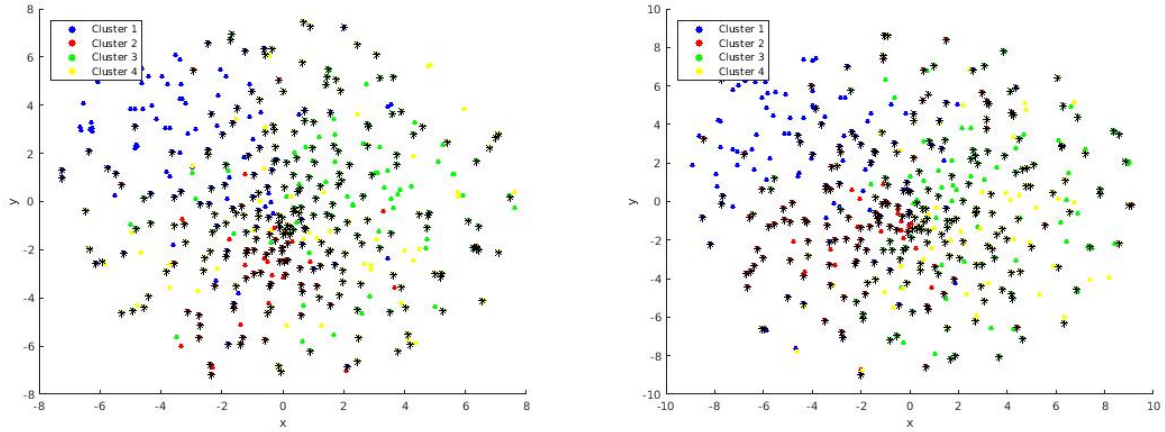
(a) Multivariate maps (visualized in 3D using t-SNE) (b) Between-cluster correlation

Figure C.1: Univariate maps (signatures responses)

Wrong assignments

Multivariate maps

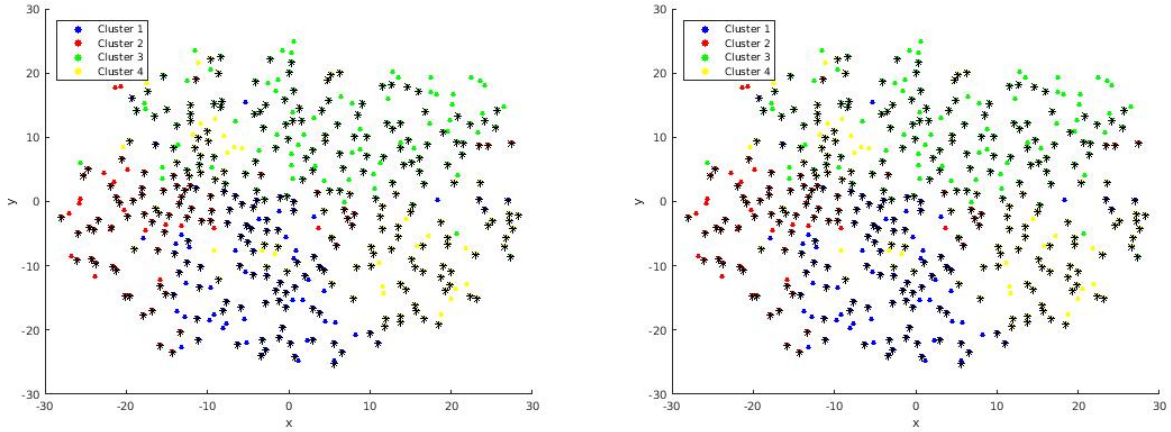
Entire maps with cosine distance



(a) First dataset (b) Second dataset

Figure C.2: Wrong assignments in the entire multivariate maps (cosine distance)

Canonical networks with cosine distance

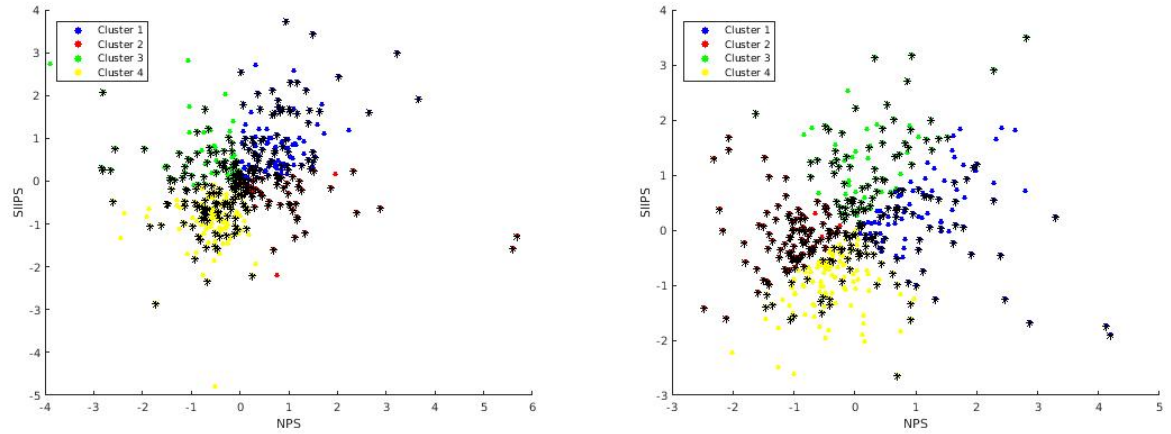


(a) First dataset (b) Second dataset

Figure C.3: Wrong assignments in the canonical networks on multivariate maps (cosine distance)

Univariate maps

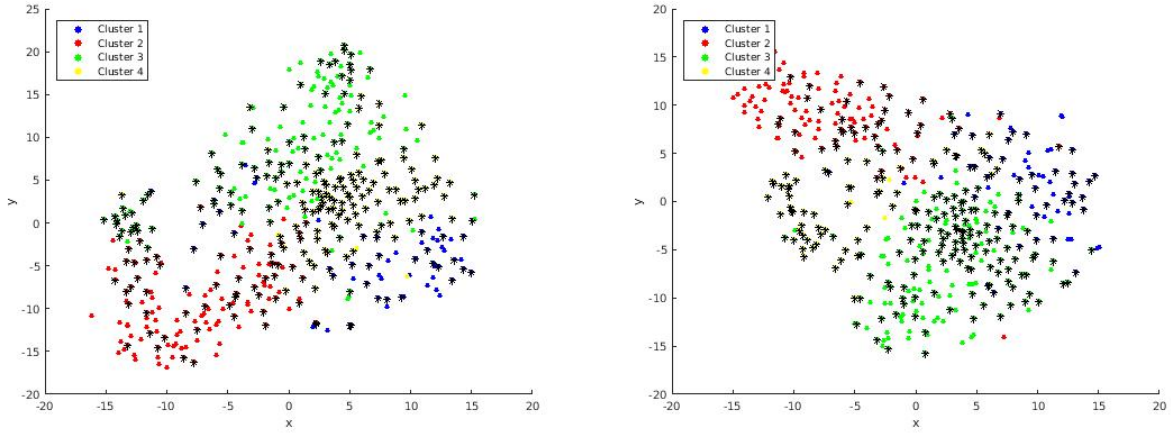
Signatures responses



(a) First dataset (b) Second dataset

Figure C.4: Wrong assignments in the signatures responses on entire univariate maps (cosine distance)

PCA reduced maps



(a) First dataset (b) Second dataset

Figure C.5: Wrong assignments in the PCA reduced univariate maps (cosine distance)

Silhouette plots

Multivariate maps

Signatures responses (Euclidean distance)

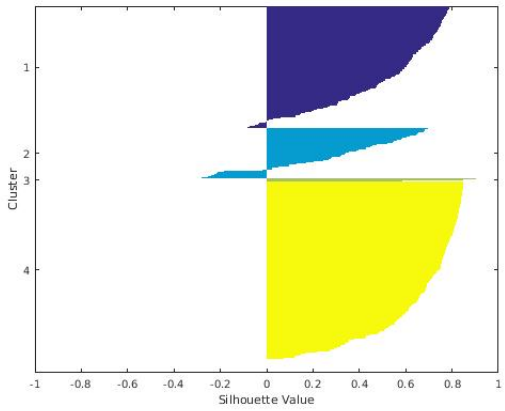
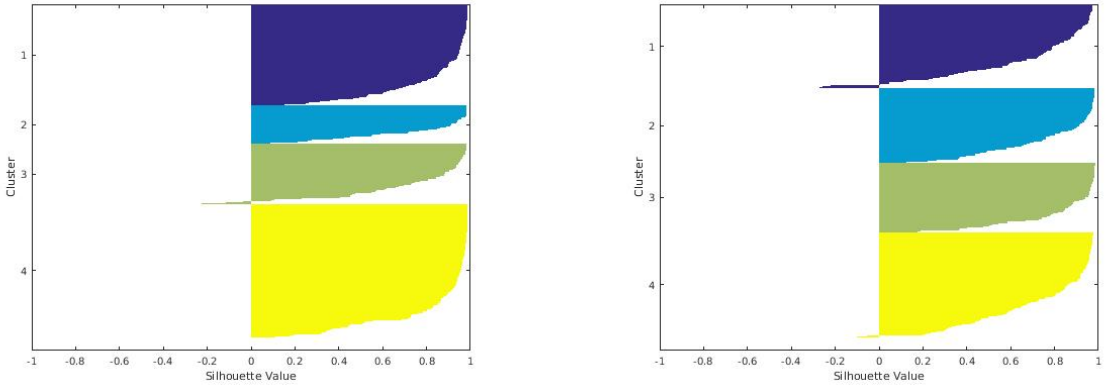


Figure C.6: Silhouette plot of the signatures responses on multivariate maps (Euclidean distance)

Univariate maps

Signatures responses (cosine distance)



(a) First dataset

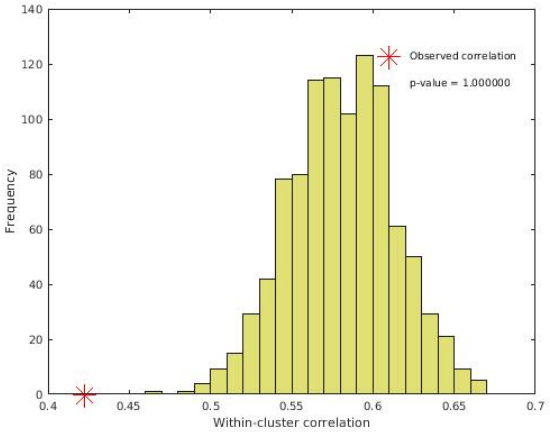
(b) Second dataset

Figure C.7: Silhouette plots of the signatures responses on univariate maps (cosine distance)

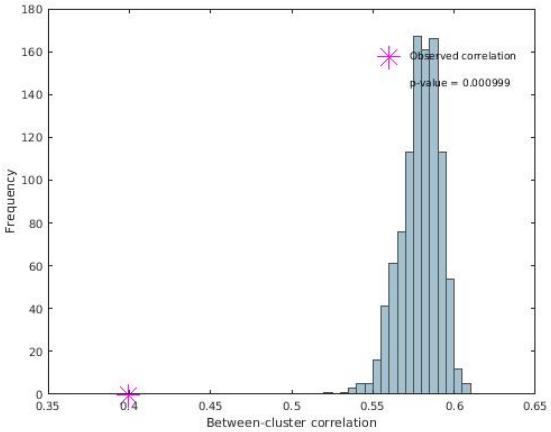
Cluster correlation

Univariate

Signatures reponses



(a) Within-cluster correlation



(b) Between-cluster correlation

Figure C.8: Permutation tests for correlation - Signatures responses on univariate maps