Università
degli Studi
di Padova

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Psicologia Generale

CORSO DI DOTTORATO DI RICERCA IN SCIENZE PSICOLOGICHE
CICLO XXIX

DETECTING COGNITIVE STATES FROM THE ANALYSIS OF STRUCTURAL AND FUNCTIONAL IMAGES OF THE BRAIN: TWO APPLICATIONS OF MULTI-VOXEL PATTERN ANALYSIS ON MRI AND FMRI DATA

Coordinatore: Ch.ma Prof.ssa Francesca Peressotti
Supervisore: Ch.mo Prof. Giuseppe Sartori
Co-Supervisore: Ch.ma Prof.ssa Sara Mondini

Dottorando : Andrea Zangrossi

Table of contents

GENERAL INTRODUCTION

Chapter 1 – From model-driven to data-driven approach in neuroimaging

Chapter 2 –Multi-Voxel Pattern Analysis: An Overview

SECTION A: Decoding Cognitive Reserve from the brain: a MVPA application

Chapter 3 – Neural correlates of Cognitive Reserve: Education and Grey Matter Volume (I)

Chapter 4 – Neural correlates of Cognitive Reserve: Education and Grey Matter Volume (II)

SECTION B: Detecting autobiographical memories from brain activity: a MVPA application

Chapter 5 – "Have you been there before?" Decoding rooms recognition in the human brain

Chapter 6 – "Do you remember it?" Decoding autobiographical emotional memories from brain activity

GENERAL DISCUSSION

SUMMARY

In recent years, the efficacy and accuracy of Multi-Voxel Pattern Analysis (MVPA) techniques on neuroimaging data has been tested on different topics. These methods have shown the ability to decode mental states from the analysis of brain scans, for this reason it has been called “brain reading”. The predictions can be applied to: general states, referring to stable conditions not related to a contingent task (e.g., a neurological diagnosis), or specific mental states, referring to task-related cognitive processes (e.g., the perception of a category of stimuli). According to several neuroscientists, brain reading approach can potentially be useful for applications in both clinical and forensic neuroscience in the future.

In the present dissertation, two applications of the brain reading approach are presented on two relevant topics for clinical and forensic neuroscience that have not been extensively investigated with these methods. In Section A, the application of this approach is tested in decoding different levels of Cognitive Reserve from the pattern of grey matter volume, in two MRI studies. Whereas, in Section B two fMRI studies investigate the possibility of decoding real autobiographical memories from brain activity (fMRI).

The aim of this thesis is to contribute to the amount of studies showing the usefulness of multivariate techniques in decoding “mental states” starting from the analysis of structural and functional brain imaging data, as well as the potential uses in clinical and forensic settings.

Chapter 1

FROM MODEL-DRIVEN TO DATA-DRIVEN APPROACH IN NEUROIMAGING

INTRODUCTION

One of the main issues in cognitive neuroscience is related to the localization of cognitive functions within the human brain. Moreover, a great amount of studies has focused on brain correlates of neurological and psychiatric diseases. Thus, the "*where?*" question seems to be the most commonly asked one, at least within the neuroimaging field. Conventional neuroimaging approaches have mainly faced the problem of identifying brain areas where activity was related to a specific cognitive function or mental state. A different approach has been introduced with the advent of multivariate analysis techniques, which can represent a step forward in the direction of a more complete understanding of human brain. In the last years, a number of neuroimaging studies have adopted multivariate techniques for the analysis of both functional and structural data. The use of these new methods can lead to important advantages for the advance of knowledge about how the brain works and about the relation between mind and brain. Indeed, as will be discussed in the following chapters, the main aim of multivariate analysis techniques applied to neuroimaging data is to infer a mental state starting from the analysis of brain imaging data. Therefore, the crucial point is that now we have the possibility to study the relation between mind and brain from a new perspective. Moreover, these novel methods allow researchers to answer to different research questions about the nature of this relation.

Multivariate approaches can be adopted for decoding cognitive states by analyzing the functional or structural images related to those states. This is the main reason why these methods can be referred to as *classifiers* or *learning machines.* These last two expressions emphasize the origin of these techniques, that come from a branch of artificial intelligence, called Machine Learning.

In this chapter, the standard approach to the analysis of neuroimaging data will be compared with the new emerging multivariate methods, highlighting their differences and their critical consequences for the cognitive neuroscience field. The discussion will start from the theoretical background from which these techniques come from.

PATTERN RECOGNITION AND MACHINE LEARNING IN COGNITIVE NEUROSCIENCE

What is Pattern recognition?

Everyday people carry out many complex tasks that are surprisingly made simple by their brain. So that, we are not aware of the real complexity of these processes, until the moment when we are asked to formalize them. For example, imagine how easily we recognize the letter “A” (Fig. 1.1a) despite the differences or the alterations of its graphical aspects, or the ease in recognizing each of the stimuli in Fig. 1.1b as belonging to a specific category (e.g., a car, or a face) even if none of them has been previously experienced.

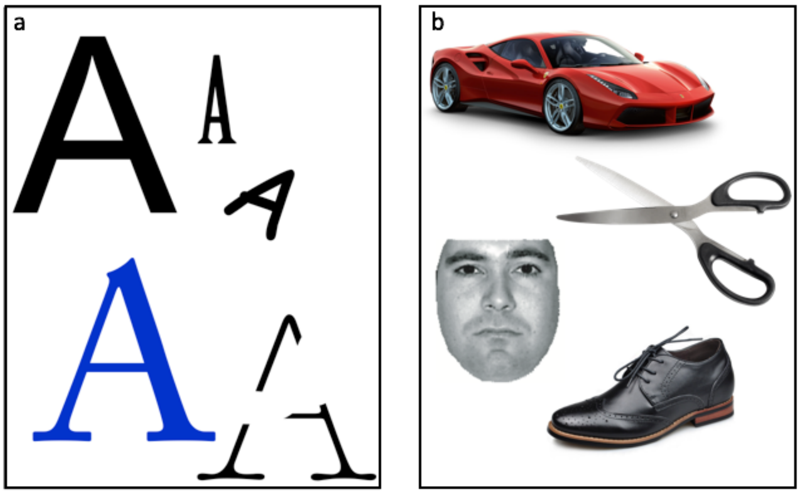

Fig. 1.1 Pattern recognition in everyday life. a) Despite the differences in shape, orientation, color, or the occlusion of some graphical parts, in all cases we recognize the letter “A”; b) An example of how easily we can associate a new stimulus to a known category

How do we associate a new stimulus to one among thousands of known categories? This cognitive process is called pattern recognition. Pattern recognition is a basic function of the human brain, and it can be realized in some milliseconds (DiCarlo, Zoccolan, & Rust, 2012). Given this speed, it is probably a product of evolution (M. W. Brown & Aggleton, 2001) as it allow humans to access the semantic information about a new stimulus and, in case, to evaluate its dangerousness. However, the apparent ease of this process hides the computational complexity of it. One of the most effective definitions of “pattern recognition” is the following:

“By the time are five years old, most children can recognize digits and letters. Small characters, large characters, handwritten, machine primed, or rotated – all are easily

recognized by the young. The characters may be written on a cluttered background, on crumpled paper or may even be partially occluded. We take this ability for granted until we face the task of teaching a machine how to do same. Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make reasonable decisions about the categories of the patterns." (Jain, Duin, & Mao, 2000)

In many real-life settings the need of computer-based systems reproducing this extraordinary human ability is continuously increasing. For example, it is common nowadays to bump into OCR websites or software. OCR stands for "optical character reader" and refers to automatic systems for the translation of handwritten documents in a digital format, and can be very useful for a number of applications. In the industry, increasing the speed and therefore the automation of the workflow is a minimum requirement for surviving in a competitive world. For this reason, several practical applications of machine-based pattern recognition are requested today, also because of the high number of variables that has to be considered simultaneously and cannot be within the range of human workers. In the next paragraph an example of automatized pattern recognition system will be presented, in order to describe the structure of the problem and how modern computer-based systems can deal with it.

Pattern recognition in practice: an example

Duda, Hart and Stork (2001) reported a good example that allows to understand the complexity of some practical problem in which the help of automatized pattern recognition system is required. Imagine that a company specialized in fish packing wants to automate the process through which fishes are differentiated on the basis of their species, before being packaged. The first species classification tested is that between sea bass and salmon. Given that the best classificatory is certainly an individual with an expertise in discriminating different fish species just looking at each example, the decision about the class (i.e., specie) has to be made on the basis of an optical sensing system. What the hypothetical researchers do as a first step is to acquire, by means of a camera, images of exemplars of the two species, in order to note the crucial differences that allow the human classifier to visually identify the species. So that, several physical *features* are highlighted, such as length, lightness, width and so on, together with the variability of these features across the images of different exemplars. Given that no kind of features emerged as uniquely represented in one of the two species (i.e., having a horn is a feature that uniquely characterizes a rhinoceros if compared to an hippopotamus), but are shared between sea basses and salmons (i.e., of course both have a length, a width and a lightness), what discriminate between the fishes is the *pattern* of features. In other words, from a mathematical perspective, the two species are described by different *models* of features. In a case like this, the classification system can be roughly represented as having the structure indicated in Figure 1.1.

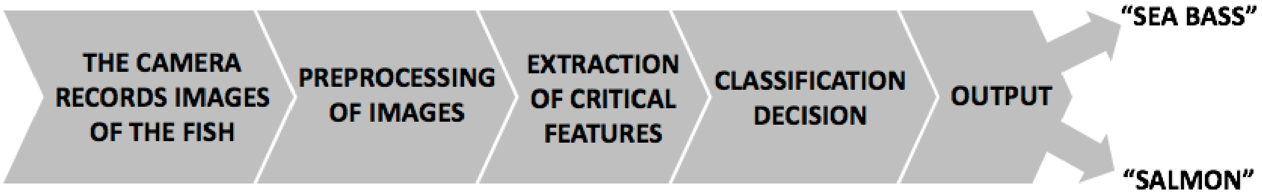

Fig. 1.1 (adapted from Duda et al., 2001). The rough macro-structure of a classification system that could be applied to the described example. First, images of fish examples are recorded. The images (input to the classifier) are then preprocessed in order to make the classification easier (noise elimination), and critical features (those that allow the classification between different species) are extracted from the whole set of physical features. Then the decision about the belonging class (either sea bass or salmon) is taken.

At this point, the critical question is how the decision is made, that is, how the classification can be implemented in an automatized system, and thus, how an automatized system can learn. What is required in this scenario is to maximize the classification accuracy, thus minimizing the number of wrongly classified fishes. Basically, a *decision boundary* has to be set-up by means of a model of the data (fishes) that, considering a set of features in input, can tell (with a certain value of accuracy) to which specie that specific pattern of features belongs to. The model on which the decision is based is not known a priori but has to be *learned* by the system from a set of examples (i.e., from a number of images of fishes belonging to the two different species). The way in which machines (e.g., computer-based systems) can learn from examples will be discussed in the next paragraph.

How machines learn: elements of the statistical learning theory

The *Statistical Learning Theory* was developed by Vapnik and Chevronenkis (Vapnik, 1995; Vapnik & Chervonenkis, 1974; Vapnik, 1999) in order to provide a framework for studying the problem of learning, in the sense of gaining knowledge. In particular, this theory is focused on explaining from a statistical point of view how we make inferences, predictions, and we construct generalizable models from a set of data (Bousquet, Boucheron, & Lugosi, 2004). Essentially, the main steps of pattern recognition are:

1) observation of a phenomenon;
2) construction of a model of the phenomenon;
3) make predictions using the model.

Basically, from a broad point of view, these steps retrace the general steps of scientific research. However, the role of the Statistical Learning Theory is to formalize this process, in order to make it understandable from a statistical perspective, while the role of Machine Learning algorithms is to automate it (Bousquet et al., 2004). Therefore, expression like "Machine Learning" or "Machine Learning algorithms" refer to algorithms that can automatically implement the learning process, as formalized by the statistical learning theory. As discussed in the example presented in the previous paragraph, essentially the machine learning process is based on extracting features from a noisy set of data (i.e., training data), and using those critical features to build a model of the data. Finally this model can be used to make prediction about new data (i.e., test data). In the example above, the algorithm can be trained on a number of examples of sea bass and salmon (images) in order to learn which pattern of extracted features is critical for distinguishing between the two species. Then, this model of the data can be applied to new examples (i.e., fishes never seen before) in order to classify them (prediction).

Given the great amount of different problems that can be addressed by automatic learning methods, is not possible to highlight one specific algorithm as the universally best choice. Indeed, if an algorithm A outperforms algorithm B for some classification problem, then there must be problems in which B outperforms A. Essentially, this concept is formalized in the so-called *No Free Lunch theorem* (Wolpert & Macready, 1997), which states, in essence, that if there is no assumption on the relation between the past (i.e., training data) and the future (i.e., test data), it is impossible to make predictions. So that, basically, this theorem states that learning is impossible without prior knowledge. This is the basic idea of the *Ugly Duckling theorem* (Watanabe, 1969) which says that things are all equivalent each other without prior knowledge, so that classification is impossible in absence of a theoretical framework. Thus, the importance of the Statistical learning Theory lies on the fact that it provides a framework allowing a statistical approach to the classification problem, and thus giving us the possibility to solve practical problems with the use of automatized machine learning systems. As Vapnik wrote: *"nothing is more practical than a good theory"* (Vapnik, 1998).

Moreover, an *a priori* restriction on the possible phenomena that are expected is needed, in order to be able to generalize a model. Without a priori restrictions, generalization is not possible, thus no algorithm is better than others. This is important also for the neuroimaging practice because, as will be discussed in the present manuscript, assumptions about the algorithm and the model which can explain data are necessary.

Indeed, different approaches can be adopted in training a machine learning algorithm. Basically, these approaches can be divided in two categories: supervised and unsupervised learning. In the following paragraphs, a description of these two approaches will be presented, in particular referring to the applications in cognitive neuroscience.

Essentially, the use of machine learning algorithm allows to deal with the classification problem also in neuroimaging. This has been possible since the introduction of multivariate analysis technique, as previously discussed. Thus, multivariate pattern based classification allows researchers to classify pattern of brain activity or of anatomical measures to specific mental states (e.g., experimental conditions) or to general conditions (e.g., Alzheimer patient vs. control), through a reverse inference process. Multivariate fMRI/MRI analysis methods can be divided into two classes (O'Toole et al., 2007):

- multivariate exploratory analysis: when the association between the pattern of brain measures and experimental conditions is not *a priori* given (unsupervised learning);
- multi-voxel pattern analysis: when the correct class labels are given before classification (supervised learning).

Supervised learning

Supervised learning (learning with a teacher) is based on creating a link between a set of features (e.g., a pattern of activation) and an associated *label* identifying the class of the example (e.g., patient vs. control; Bray et al., 2009). This type of learning is called "supervised" as the information about the correct label assigned to each example is given *a priori.* Basically, a set of data is used by the algorithm to learn (estimate) the parameter of a data-driven model which represents the relation between patterns of features and class labels. Then, the classifier can use the learned parameters to predict the class of new (i.e., never seen before) examples. There are two different types of "supervised learning problem", based on the type of variable that has to be predicted. In the case of a discrete variable, such as two or more class labels (e.g., object vs. animal) or different diagnosis (e.g., Alzheimer Disease vs. Frontotemporal Dementia), it is called *classification.* By contrast, when the predicted variable is continuous (e.g., age), it is called *regression*

(Haynes, 2015). The great majority of neuroimaging studies adopting a multivariate approach use supervised learning algorithms for performing binary classification.

Among the supervised learning algorithms, different classification methods are available, each one making different assumptions about the data and proposing a specific model through which the relation between features and target labels can be described (Bray et al., 2009). Two main classes of supervised classifier algorithms can be identified: linear and nonlinear classifiers. Linear classifiers aim at predicting the labels of new examples through a linear combination of features. These classifiers are widely adopted in neuroimaging studies, because of their simplicity and easy interpretability (Haynes, 2015), together with good performances. Moreover, the use of “simple” classifiers usually provides a better generalization.

The most commonly used linear classifiers in neuroimaging studies include the linear discriminant analysis (LDA; e.g., Haynes & Rees, 2006), the logistic regression (LR; e.g., Yamashita, Sato, Yoshioka, Tong, & Kamitani, 2008), and the Support Vector Machines (SVM; Haynes & Rees, 2006; Kamitani & Tong, 2005; LaConte, Strother, Cherkassky, Anderson, & Hu, 2005). Although is not possible to tell which classifier is the best choice in general, some studies (Cox & Savoy, 2003; Mourão-Miranda, Bokde, Born, Hampel, & Stetter, 2005; Pereira, Mitchell, & Botvinick, 2009) have compared the performance of the abovementioned linear classifiers. The findings showed that SVM outperforms LDA in some cases (e.g., Cox & Savoy, 2003), and the level of performance is the same of LR (Pereira et al., 2009), however SVM is more efficient in managing a high number of features as commonly experienced in neuroimaging (Bray et al., 2009).

By contrast, nonlinear classifiers use a nonlinear function to describe the relation between features and labels, leading to more complex and less interpretable models. Moreover, the performance level of these techniques does not seem to be better than the linear algorithms (Cox & Savoy, 2003).

For all the considerations expressed above, among the wide range of machine learning algorithm that can be used to perform a binary classification, supervised learning classifiers are preferred to unsupervised classifiers (described below). Moreover, among these, linear SVMs are those most common in neuroimaging study both using functional (i.e., fMRI) and structural (i.e., MRI) data.

Unsupervised learning

While supervised learning requires both features and class labels of the training examples to be specified, *unsupervised learning* requires only the features and tries to learn the structure of the data, and whether data can be distinguished through a pattern analysis (Bray et al., 2009). Thus, the target here is to teach an algorithm to learn without any *a priori* given association between data and target labels. For this reason, the analyses methods belonging to this category are called *exploratory analyses.* Three main approaches are those most commonly adopted in fMRI multivariate exploratory studies: Independent Component Analysis (ICA), Principal Component Analysis (PCA) and clustering algorithms (e.g., k-means clustering, hierarchical clustering). Exploratory analysis, even if overcoming some lacks of the voxel-wise analysis, have some disadvantages and limits. For example, the interpretation of the resulting "components" is made *post-hoc* and is completely left to the subjectivity of the experimenter, moreover there are no systematic ways to relate patterns of activity to class labels or experimental conditions.

After this brief discussion about the different multivariate machine learning approaches, a number of issues are still open. However, it is clear that in general the introduction of data-driven approaches can overcome some of the limitations of the traditionally adopted methods. For this reason, in the last years an increasing amount of studies have adopted multivariate techniques for the analysis of neuroimaging data in an innovative fashion. The introduction of pattern based multivariate classification methods can play a key role in the future in the analysis of neuroimaging data. Indeed, the great advantage of these technique is that they can provide a direct link between brain activity or anatomy and explanatory variables. Therefore, the neuroscientific perspective to the study of brain and mind can be reversed in a way that can lead to a more deep understanding of how the brain react to the environment and how researcher can ask the brain in order to predict behavioral or clinical variables. These advantages are possible thanks to the key concepts on which the multivariate analysis approach lies. These crucial aspects, and their differences with the standard approach will be discussed in the next paragraph.

UNIVARIATE vs. MULTIVARIATE ANALYSIS OF NEUROIMAGING DATA

Traditionally, the statistical analysis of both structural and functional MRI data has been performed in a univariate sense. Basically, univariate fMRI data analysis is realized correlating an experimental variable (e.g., task conditions) with the brain activity (i.e., the BOLD signal; Ogawa, Lee, Kay, & Tank, 1990) through a General Linear Model (GLM) computed separately for each voxel in the brain. This approach is called *Statistical Parametric Mapping* (SPM; Friston et al., 1995). Under this perspective, the GLM can be represented with the general formula: $Y_j = \beta_1 X_1 + \cdots + \beta_v X_v + \varepsilon_j$ (Penny, Friston, Ashburner, Kiebel, & Nichols, 2011), where:

- Y_j is the intensity of the BOLD signal in the voxel j;
- $X_{(1 \dots v)}$ are explanatory variables (regressors);
- $\beta_{(1 \dots v)}$ are (unknown) coefficients.

In these approach, data are modeled assuming the shape of the *Hemodinamic Response Function* (HRF), that is convolved with each task condition and used for detecting correlations. This method is voxel-wise, that is it analyzes each voxel timeseries independently for each of the thousands of voxels in the brain. The output of this analysis is a statistical map indicating voxel-by-voxel the intensity of the BOLD signal related to the regressors of interest, and thus where (i.e., in which voxels) the activity related to the experimental variable is located (Worsley & Friston, 1995). Therefore, this method addresses the research question about whether BOLD signal is related to an experimental variable, and where the activity is mainly located in the brain, starting from a predefined generative model that is fitted to the data separately for each voxel (Bray et al., 2009; Haller, Lovblad, Giannakopoulos, & Van De Ville, 2014). The mechanism through which the inferential process is realized with this method can be called *forward inference* (Haller et al., 2014; R. Henson, 2006). Basically, this process tell us how well the model fits (i.e., explains) the data, and provide inference on which brain areas are supposed to be involved (i.e., more active) in a specific mental process or pathological condition. However, this analysis is based on a correlational logic, so that it is not possible to infer the causality of the resulting relation between brain activity and mental states. Moreover, the brain correlates of a number of mental/cognitive processes are not localized in a specific brain area, but are represented across patterns of brain activity (or patterns of anatomical indexes such as Grey Matter volume).

Recently, new analysis methods have been introduced in the neuroimaging scenario providing a different prospective through which the brain can be studied. These methods, referred to the general framework of non-parametric approaches and belonging to the category of *multivariate analysis,* may improve our knowledge on how the brain react to stimuli (Di Bono & Zorzi, 2008). However, as discussed in the present dissertation, have many advantages in the investigation of anatomical data, as well. The main strength in the multivariate approach, if compared to the traditional one, is that the information of thousands of voxel is analyzed simultaneously, in order to detect the distinctive pattern of activity (or anatomical differences) related to a cognitive process or a general mental state. Therefore, rather than asking how much each voxel's activity is related to an experimental condition, in multivariate analyses the question is about whether there is a pattern of brain activity (i.e., spread across many voxels) which discriminate between experimental or clinical conditions (Bray et al., 2009). The key idea that critical information about a cognitive process or, in general, a mental state, are not encoded is a single location but are distributed across many brain regions, is certainly more biologically plausible, given the widespread nature of brain responses (Uttal, 2015). Moreover, in multivariate analysis, no (parametric) models are preliminarily specified (e.g., HRF model) as there can be strong differences across subjects (Handwerker, Ollinger, & D'Esposito, 2004; Thomason, Burrows, Gabrieli, & Glover, 2005), but models are learned from a set of data in order to predict the explanatory variable on new data (Haller et al., 2014). This approach, can therefore be defined *data-driven* as it starts from the detection of patterns of information encoded in the brain, in order to come to the mental process related to the detected pattern. In other words, it is possible to infer a mental state by looking at neuroimaging data. This inferential process is called *reverse inference* (Cox & Savoy, 2003; Haller et al., 2014; Poldrack, 2006, 2008) or *inverse inference* (Haynes & Rees, 2006; Haynes, 2015). Basically, this mechanism allows to infer a mental process or condition starting from the brain. From a logical point of view, this reasoning is based on the *fallacy of the converse,* that is affirming the antecedent from the consequence. This fallacy is based on the fact that a specific brain pattern is not the only one possible when performing a specific cognitive task or belonging to a specific clinical population. Indeed, the detected brain pattern and the co-occurring mental state are not necessarily connected by a causal link (Haynes & Rees, 2006). This is the main reason why such inference mechanism has been criticized in the neuroimaging field (Poldrack, 2006). Anyway, given that research in cognitive neuroscience is aimed at explaining behavioral events, rather than inferring the general laws that rule behavior, the use of methods relying on reverse inference is more robust than that of the traditional inferential approach (Di Bono, 2009). Importantly, the possibility of combining information encoded in many voxels, can potentially increase the analysis sensitivity, as univariate approach can fail to map neural basis of conditions having a distributed pattern of effects on the brain (Bray et al., 2009; Davis & Poldrack, 2013). Therefore, when the difference between two mental processes lies on the functional connectivity pattern, multivariate maps might be more informative if compared to standard univariate statistical maps (Sato et al., 2008). Finally, the use of multivariate analysis techniques allow to predict mental states from brain imaging data, by recognizing the pattern of brain activity related to a specific state (e.g, a cognitive process) or a general state (e.g., a neurological condition, such as a diagnosis of Alzheimer Disease).

As discussed in this chapter, multivariate methods have changed the way of asking research questions in neuroimaging studies. From a standard approach focused on localizing the brain correlates of cognitive functions to a more flexible approach allowing to understand in a more exhaustive way which information are encoded in the brain, how they are represented. The two approaches, highlighted in the previous paragraphs, within the family of multivariate methods, can be distinguished on the basis of the research question that they allow to deal with. For the aims of the present dissertation only one question was investigated, in different applications. In particular I was interested in analyzing whether and how consistently patterns of functional and structural brain indexes can predict mental states. This question is addressed by the pattern-based classification methods, also called Multi-Voxel Pattern Analysis (MVPA), that will be discussed in the next chapter.

Chapter 2

MULTI-VOXEL PATTERN ANALYSIS: AN OVERVIEW

INTRODUCTION

In the last years, a growing number of studies have adopted multivariate analysis approaches for investigations in a wide range of scientific areas. Recently, the multivariate approach known as Multi-Voxel Pattern Analysis (MVPA) has strongly entered the common practice in neuroimaging studies. In Figure 2.1 the number of neuroimaging publications in which MVPA methods have been adopted or discussed between 1992 and 2016 is shown. Considering the year in course, the number of already published neuroimaging papers using MVPA (515) is almost double of those published only three years ago (277). This data indicates the increasing interest on the multivariate techniques for the analysis of neuroimaging data.

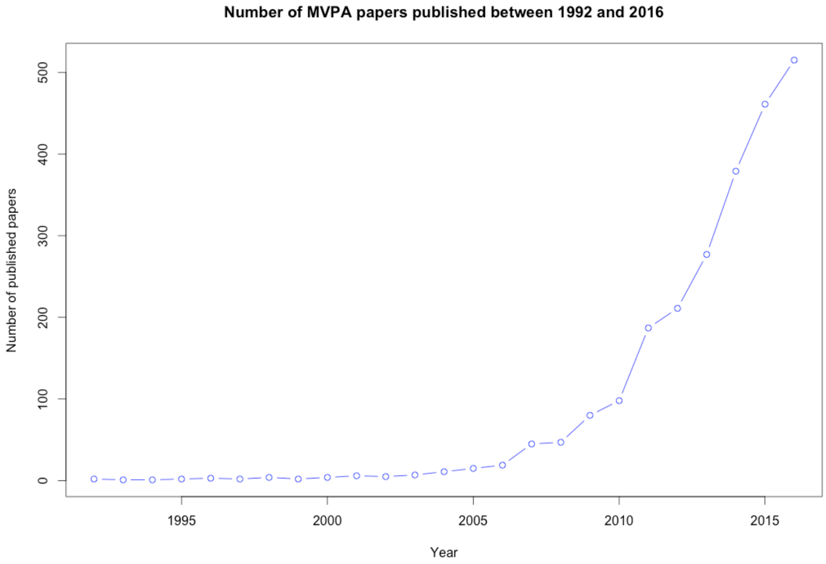
Number of MVPA papers published between 1992 and 2016

Fig. 2.1. Number of published papers in neuroimaging field, using MVPA (Source: pubmed.org - Key: (Multi-Voxel Pattern Analysis[Title/Abstract]) OR MVPA[Title/Abstract]"). Research made on November, 7th 2016.

As previously discussed, the multivariate approach has shown several advantages if compared to the standard voxel-wise approach to functional and structural neuroimaging data. These advantages suggest that the increasing trend in the use of MVPA in the neuroimaging field will not stop in the next years, and that this approach will potentially become a gold-standard for the statistical analysis of neuroimaging data.

In neuroimaging field, MVPA has been used to investigate both structural and functional neural correlates of a number of mental states and traits including conscious and unconscious racial attitudes (Cunningham et al., 2004), emotional states and self-regulation (Beauregard, Lévesque, & Bourgouin, 2001), personality traits (Canli & Amin, 2002), and many others.

Basically, MVPA has been mainly employed in two different ways. The first application is to decode the specific mental state of a subject while he is performing task. An example of this application is the study by Haxby and co-workers (2001), considered as the first important study adopting a pattern recognition approach with fMRI data. In this study authors investigated the discriminability of the brain activity pattern associated to the processing of faces or objects pictures. Basically, the kind of perceived stimulus is inferred based on the information derived from the fMRI BOLD signal, thus leading this procedure to be called "brain reading" or "mind reading" (Haynes & Rees, 2006) or "decoding" (e.g., Hebart, Görgen, Haynes, & Dubois, 2015) techniques.

The second most common application of MVPA is more related to the clinical field and concern the prediction of the general state of an individual on the basis of functional or structural data. For example, several studies have shown the possibility to discriminate prodromal Alzheimer Disease (AD) patients from healthy subjects (e.g., Davatzikos, Fan, Wu, Shen, & Resnick, 2009), showing that these techniques can be useful in the automatized early diagnosis of AD (Salvatore, Battista, & Castiglioni, 2016).

In both cases, the analysis is focused on the classification of *examples* as belonging to one among two classes. The classification is achieved by training an algorithm in discriminating between the two classes, on the basis of a pattern of provided *features*. This procedure produces an accuracy map or a single accuracy value telling how much two stimuli or two neurological state can be distinguished and/or where the information for the classification is mainly encoded.

Types of classifier

An important step that has to be performed before training a classifier is choosing which classifier. Indeed, many different kinds of classification algorithms are available, and can fit to different research questions or different dataset characteristics. One of the most simple classification algorithms is the *nearest-neighbour*. This procedure is based on finding the example in the training set that is most similar to a new example that has to be classified. Thus, the class label of that training example is assigned to the test example. Other classifier types are the *discriminative* and the *generative* models. Discriminative models are those that predict the class of a new example by setting the models parameters learned from the training set. By contrast, generative models basically learn a model that can generate an example of one specific class (Pereira et al., 2009). An exhaustive comparison of the different typed of classifiers is not consistent with the aims of the present dissertation (for further details see e.g., Bray, Chang, & Hoeft, 2009; Pereira et al., 2009; Poppenk & Norman, 2012).

A crucial question, when choosing a classifier, is about the shape of the function, that is the choice between a linear and a non-linear classifier. Despite this issue is strongly related to the specific research question and features set employed, in neuroimaging there is a general “rule” that has to be considered. Indeed, as specified in other subsections of the present manuscript, typically in fMRI/MRI studies the number of features exceeds the number of examples. In other words, for each example (e.g., a single MRI scan of a patient) several thousands of features (i.e., voxels) are considered. This can entail the model to be excessively fitted to training data, and thus not able to generalize the accuracy performance to new unseen data (see the paragraph about the “overfitting” problem). The risk of overfitting can be reduced by adopting a les flexible classifiers. Therefore, among the variety of available different classification algorithms, the great majority of MVPA studies has adopted linear classifiers. The main reason for this choice is that the use of linear classifiers redues the risk of overfitting, as this approach is less flexible than non-linear approaches (Misaki, Kim, Bandettini, & Kriegeskorte, 2010). In neuroimaging, the most common classifier is the linear support vector machine (SVM; Cortes & Vapnik, 1995; Vapnik, 1995), that has been adopted in a number of studies (see e.g., Kamitani & Tong, 2005; LaConte, Strother, Cherkassky, Anderson, & Hu, 2005; Mourão-Miranda, Bokde, Born, Hampel, & Stetter, 2005) as it has been demonstrated to be a valid and efficient approach for pattern analysis, both with functional (e.g., Pereira, Mitchell, & Botvinick, 2009; Schmah et al., 2010) and structural (e.g., Casanova, Hsu, & Espeland, 2012; Ung et al., 2014) neuroimaging data. In the following subsections, we will discuss issues a user has to deal with when facing binary classification problems, as well as methodological aspects of the application of linear SVM algorithms.

BINARY CLASSIFICATION: THE SUPPORT VECTOR MACHINES (SVM)

SVM classifiers have reached increasing popularity in the last 15 years. One reason of the diffusion of these algorithms, among all the available approaches, is the relative ease in implementing the analysis and in interpreting the results, if compared to other classifiers. The workflow of a MVPA study employing an SVM algorithm has to follow several steps (Haller et al., 2014; Haynes & Rees, 2006; Norman, Polyn, Detre, & Haxby, 2006; Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012; Pereira et al., 2009). There could be slight differences in the pipeline when using structural or functional neuroimaging data, as shown in Figure 2.2.

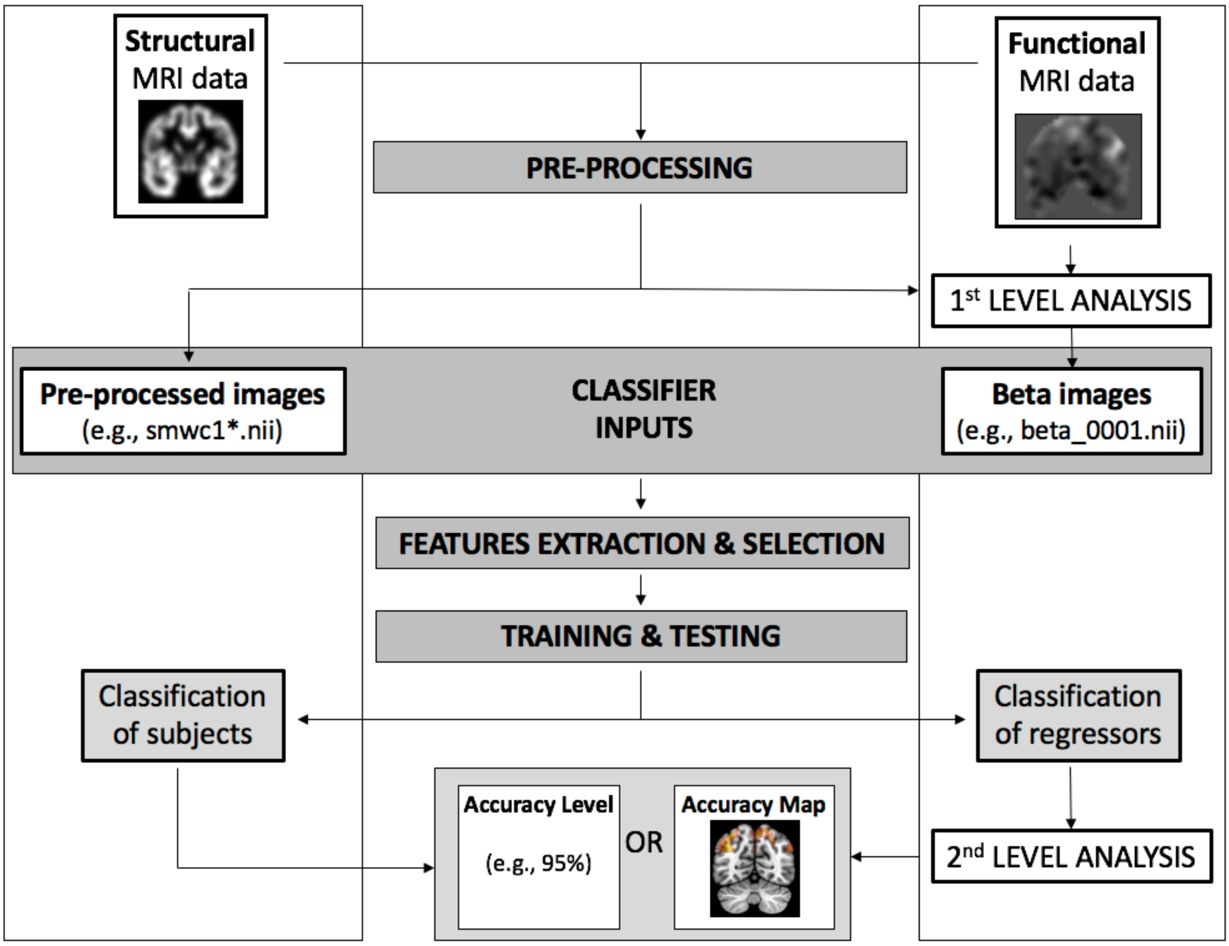

Fig. 2.2. Graphical overview of a basic workflow of a MVPA study using functional or structural MRI data.

However there are some main steps that should always be performed. These steps include the selection and extraction of features, the use of the selected features to train and test a classifier, and finally the evaluation of classifier performance. In the following, each of the main steps will be discussed.

Feature extraction and Feature selection

The word "*features*" in the pattern recognition field refers to independent variables (predictors) that are used to predict the class an example belongs to, where an "*example*" is defined as a pattern of independent variables (Pereira et al., 2009). In neuroimaging, each brain voxel considered in the analysis is a feature, and examples can be the type of stimuli investigated (e.g., Object vs. Animal) in the case of functional data (i.e., fMRI) or a global condition (e.g., AD patient vs. Control) in the case of structural MRI data.

Once participants' scans (*examples*) have been collected, data have to be prepared for training the classifier. First, the original data are transformed into a set of "features", which will be subsequently used for the pattern classification. This is the so-called *features extraction.* In the neuroimaging practice, this step consists in transforming each brain scan in a vector of *features* with each value corresponding to a single voxel (e.g., GM volume or activation in that voxel; Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012). Therefore, an *example* $\boldsymbol{x}$ will be in the form of a vector $x = [x_1, x_2, ..., x_v]$, where v is the number of voxels and $x_{1...v}$ are the voxel-by-voxel values (Pereira et al., 2009). Based on the type of neuroimaging data that are analyzed, the feature vector could encode the amount of Grey Matter (GM) or White Matter (WM) or the cortical thickness (in structural MRI data), rather than a measure of brain activity (in fMRI data) for each considered voxel (depending on the analysis approach, e.g., whole-brain vs. ROIs-based).

Then, the most useful features are selected and those redundant or minimally important are discarded (Orrù et al., 2012). This process is called *features selection.* The reduction of the number of features used for the classification can allow to:

- maximize the classifier performance. A less noisy set of features can foster the highest reachable accuracy in discriminating between the classes;
- make results more understandable. For example, focusing only on brain areas highlighted by previous literature as related to a certain cognitive process, can make the interpretation of results easier. Indeed, when the number of features is excessively greater than the number of data-points in the training set, the model will be barely interpretable (Friedman, Hastie, & Tibshirani, 2001);
- save time and computational load. Sometimes MVPA algorithms (e.g., the searchlight) can require a high computational power, as well as a great amount of time to achieve results. So that, loading less data (i.e., less features) can speed-up the process.

As abovementioned, one possible way to perform a feature selection is to perform a Regions-Of-Interest (ROIs) analysis. ROI analysis consists in selecting specific brain regions derived from the literature, to focus the analysis only on a subset of voxels. For example, only areas that the literature has shown to be related to the psychological process under investigation, are selected. This option is advantageous when there are specific spatial priors that can be used to drive the results, on the other hand the disadvantage is that the analysis is restricted to a small set of spatial hypotheses and thus the possibility of finding unexpected results is less probable. The second option for the feature selection is to use automatic algorithm for the features elimination. These methods include the *recursive features elimination* (RFE), in which features giving a small contribution to the classifier (i.e., having small SVM weights; see below) are iteratively discarded (De Martino et al., 2008) or the Greedy algorithms (see Guyon & Elisseeff, 2003), and the *sparse methods,* which consists of classification and regression algorithms that perform an implicit selection of a subset of features that will be used in the model (Carroll, Cecchi, Rish, Garg, & Rao, 2009; Grosenick, Greer, & Knutson, 2008; Yamashita, Sato, Yoshioka, Tong, & Kamitani, 2008). Also Independent Component Analysis (ICA) and Principal Component Analysis (PCA) can be used for the feature selection.

Training and Testing

As abovementioned, the target of a MVPA study is to reach a good classification accuracy in discriminating between examples belonging to different classes (two in the binary case). This objective can be achieved following two steps: estimating and optimizing the model parameter through the use of a set of training data, and then testing the generalization of classifier performance to new examples (i.e., not used for the training). As previously stated, given its diffusion in neuroimaging application of MVPA methods, here we will focus on how an SVM linear classifier deals with a binary classification problem (i.e., where examples belong to two classes).

Support Vector Machines (SVM) are linear classifiers, so that the classification function is based on a linear combination of the features (Pereira et al., 2009) and has the form

$y = g(w_1x_1 + \ldots + w_nx_n)$

where:

- $x_1 \ldots x_n$ is the set of features;
- $w_1 \ldots w_n$ are the weights associated to each feature, estimated during the training phase;
- y is the function of the pattern of features.

In the training phase, the SVM algorithm learns specific parameters called *weights* (w), from a set of training data. The *weights* can be defined as the influence of each feature on the classification performance. For instance, the value of individual weights indicates the relative importance of each feature in the classification process (Bray et al., 2009). Basically, a SVM is trained on a set of data in order to estimate a function called “decision function” or “hyperplane”, which optimally discriminates between the two considered classes (e.g. AD patients vs. Controls). In Figure 2.2, a representation of the SVM performance is presented in a pattern recognition problem between two classes (red and blue spheres; e.g., AD patients vs. Controls). This simplified case refers to a three-dimensional case (Fig. 2.2a), where only three features (i.e., voxels) are considered (one for each axis), and the position of each example (either a red or a blue sphere) in the k-dimensional feature space (in this case k=3) is defined by its pattern of features.

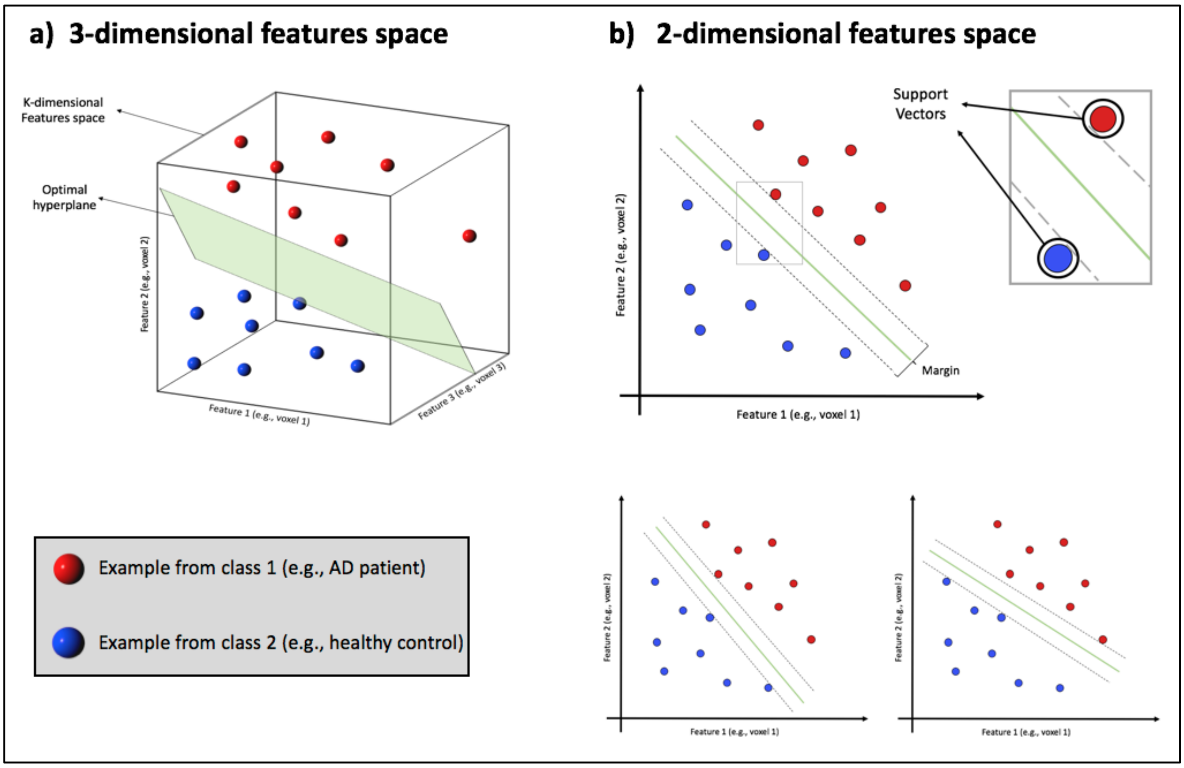

Fig. 2.3 a) Illustration of a 3-dimensional feature space in which 2 classes of examples (blue and red spheres) are separated by a linear SVM hyperplan (in light green). b) Illustration of a separating hyperplane (green line) in a 2-dimensional space, with a maximal margin determined by a linear SVM. Dashed lines are the boundary hyperplanes of the two classes. The portion of plan including the closest examples to the line is zoomed-in on the right: the circled examples are the support vectors. The margin is the distance between the closest examples of the two classes. In the lower right, two different possible separating hyperplans for the same examples are illustrated. The optimal one is the one with the maximum margin.

The two classes of examples in Fig. 2.2 can be separated by different hypothetical lines (i.e., functions). The target of the SVM, in the training phase, is to find the hyperplan that maximizes the distance between the examples lying closest to the separating plane (and thus the most difficult to classify) and the plane itself. These examples (i.e., vectors of features) are called *support vectors* and are those maximally influencing the classifier. In Figure 2.2b the same classification problem is illustrated, but in a more understandable two-dimensions representation, and support vectors are highlighted (circled examples). Therefore, in this simplified two-dimensional features space, learning a linear classifier means learning a line that separates points (examples) as accurately as possible, the so-called *decision boundary* (Pereira et al., 2009). The same principle is applied in a k-dimensional features space.

In the next step, the testing phase, the estimated model (the hyperplan) is used to predict the class of new examples (not involved in the training phase). At this point it is important to assess the robustness and generalizability of the classifier performance. Indeed, a good performance in fitting the model to the training data is not necessarily followed by a good classification accuracy of the examples in the test set. One possible cause of this mismatch is the so-called *overfitting*.

The overfitting problem

When dealing with complex systems of features, a classifier could make a performance close to 100% of accuracy on the training set but the performance can be poor when the model is applied to new datasets. This problem is commonly referred as "overfitting" (Duda et al., 2001). Overfitting can result from a biased analysis procedure. Indeed, to avoid biased estimation of generalizability, the data used as training set should not overlap with those used for testing.

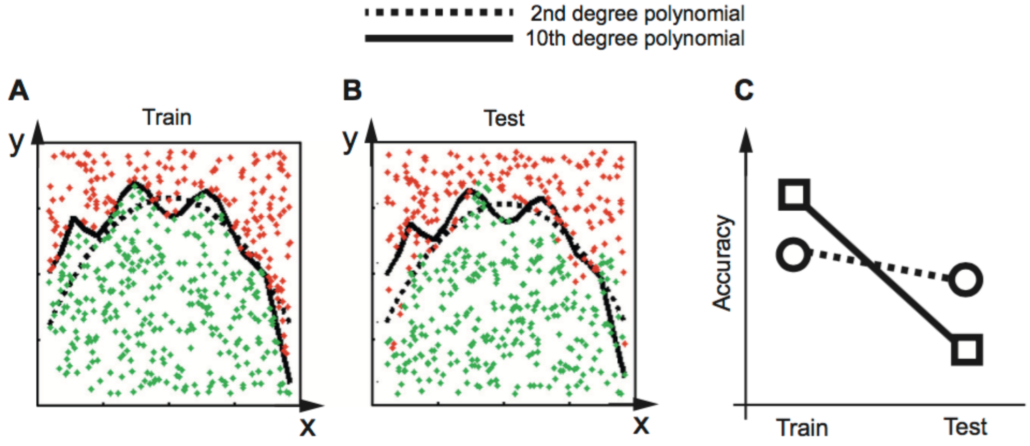

Fig. 2.4 (Haynes, 2015) Example representing two different classifiers used in a two-class dataset (red and green dots). The 10th-degree classifier shows a good performance on training data (actually due to overfitting) but it does not generalize to the test set. The 2nd-degree classifier has a lower classification accuracy but it is less overfitting-prone.

However, even in presence of an unbiased design, when the number of features is high (as often happens in neuroimaging where the number of features exceeds the number of examples), the classifier can be excessively fitted to the training data, leading to a lack of generalizability (Pereira et al., 2009). Four main options can be chosen to deal with this phenomenon:

1) use of different (i.e., independent) sets of data for training and testing. However this solution increases the amount of data that should be acquired;
2) reduction of the number of features are used for the classification. As discussed above, one way is to perform the feature selection step, often crucial to improve the classifier performance. Another possibility is to consider some dimensionality reduction method, such as Principal Component Analysis (PCA; Hansen et al., 1999) or Independent Component Analysis (ICA; Calhoun, Adali, Hansen, Larsen, & Pekar, 2003);
3) choice of the classifier. In many neuroimaging papers a linear classifier is chosen in order to avoid overfitting.
4) use of cross-validation (CV) schemes, where a subset of data is iteratively left out from the training phase and used as test set, and the resulting classification accuracy is averaged across repetitions (Lemm, Blankertz, Dickhaus, Muller, & Müller, 2011). In the next paragraph cross-validation techniques will be discussed.

Dealing with overfitting: Cross-Validation schemes

From a general point of view, all cross-validation techniques are based on a common principle, that is the independence between training and test set within a single iteration. This means that data are split in subsets: one is used in the training phase in order to build a model and to estimate model parameters, the other one (i.e., not considered in the training process) is used to evaluate the quality of the learned mapping between the patterns of brain activity (or the pattern of GM volume when using structural data) and the experimental conditions. This process is repeated a variable number of times (iterations) based on the number of created subsets. Cross-validation can be considered a method that allow the user to test the validity, robustness and reliability of a classifier performance. In the following, the most common CV-schemes are presented.

Hold-out cross-validation

In this cross-validation scheme a subset of data is used as training set, in order to estimate the model parameters, and the remaining set of examples is used as test set for the validation of classifier performance. Thus, in the testing phase, the model built on training data is applied to the new data in order to predict the class of each data point, and the number of correctly classified examples is collected.

K-fold cross-validation

A common cross-validation design in neuroimaging is the *k-fold cross-validation*. In this scheme, *k* is the number of partitions the dataset is split into (*N* is the total number of examples). Then, *k* steps are performed in which the classifier is trained on *N-k* folds and tested on one. The algorithm stops when each folds has been used as test set once. Then, the classification accuracies are averaged across the *k* iterations (Bray et al., 2009) and a final value of accuracy is reached.

Leave-One-Out cross-validation (LOO-CV)

Another cross-validation scheme often used in neuroimaging studies is the LOO-CV. This scheme can be considered a particular case of the *k-fold cross-validation* scheme where *k* is set to *1*. For example, one subject per group can be iteratively left out and the classifier in each iteration is trained on *N-1* examples per group, and tested on the left-out pair. Based on the number of examples, this scheme can potentially require more computational effort (Bray et al., 2009). Leave-one-out cross-validation scheme actually refers to a family of cross-validation designs, as the decision on what to leave out can be customized. One of the most common implementation of this scheme in fMRI studies is the Leave-One-Run-Out cross-validation, where the classifier is trained on all but one runs and tested on the remaining run (the process is repeated until each run is left out once).

Other possible cross-validation schemes that will not be discussed in the present manuscript are referred to the *jackknifing* or the *bootstrapping* techniques (Efron & Tibshirani, 1993), and the *reproducibility resampling* (Strother et al., 2004).

The important concept is that, independently from the employed cross-validation design, users should be aware of possible problems related to the analysis of functional and structural MRI data, either related to the nature of the data itself (e.g., the temporal correlation of fMRI scans) or to the kind of analysis approach adopted (e.g., in the case of MVPA, the overfitting problem).

MVPA classification approaches

As shown in Figure 2.2, the decoding (classification) analysis can be realized on preprocessed images (e.g., in the case of structural data) or on beta images, representing the brain activation related to a specific stimulus type (in the case of functional data). When analyzing functional data in a task-based design, the decoding can be performed with cross-validation schemes at the subject-level (i.e., examples are betas resulting from the first level analysis of a single subject) or at the group-level (i.e, the dataset includes betas resulting from the first level analysis of different subjects). In all these cases, the results of the analysis are related to the type of chosen approach. One decoding option is the *whole-brain* analysis, in which each voxel in the brain is used as input feature for the classifier. Another possibility is the *ROI-based* analysis, discussed in a previous paragraph as it can be considered a way to reduce the number of features. An interesting alternative to these two analysis approaches has been introduced by Kriegeskorte and colleagues (2006) and is called *searchlight.* In this analysis, classification is performed separately for each "searchlight" sphere (or square) of a specified radius (e.g., 3 voxels radius). The analysis is repeated until each voxel in the brain has been used as central voxel of a searchlight sphere (see Etzel, Zacks, & Braver, 2013).

These different decoding approaches lead to different results format. In the case of a whole-brain the result will be a single value of accuracy that indicated the classification performance considering all the voxels in the brain. Similarly, an ROI-based analysis will result in one value of accuracy for each of the selected ROI. Whereas, in a searchlight analysis the result is expressed as an accuracy map in which, for each considered voxel, an accuracy value is reported indicating the accuracy reached in the multivariate analysis performed on the searchlight sphere having that voxel as the central one.

In the following paragraph, the next step of testing the significance of the obtained result will be discussed.

Interpretation of classifier performance

After performing a multivariate analysis for a classification purpose, the output should be properly interpreted. Indeed, it is important to note that pitfalls may be encountered when interpreting the results of classification experiments. For example, it is not possible to directly relate the information contained in a voxel or cluster to the information encoded in neurons (Haynes, 2015), as should be done in general with techniques involving the analysis of BOLD signal (Logothetis & Wandell, 2004), also in the univariate perspective.

For instance, it is possible that a voxel sample the signal coming from a large vessel that brings blood to a large population of neurons without sharing a direct anatomical correction (Gardner, 2010; Haynes, 2015). Moreover, there are many factors that have to be evaluated when interpreting the overall output of a classifier (Allefeld & Haynes, 2014), including the experimental design and the size of the training set.

However, on the basis of the different MVPA approaches previously described (whole-brain, ROI-based and searchlight analyses), there are differences in the way output is presented. In the whole-brain analysis a single number is obtained, simply indicating the level of classification accuracy considering all the voxels in the brain volume. So that, in this case we would not have a real accuracy map, but only a general accuracy value (e.g., "80%"). The same happens with ROI-based analyses, with the only difference that one accuracy value is obtained for each of the ROIs used.

More interesting is the output of searchlight analyses. This approach provides maps in which for each point (voxel) one accuracy value is reported. This value indicates the classification accuracy of local clusters of voxels surrounding that point. Therefore, each accuracy value does not indicate the accuracy reached in that specific voxel, but the accuracy reached in the cluster centered on that voxel (Haynes, 2015). This important aspect should be taken into account as information might be distributed in many ways within the cluster, not necessarily including the central voxel (Etzel et al., 2013).

Statistical evaluation of classifier performance

Once a classification result has been obtained (e.g., either an accuracy value or an accuracy map), the statistical significance of that result should be tested. In other words, we should test whether the classification accuracy reached has been obtained by chance. From a statistical point of view, a significant result is one that allow us to reject the null hypothesis that the analyzed features do not encode information about the variable of interest (i.e., the class). In the classical NHST (Null-Hypothesis Significance Testing) framework, the statistical significance is determined by quantifying the probability (p-value) of obtaining the observed result (in this case the classification performance) under the null hypothesis that there is no difference between the considered classes, and thus that the classification accuracy is 1/2 = 0.5 (50%; if we have had four classes, it would have been 25%). Consider the example of tossing a coin for n times. In this case we have two classes (heads vs. tails) and assume that "heads" is obtained in k out of n trials. The null hypothesis is that the coin is well-balanced, and thus that in each trial there is a 50% probability for each class to occur. The alternative hypothesis is that the coin is unbalanced, so that one class is more probable than the other. At this point, we want to test whether the k successes we obtained are significantly different (or above/below) than chance-level. Each trial in this case is independent, so that it can be modeled as a Bernoulli trial with 50% probability of success under the null hypothesis (Pereira et al., 2009), so that one possibility is to perform a Binomial test. The binomial test gives us the probability that the k successes occurred by chance. However, in most MVPA neuroimaging studies, the assumption of independence is not satisfied, mainly because of the use of cross-validation designs, where the same examples are used both as training and test data. Although this consideration, the use of binomial test has been common practice for years (Kohavi, 1995). Recently, several papers (e.g., Noirhomme et al., 2014; Stelzer, Chen, & Turner, 2013) have faced this issue and have argued that when a standard cross-validation design is applied, the independence of trials is not provided, thus the distribution of observations cannot be approximated to a binomial distribution. Therefore, in these cases, a binomial test is not appropriate and results can be biased. According to a recent paper (Noirhomme et al., 2014), however, when an independent leave-one-out cross-validation design is adopted (e.g., using an independent test set, the examples of which are not used during the training phase) a binomial test can be applied.

In the lots of cases in which assumptions of independence do not hold, a good option is the use of a permutation test (Etzel, 2015; Etzel & Braver, 2013; Pereira et al., 2009). Permutation tests (for neuroimaging application see Nichols & Holmes, 2002; Nichols & Holmes, 2003) are based on minimal assumptions (P. Good, 2013) and their efficacy in the context of classification has been proven (Golland & Fischl, 2003; Golland, Liang, Mukherjee, & Panchenko, 2000; Stelzer et al., 2013). Basically, permutation test is based on shuffling the class-labels in the training set, and training and testing the classifier with the relabeled data. This procedure is repeated a number of times (usually at least 1000) using the same CV-scheme adopted in the original classification, in order to build an empirical distribution of accuracy results under the null hypothesis (Pereira et al., 2009; Stelzer et al., 2013). Then, the accuracy obtained using the correct labels is compared to the built distribution (Etzel & Braver, 2013), and thus significance value is obtained. Although the principle of permutation test is the one described above, different permutation designs are possible, for example one can shuffle the class labels across all dataset (*dataset-wise permutation schemes*), or for each independently (*fold-wise permutation schemes*; Etzel, 2015; Etzel & Braver, 2013).

The abovementioned approaches to significance testing can be applied to both single-subject and group-level decoding. In neuroimaging studies, one common practice is first to realize a within-subject classification (i.e., on not normalized, unsmoothed data), and then perform a group analysis on the resulting accuracy maps (e.g., in the case of searchlight approach) using a standard approach such as one sample t-test. This procedure, however, can be realized only when the classification is possible within-subject, that is when a task is performed and functional data are acquired. In MVPA studies where the examples that has to be classified are subjects (e.g., patients vs. controls, using structural data or resting state functional data), the most common approach is to perform a permutation test (or a binomial test when the assumption of independence holds).

Applications of MVPA

As mentioned above, many applications of MVPA have been tested in the last 20 years. Following the important study by Haxby and co-workers (2001) several studies investigated the association between visual object categories and patterns of activity in ventral temporal cortex (e.g., Carlson, Schrater, & He, 2003). However, applications of MVPA on functional MRI data have covered a wide range of fields not related with visual perception, such as psychopathy (Pridmore, Chambers, & McArthur, 2005), criminal tendencies (Raine et al., 1998), drug abuse (Childress et al., 1999), product preferences (McClure et al., 2004), decision-making (Heekeren, Marrett, Ruff, Bandettini, & Ungerleider, 2006), and many others. Importantly, for the aims of the present thesis, some studies have focused on aspects related to possible forensic applications, such as lie detection (Davatzikos et al., 2005) or intentions (Haynes et al., 2007). Particularly interesting for the potential future application to forensic settings is the use of MVPA for the detection of memories (Rissman, Chow, Reggente, & Wagner, 2016; Rissman, Greely, & Wagner, 2010; see Section B and Chapter 7).

In addition, the analysis of structural MRI brain scans has been used to discriminate between healthy controls and patients with different pathologies, including Alzheimer's Disease (Davatzikos et al., 2009), fragile-X syndrome and autism (Hoeft et al., 2011), psychosis (Gothelf et al., 2011), depression (Costafreda, Chu, Ashburner, & Fu, 2009), multiple sclerosis (Weygandt et al., 2011) and so on.

In the next chapters two applications of MVPA will be presented, one using structural MRI data (Section A) and one using functional MRI data (Section B).

SECTION A

DETECTING COGNITIVE RESERVE FROM THE BRAIN: A MVPA APPLICATION

Introduction

The concept of "reserve" has been introduced about thirty years ago (Katzman, 1993; Katzman et al., 1989) to account for the disjunction between the level of brain pathology and its clinical manifestations (Stern, 2009) and to explain the differences in age-related cognitive decline. According to a general definition (OED, 2016), "reserve" refers to the *"supply of a commodity not needed for immediate use but available if required"*. Some studies have reported discrepancies between cognitive functioning in aging and brain signs of pathology at a post-mortem examination (Ince, 2001). In cases like these, the amount of reserve is assumed to determine the relation between pathological brain alterations and manifestations of symptoms. Within the concept of reserve, a number of models have been proposed, such as Cognitive Reserve (CR; Stern, 2002), Brain Reserve (BR; Satz et al., 1993), Brain Maintenance (BM; Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012), and neurocognitive scaffolding (STAC; Park & Reuter-Lorenz, 2009). In the present manuscript, we will focus mainly on CR and BR models and on their relation.

Actively building reserve: the Cognitive Reserve model

The expression Cognitive Reserve (CR) is defined as "*the ability to optimize or maximize performance through differential recruitment of brain networks, which perhaps reflect the use of alternate cognitive strategies"* (Stern, 2002). CR refers to individual differences in cognitive processing as a function of lifetime intellectual activities (as well as environmental factors) that can explain differences in the susceptibility to cognitive impairment due to a neurological condition (Stern, 2002, 2009). Thus, basically, the whole set of variables including education level, QI, learning experiences and knowledge that a person acquires throughout life contributes to his CR (Mondini et al., 2016). Among the components of reserve, CR is the "active" one, as it suggests that the brain actively try to cope with brain pathology both using cognitive processes in a more efficient way and attempting to compensate the damage (Stern, 2002).

According to Stern (2009), the way in which CR may influence the clinical expression of neurological pathology (e.g., Alzheimer disease, AD), is represented in Figure A.1. Starting from epidemiological and imaging studies on the relation between CR and AD-related cognitive decline, the moment in which pathology begins producing clinically visible cognitive symptoms (the so-called "point of inflection") is later in time in persons with high CR if compared to persons with low CR. The logical consequence is that a person with higher CR will meet clinical diagnostic criteria for AD at a later stage of neuropathology, thus, from a clinical perspective, the onset of the disease is delayed. The second implication is that, once cognitive decline begins, it is faster in the person with higher CR. Therefore, the impact and efficacy of a cognitive training is different in AD patients with high and low CR. Indeed, as recently demonstrated, cognitive training is less efficient in AD patients with high if compared to patients with low CR, because of the different underlie stage of pathology at the moment of clinical manifestation (Mondini et al., 2016).

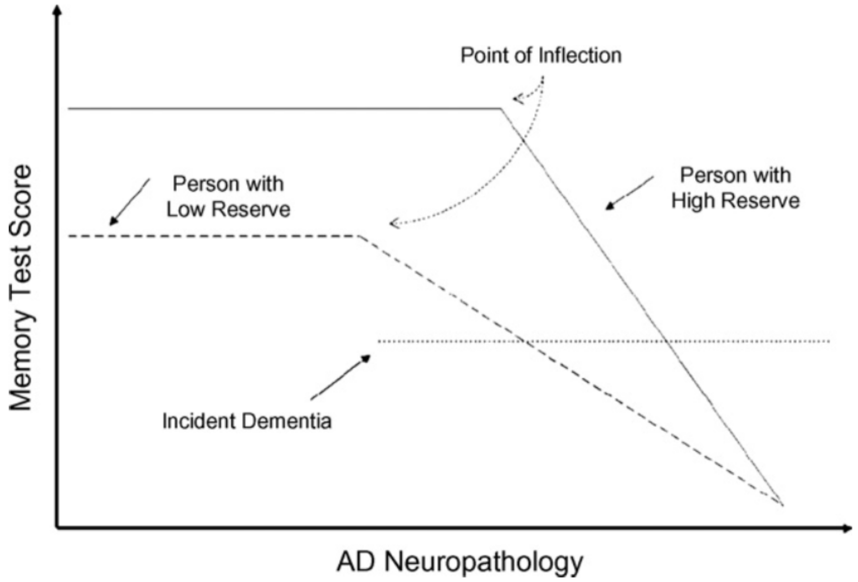

Fig. A.1. (Stern, 2009) Theoretical representation of the role of CR in mediating the relation between AD neuropathology and its clinical expression. The x-axis represents the progression of AD neuropathology over time, while the y-axis represents the level of cognitive performance (i.e., cognitive functioning). The basic assumption is that the progression of AD pathology is the same independently from the level of cognitive reserve. What is modulated by the CR level (high vs low) is the moment when the clinical expression of symptoms (point of inflection) begins. With higher CR this point is postponed.

It has been largely demonstrated that life experiences, such as educational and occupational attainment, or involvement in leisure-time activities (e.g., social or, in general, cognitively relevant activities) have been associated with a reduced risk of degenerative diseases such as dementia (Scarmeas, Levy, Tang, Manly, & Stern, 2001; Stern et al., 1994; Valenzuela, Sachdev, Wen, Chen, & Brodaty, 2008; Wilson, 2009) and better progress in other neurological conditions such as traumatic brain injury (Fay et al., 2010), Parkinson's disease (Poletti, Emre, & Bonuccelli, 2011), multiple sclerosis (Sumowski et al., 2016; Sumowski, Chiaravalloti, Wylie, & Deluca, 2009; Sumowski & Leavitt, 2013).

Taken together, all these studies suggest that the threshold for functional decline can be modulated by life experiences (e.g., education, occupation, leisure time activities), therefore it is not fixed throughout the life-span. A slightly different perspective is suggested by "passive" models of reserve, the so-called Brain Reserve (BR), that will be presented in the following.

Quantifying Cognitive Reserve

Given the theoretical nature of CR, there are no direct measures through which it can be quantified. However, in literature, several variables have been indicated as "proxies" of CR, that is, measures used as indirect indicators of CR level. The most commonly used proxies of CR are:

- educational attainment (e.g, Arenaza-Urquijo et al., 2013; Fitzpatrick et al., 2004);
- occupational achievement (e.g., Garibotto et al., 2008; Staff, Murray, Deary, & Whalley, 2004; Stern et al., 1995);
- socio-economic status (e.g., Scarmeas & Stern, 2003);
- intelligence (premorbid IQ; e.g., Sole-Padulles et al., 2009);
- leisure activity (e.g., Nikolaos Scarmeas & Stern, 2003; Sole-Padulles et al., 2009);
- bilingualism (e.g., Schweizer, Ware, Fischer, Craik, & Bialystok, 2012)

Moreover, several composite measures of CR have been introduced in literature. These tools take into account the combined contribution of several variables such as education, leisure-time activities, occupation and so on. Examples of these tools are the Cognitive Reserve Index questionnaire (CRIq; Nucci, Mapelli, & Mondini, 2012) or the Cognitive Reserve Scale (CRS; León, García-García, & Roldán-Tapia, 2014; Leon, Garcia, & Roldan-Tapia, 2011).

Among the abovementioned indicators, educational attainment is the most widely studied, and in many cases educational level and CR level are considered as the same concept (Valenzuela & Sachdev, 2006a, 2006b). It has been argued that education may foster CR by contributing to the creation of alternative cognitive strategies (Manly, Byrd, Touradji, & Stern, 2004; Stern, 2002). Moreover, it may promote synaptic growth (Katzman, 1993), also because of the lifetime period in which educational path begins is the childhood, a critical temporal window for brain plasticity plastic. So that, education may have effects on brain structure and contribute to brain protection. Indeed, it has been widely demonstrated that education modulates the relation between neuropathology degree and neuropsychological test performance (Bennett et al., 2003; Dufouil, Alpérovitch, & Tzourio, 2003; Rentz et al., 2010).

From Cognitive Reserve to Brain Reserve: two sides of the same coin?

Opposite to CR model, the Brain Reserve model (BR; Satz et al., 1993) suggests that quantitative (passive) measures, such as the overall size of the brain, the number of neurons or the number of synapses, explain the differential susceptibility to brain pathology, on the basis of the concept of brain reserve capacity (BRC). Under this perspective, functional decline begins when the BRC is depleted by an high level of neuropathology.

Despite BR and CR models have been developed as mapping distinct and independent aspects of reserve, a number of studies suggest that boundaries between the two models are thinner than expected. Thus, variables that contribute to CR may have a role in BR, as well. For example, higher CR seems to be associated with a reduced rate of hippocampal decline in aging (Valenzuela et al., 2008).

Some recent studies (e.g., Bartrés-Faz & Arenaza-Urquijo, 2011) tried to shed light on the correspondence between BR and CR models suggesting that a complex relation exists between the passive and active components of reserve. For example the model of Brain Maintainance (Nyberg et al., 2012) suggests the existence of a combination of both genetic and environmental factors, such as particular life experiences, that can have a protective role on brain. Barulli and Stern (2013) argued that, taken together, CR, BR, and other models such as brain maintenance, should be considered as different perspectives in explaining the same concept, that is the resilience to pathology. Lövdén and co-workers (2010) suggested that when people have to perform tasks, they use cognitive strategies taken from their own experience and knowledge. However, when they experience for a long time a mismatch between the efficacy of these strategies and tasks' requests, brain plasticity comes into play to compensate for this mismatch and this may lead to structural modifications. This mechanism is similar to that theoretically defined by Stern (2009) as neural compensation. Thus, essentially, the brain changes itself to properly deal with challenges coming from environment.

Cognitive Reserve and brain structure

Several accounts of the effects of environment on brain structure have been proposed in the last century, after the introduction of the enriched environment as an experimental variable (Rosenzweig, 1966; Rosenzweig, Bennett, & Krech, 1964). Many studies on rats have shown that environmental stimulation can act on the structure of adult brain in many ways, such as promoting dendritic arborization or neurogenesis (e.g., Cummins, Walsh, Budtz-Olsen, Kostantinos, & Horsfall, 1973; for a review see van Praag, Kempermann, & Gage, 2000). In addition, a number of studies on humans have shown that experience can have effects on brain structure. Volumetric brain changes have been documented following prolonged life experiences, such as years of intellectual stimulation (Coffey, Saxton, Ratcliff, Bryan, & Lucke, 1999), literacy (Carreiras et al., 2009), and musical learning (Gaser & Schlaug, 2003; Pascual-Leone, 2001). Moreover, involvement in physical, social and cognitive activities seems to have an effect on brain structure (Seider et al., 2016) also after relatively short periods of time. For example, Erickson and co-workers (2011) showed an increase in hippocampus size after six months of aerobic exercise. Increases in hippocampal volume has been shown also in the well-known study on London taxi-drivers (E. a Maguire et al., 2000), after intensive learning of city maps. Training-induced structural brain changes have been demonstrated also after three months of training on juggling skills, showing increased GM volume in temporal and parietal cortices. Moreover, a number of investigations on elders found effects of cognitive or physical training on brain structure, involving several measures of brain volume such as the regional volume of GM (e.g., Boyke, Driemeyer, Gaser, Büchel, & May, 2008) and WM (Colcombe et al., 2006), as well as increased regional cortical thickness (Engvig et al., 2010). This finding is consistent with the idea that brain plasticity is an intrinsic property of the nervous system and it is maintained throughout all life-span (Bartrés-Faz & Arenaza-Urquijo, 2011; Greenwood & Parasuraman, 2010; Pascual-Leone, Amedi, Fregni, & Merabet, 2005).

Only a few neuroimaging papers have explicitly investigated the relation between measures of CR and brain structure in elders, showing that higher levels of CR in healthy elders were associated to larger brains (e.g., Bartrés-Faz et al., 2009; Bosch et al., 2010; Sole-Padulles et al., 2009), increases in GM volume in frontal and parietal areas (Bartrés-Faz et al., 2009) and reduced age-related hippocampal volume decline (Valenzuela et al., 2008). To our knowledge, few studies have focused on the relation between the most used proxy of CR (i.e., education) and the brain structure. For example, Arenaza-Urquijo and co-workers

(2013) found that a higher amount of years of education was related to higher volume in the superior temporal gyrus, insula and anterior cingulate cortex.

Liu and colleagues (2012) showed that more educated subjects had larger regional cortical thickness in transverse temporal cortex, insula, and isthmus of cingulate cortex than less educated subjects. A low age-related GM loss was also found in the left anterior cingulate cortex (anterior portion) in high versus low educated elders (Rzezak et al., 2015). Moreover, in two recent important studies on more than 300 participants each (Foubert-Samier et al., 2012; Steffener et al., 2016), education was associated with higher GM volume in a wide range of brain regions. However, some other investigations (e.g., Coffey et al., 1999) did not find significant relationship between education and brain volume.

Aims of this section

Although some studies have focused on the relation between CR proxies and the brain, the effects of cognitively relevant life experiences on brain structure is still unclear. Moreover, to our knowledge only one study (Steffener et al., 2016) used a multivariate approach, and no studies adopted a pattern recognition approach for the study of the relation between CR and BR. In Chapter 3 and 4 two neuroimaging studies are presented, in which Multi-Voxel Pattern Analysis (MVPA) techniques are employed to decode subjects' educational level starting from regional GM volume. In both studies, freely accessible MRI images databases have been used (i.e., the OASIS database in Chapter 3 and the IXI database in Chapter 4, see the following chapters for further details). These databases have been chosen because, among the set of demographic available data, the educational level of each participant was collected. Therefore, in the next experiments we investigate the relation between CR and brain structure in healthy elders (as this is the most studied population in this field) by analyzing the differences in patterns of GM volume between individuals with high and low education, considering education as an indirect indicator of CR.

Chapter 3

NEURAL CORRELATES OF COGNITIVE RESERVE: EDUCATION AND GREY MATTER VOLUME (I)

INTRODUCTION

Variables reflecting lifetime experiences have been extensively used in the literature as proxies of reserve capacity. Among these variables, education is the most commonly adopted in studies of Cognitive Reserve (CR; e.g., Stern, Alexander, Prohovnik, & Mayeux, 1992). Moreover, level of education, as well as sex and socioeconomic status, have been identified as having a role in changes in brain structure in aging (Coffey et al., 1999; Rzezak et al., 2015). A number of studies have shown a protective effect of higher levels of education against dementia (for a review see Fratiglioni & Wang, 2007). Indeed, longitudinal studies have demonstrated that subjects with higher levels of education (acquired during early stages of life) showed a diminished risk of dementia during elderly (Brayne et al., 2010; Letenneur et al., 1999; Ott et al., 1995; Stern et al., 1994).

However, the relation between education and brain structure remains unclear. Some studies (Arenaza-Urquijo et al., 2013; Sole-Padulles et al., 2009) found a positive correlation between educational level and measures of brain structure, indicating for example an increase in GM volume in higher educated individuals. By contrast, other studies (e.g., Coffey, Saxton, Ratcliff, Bryan, & Lucke, 1999) found no significant relationship.

In the present study, our aim is to investigate the neural correlates of education, considered as one of the proxies of CR, as well as one of the factors contributing to Brain Reserve (BR; Katzman, 1993). As specified in the introduction of the present section, BR is a passive model of reserve and it refers to inter-subject differences in the brain that allow some individuals to cope better than others with pathological brain conditions (Stern, 2009). These BR differences can be quantified by measuring brain size, or counting neurons or synapses. Nevertheless, life experiences can act on brain structure through several processes (e.g., neurogenesis, promoting resistance to apoptosis), basically promoting brain plasticity.

In addition, recent neuroimaging studies highlighted that more indirect indexes of neuropathology, such as anatomical changes (Liu et al., 2012), cerebral blood flow reduction (Kemppainen et al., 2008) and metabolic alterations (Hanyu et al., 2008) are found in people with higher level of CR suffering from dementia, despite their quite normal cognitive performance.

A recent population-based study (Foubert-Samier et al., 2012) on more than 300 participants, investigated the influence of several CR proxies (including education) on the passive component of reserve (BR). The authors found that education was significantly associated with increases in cerebral volume (involving both Grey and White Matter) mainly located in temporoparietal and orbitofrontal lobes, suggesting that education may contribute to the constitution of cerebral reserve capacity. As stated before in the manuscript, another study (Bartrés-Faz et al., 2009) larger GM volumes in frontal and parietal regions.

However, the role of education (considered one of the most important proxies of the active component of reserve; CR) in promoting the passive component (BR), is still unclear. Here, we investigated the structural brain differences due to educational level by means of two analysis approaches. First we performed a standard univariate analysis investigating the effects on educational level on GM volume, then we adopted a multivariate approach (MVPA) in order to test whether a machine-learning classifier can discriminate subjects with low educational level from those with higher education, on the basis of the pattern of GM volume.

MATERIALS AND METHODS

Description of the dataset

Cross-sectional structural MRI (T1) images used in this study were extracted from the freely accessible Open-Access Series of Imaging Studies dataset (OASIS, oasis-brains.org; Marcus et al., 2007). This dataset consists of 416 participants across the adult life span (age range: 18-94). The sample of individuals aged 60 to 94 years, included patients with a diagnosis of probable mild to moderate dementia of Alzheimer type (AD). In this dataset, the presence and the level of pathology were established by means of the Clinical Dementia Rating (CDR) scale (Morris et al., 2001; Morris, 1993). The CDR is an instrument that allows to rate cognitive impairment in six domains: memory, orientation, judgment and problem solving, function in community affairs, home and hobbies, and personal care. On the basis of different sources of information, including interviews, a global CDR score was derived for each subject. A global CDR score of 0 indicates no dementia, and CDR score greater than 0 indicates very mild (CDR=0.5; e.g., “mild-cognitive impairment”; (Petersen et al., 1999; Storandt, Grant, Miller, & Morris, 2006), mild (CDR=1), moderate (CDR=2), severe dementia (CDR=3). Moreover, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) score was provided as a measure of global cognitive status.

Participants

In order to be consistent with the aims of the present experiment, only participants aged between 60 and 94, with a CDR score of 0 and a MMSE score above 24 (indicating the absence of clinically relevant dementia) were selected from the OASIS dataset. The final sample was composed by 98 healthy participants (72 F). Demographic and clinical characteristics of the sample are presented in Table 3.1.

Variable	Mean	SD	Range
Sex (female/male)	72/26		
Age (years)	75.9	8.9	60-94
Education (years)	14.5	2.9	8-23
Education (level)	3.3	1.3	1-5
MMSE score	29	1.2	25-30

Table 3.1. Demographic and clinical sample characteristics. MMSE=Mini-Mental State Examination.

Education was coded in five levels. Level 1: less than high school graduation; level 2: high school graduation; level 3: some years at college; level 4: college graduation; level 5: beyond college education. In Table 3.2 the number of subjects for each education level is reported.

Number of subjects	**Educational level**				
	1	**2**	**3**	**4**	**5**
F	5	18	19	13	17
M	3	7	4	4	8
Total	8	25	23	17	25

Table 3.2. Number of subjects for each educational level.

MRI Data acquisition parameters

In the OASIS dataset, three to four T1-weighted MPRAGE images (Mugler & Brookeman, 1990) were acquired for each subject, on a 1.5-T Vision scanner (Siemens, Erlangen, Germany) within a single session. Acquisition parameters were optimized for gray–white contrast (TR=9.7 ms; TE=4ms; Flip angle=10°; TI=20ms; TD=200ms; Orientation=sagittal; Thickness=1.25mm [gap=0]; slice number=128; resolution [pixels]=256×256; voxel size=1×1mm).

MRI Data Pre-processing

All MRI scans were visually inspected for artifacts or structural abnormalities before voxel-based morphometry (VBM; Ashburner & Friston, 2000; Mechelli, Price, Friston, & Ashburner, 2005) analysis was applied by using SPM12 (Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London, UK; http://www.fil.ion.uncl.ac.uk/spm). For the VBM analysis the following steps were performed. First, all structural images were manually set the origin of the stereotassic space to the anterior commissure and realigned along the AC-PC (Anterior Commissure – Posterior Commissure) line. Then, images were segmented into GM, white matter (WM), and cerebrospinal fluid (CSF) and imported into a rigidly aligned space (Ashburner & Friston, 2000). The resulting GM segmented images were then iteratively warped into a study-specific template by means of a fast diffeomorphic image registration algorithm (DARTEL; Ashburner, 2007), available in SPM12 software. In this step, subject-specific deformation fields were used to warp the GM segmented images to the new space, as it has been previously demonstrated that this procedure can maximize accuracy and sensitivity (Yassa & Stark, 2009). The resulting warped GM images were then spatially normalized into the MNI space through an affine spatial normalization and modulated in order to ensure that the total amount of gray matter in each voxel was conserved after the registration (Ashburner & Friston, 2000; C. D. Good et al., 2001; Mechelli et al., 2005). Finally, the images were smoothed with an 8 mm full-width at half-maximum (FWHM) Gaussian kernel. This amount of smoothing is commonly adopted in VBM studies (Scarpazza, Sartori, De Simone, & Mechelli, 2013).

Therefore, the resulting smoothed, modulated, normalized data were then used as inputs for the VBM analysis.

Univariate analysis: Voxel-Based Morphometry (VBM)

A general-linear regression model (GLM) was built in order to test the effect of education level on the amount of GM volume in a voxel-by-voxel fashion. Previous VBM studies have demonstrated strong effects of age (e.g., Salat et al., 2004) and sex (Buckner et al., 2005; Marcus et al., 2007) on GM volume, therefore in the present experiment these variable were used as covariates of no interest in the regression model in order to remove their effect from the results. Moreover, the estimated Total Intracranial Volume (eTIV; for details on its calculation see Buckner et al., 2004; Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Marcus et al., 2007) was used to global scale data, in order to avoid biased results due to inter-individual differences in brain size.

Statistically significant effects of education on regional GM volume were identified using a FWE multiple-comparisons correction at voxel-level.

Multi-Voxel Pattern Analysis

The decoding analyses were performed using the Pattern Recognition for Neuroimaging Toolbox (PRONTO; Schrouff et al., 2013), running under Matlab R2014b. According with Schrouff and co-workers (2013), non-smoothed images were used as input for the multivariate analyses.

In this experiment we aimed at using a whole-brain MVPA technique in order to test the accuracy of a SVM classifier in discriminating subjects with "low" education from those with "high" education. Given that SVM is a binary classifier, as a first step we created two classes of subjects selecting from the whole sample only subjects with extreme educational levels (level 1, 2, 4 and 5). Thus, we categorized those with an educational level equal to 1 or 2 as belonging to the "low education" group (LED). By contrast, "Highly" educated (HED) individuals were, instead, those with an educational level equal to 4 or 5. Given the differences in the newly created sample sizes (low education = 33; high education = 42), we randomly extracted from the high education group 33 individual scans in order to match the size of the low education sample. Thus, the multivariate analysis was performed on 66 scans.

Two decoding analyses were performed. The first analysis was focused on the whole-brain GM volume (by using a whole-brain GM mask). As previously reported, a recent study by Foubert-Samier and colleagues (2012) investigated the effects of several proxies of CR (including education) on GM volume. In particular they investigated, on a huge sample of participants (n=331), the effects of educational level on the amount of regional GM volume. The areas showing a significant effect in the abovementioned study involved frontal, temporal, parietal, occipital and limbic lobe (see Table 3.3 for a summary of significant areas obtained). These areas were used to build 5 Regions-Of Interest (ROIs) masks (one for each lobe) using the Automated Anatomical Labeling (AAL) atlas as implemented in Pickatlas software (http://fmri.wfubmc.edu/ software/PickAtlas). Then, we repeated the decoding analysis, as previously described, focusing only on the areas that previous literature (Foubert-Samier et al., 2012) showed to be more sensible to effects of education on GM volume.

Region	**Lat.**	**Lobe**
Superior medial frontal gyrus	L, R	Frontal
Supplementary motor area	L	
Precentral gyrus	L, R	
Superior orbitofrontal gyrus	L, R	
Middle frontal gyrus	L, R	
Rectus	L	
Medial orbitofrontal gyrus	R	
Inferior frontal gyrus	L	
Superior frontal gyrus	R	
Middle orbitofrontal gyrus	R	
Inferior operculaire frontal gyrus	R	
Inferior orbitofrontal gyrus	R	
Superior temporal gyrus	L, R	Temporal
Fusiform gyrus	R	
Middle temporal gyrus	R	
Inferior temporal gyrus	R	
Anterior cingulate	L, R	Limbic
Supra-marginal gyrus	R	Parietal
Precuneus	L, R	
Paracentral lobule	L	
Postcentral lobule	L, R	
Lingual gyrus	R	Occipital

Table 3.3. Area derived from Foubert-Samier et al. (2012) that showed a more important grey matter volume increase in subject with high educational level compared with low level (adjusting for age, sex and Total Intracranial Volume). Lat. = lateralization.

In both kinds of analyses (whole-brain and ROIs-based) a Leave-One-Out Cross-Validation (LOO-CV) scheme was adopted, with one subject per group iteratively used as test set (see Figure 3.1). The performances of the classifier were measured both with the accuracy level (i.e. the rate of correct classifications) and the Area Under Curve (AUC) derived from a Receiver Operator Curve (ROC) analysis. Statistical significance of the whole-brain decoding accuracy was tested with a permutation test (Ojala & Garriga, 2009) as this test is preferable when the assumption of independence between test examples is violated such as when using a cross-validation scheme (Schrouff et al., 2013). Here the permutation was repeated 1000 times and a p-value lower than 0.05 was considered significant.

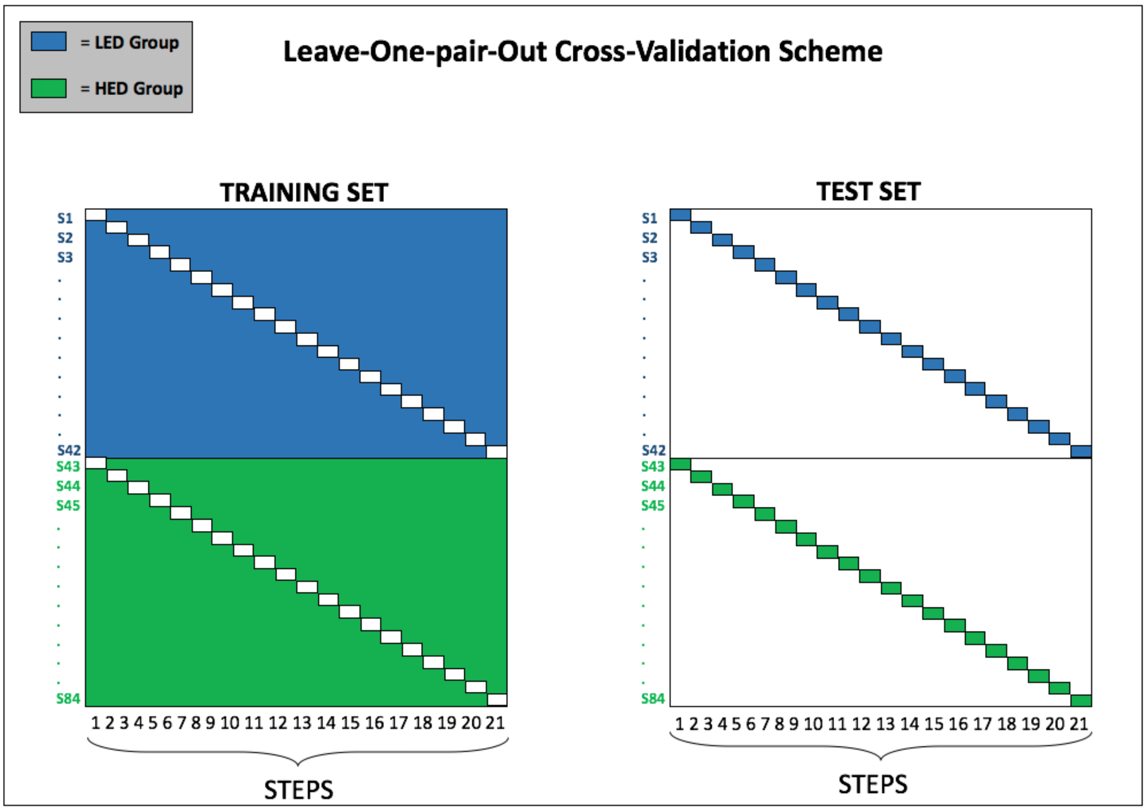

Fig. 3.1. Leave-One-Out Cross-Validation scheme adopted in this MVPA study. For each iteration (total iterations = 21), one subject from the LED and one subject from the HED were left out and used to test the classifier. For example, in the 1st step, the first subject of each group was left out from the training phase, in the 2nd step the second subject of each group was let out and so on. Thus, in the 1st step, subjects from S2 to S42 (LED) and from S44 to S84 (HED) were used to train the classifier, while subjects S1 (LED) and S43 (HED) were used to test the classifier performance. BLUE indicates LED group; GREEN indicates HED group; WHITE indicates unused subjects.

RESULTS

Voxel-Based Morphometry

As shown in Figure 3.2, no significant differences emerged in Total Intracranial Volume between subjects with different levels of education.

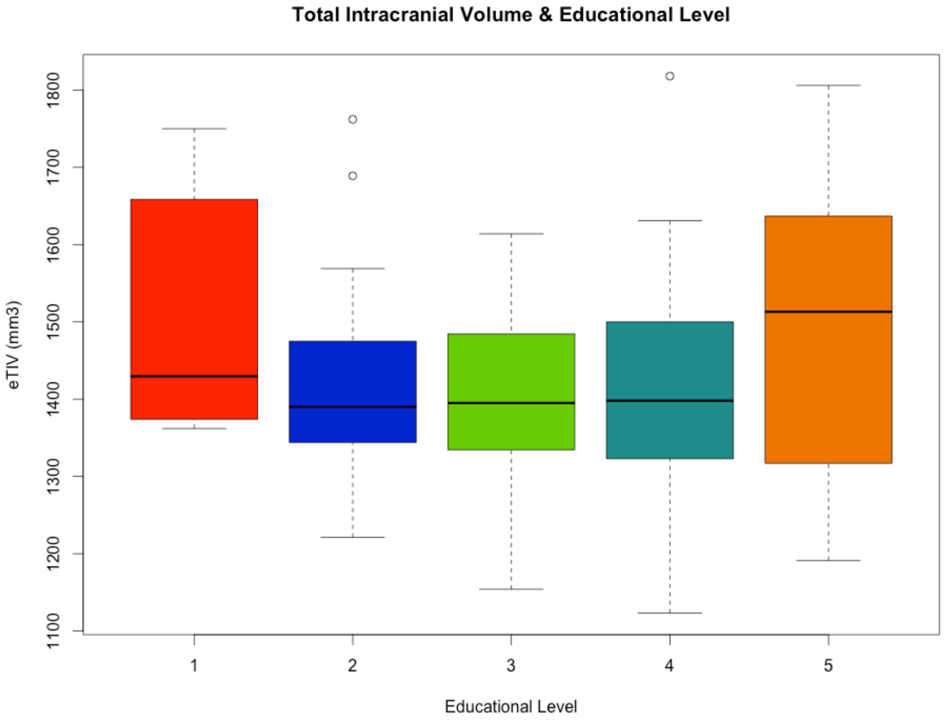

Fig. 3.2. Relation between Total Intracranial Volume (eTIV; in mm^3) and levels of education.

With the VBM analysis, an effect of an increase in the educational level on GM volume emerged in several brain regions, adjusting for age, sex and eTIV. Significant clusters were found in the left temporal lobe (parahippocampal gyrus, fusiform gyrus, hippocampus, lingual gyrus, temporal inferior gyrus; peak-voxel p=.002), right parietal lobe (precuneus, paracentral lobule, parietal superior, middle cingulum; peak-voxel p=.011) and in the right inferior temporal gyrus (peak-voxel p=.037). See Fig. 3.3.

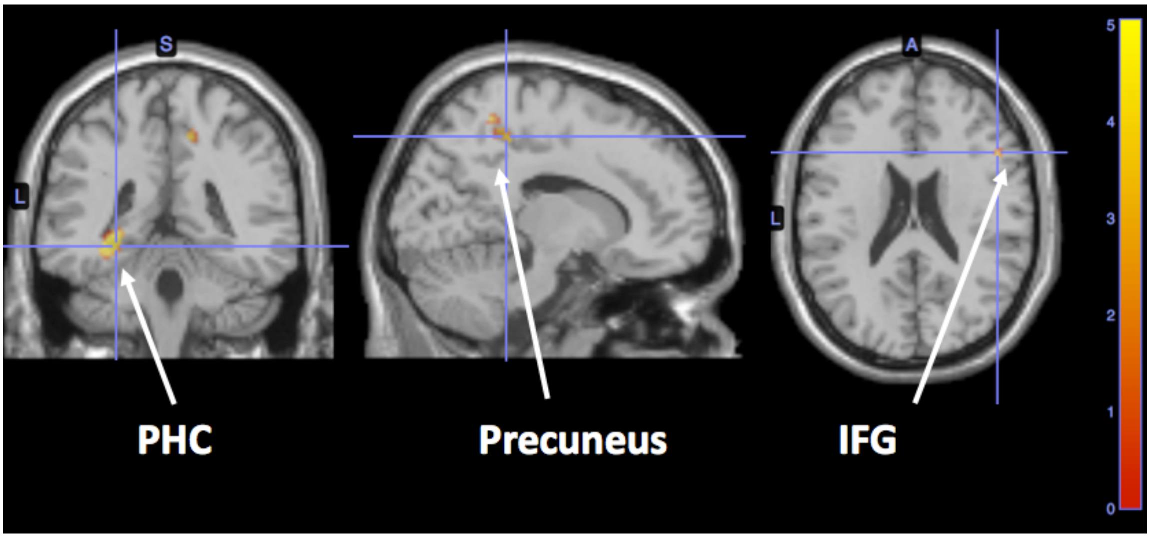

Fig. 3.3. Brain-areas showing significant effects of educational level on GM volume, adjusting for age, sex and eTIV. Statistical T maps are thresholded at p<.05 FWE-corrected for multiple comparisons at voxel-level. PHC= Parahippocampal cortex; IFG= Inferior frontal gyrus.

In Table 3.4 peak-voxel coordinates and T-values are reported of significant clusters (all reported results are significant at p<.05 FWE-corrected at voxel-level). By contrast, no brain-regions emerged as having higher GM volume with decreasing educational level.

Peak	**Lat.**	**Cluster size**	**T**	**MNI peak coordinates** (mm)		
				x	**y**	**z**
Parahippocampal Gyrus	L	415	5.04	-32	-44	-4
Precuneus	R	83	4.50	14	-45	54
Frontal Inferior Gyrus	R	53	4.13	50	20	21

Tab. 3.4. Peak-voxels of significant clusters FWE-corrected at voxel-level at p<.05. Lat.= lateralization.

Multivariate analysis

As specified above, two MVPA were performed investigating the accuracy of a binary SVM classifier in discriminating between LED and HED on the basis of the volume of GM, in a whole-brain and a ROI-based analysis respectively. Regions-of-interest for the second analysis were derived from a previous study by Foubert-Samier and colleagues (Foubert-Samier et al., 2012). All significant results have been tested through a permutation test with 1000 repetitions.

Accuracy of GM in discriminating LED from HED subjects: whole-brain analysis

The resulting accuracy in classifying brain images as belonging to the LED or the HED group was at chance level (48.5%), in the whole-brain analysis. So that, as shown in Table 3.5, about half of participants were correctly classified (17 LED and 15 HED).

		PREDICTED	
		LED	**HED**
TRUE	**LED**	17	16
	HED	18	15

Table 3.5. Number of predicted vs true subject-group (LED vs. HED) association in the whole-brain MVPA analysis.

The ROC curve and the AUC value (Fig. 3.4) confirm the poor performance of the SVM classifier algorithm in classifying subjects having high vs. low educational level on the basis of the whole-brain pattern of GM volume.

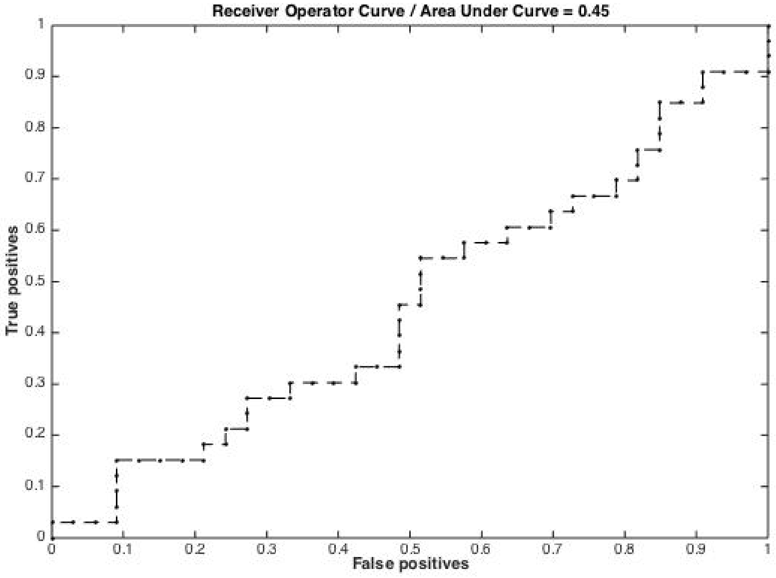

Fig. 3.4. ROC curve showing the classification performance in the whole-brain analysis (AUC=0.45)

Accuracy of GM in discriminating LED from HED subjects: ROI-based analyses

The five ROI-masks considered in this paragraph, overlaid on a standard anatomical template, are presented in Figure 3.5. No overlap is shown between the masks.

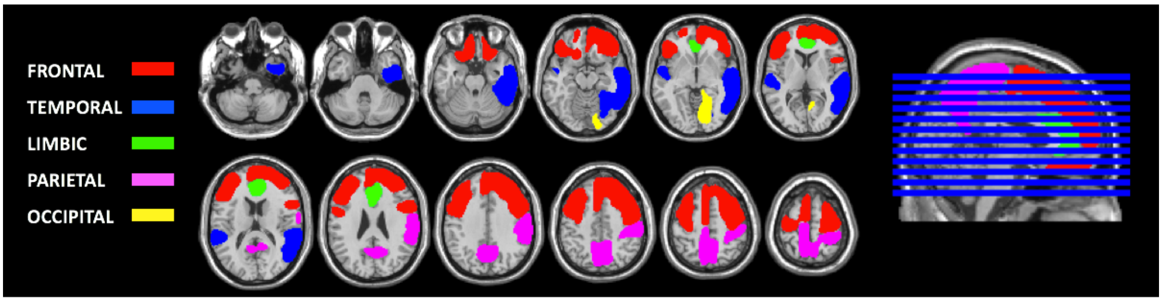

Fig. 3.5. Overlay of the five ROI-masks based on Foubert-Samier et al. (2012) on a standard template.

When using ROIs derived from the previous literature, we found strong differences in the performances of the classifier using the different ROIs-masks. The accuracy was basically at chance-level when using Frontal, Temporal, Limbic and Occipital ROI-masks as shown in Table 3.6.

ROI-Mask	Accuracy (%)	Significance
Frontal	50	n.s.
Temporal	37.9	n.s.
Limbic	48.5	n.s.
Parietal	75.8	p=.001
Occipital	45.5	n.s.

Table 3.6. Accuracy reached using each ROI-mask and significance. n.s.=not significant.

The performance of the SVM classifier strongly improved using the ROI-mask focused on parietal areas, leading to an accuracy value of 75.8% (p= .001). Twenty-six out of 33 LED individuals were correctly classified (see Table 3.7; sensitivity = 78.8%; p=.001), as well as 23 out of 33 HED subjects (specificity = 72.7%; p=.009).

		PREDICTED	
		LED	**HED**
TRUE	**LED**	26	7
	HED	9	24

Table 3.7. Number of predicted vs true subject-group (LED vs. HED) association in the ROIs-based MVPA analysis using Mask 4 (Parietal lobe).

In Figure 3.6 ROC curve and AUC value related to the performance of the SVM classifier using the ROI-mask 4 (Parietal lobe) are reported. The resulting AUC value is 0.80, indicating an accuracy of 80%.

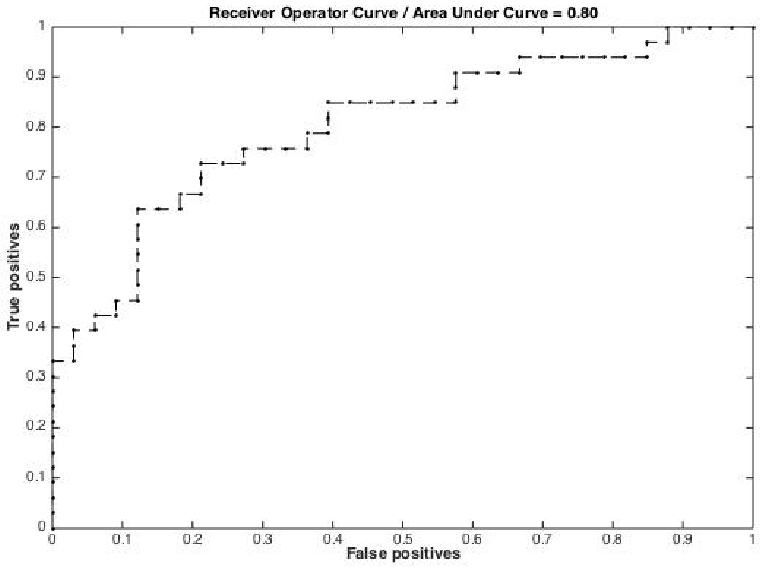

Fig. 3.6. ROC curve showing the classification performance in the ROIs-based analysis using Mask 4 (Parietal lobe). The corresponding AUC was 0.80.

DISCUSSION

In the present study we investigated the neural basis of Cognitive Reserve (CR) analyzing the effect of educational level (considered in literature as a proxy of CR) on GM volume, in a sample of healthy elders selected from a free database of MRI images (OASIS; oasis-brains.org). To our knowledge, no studies investigated the effects of education on GM through a Multi-Voxel Pattern Analysis approach. Thus, the main aim of this study was to test the performance of an SVM algorithm in discriminating subjects with high educational level from those with low educational level on the basis of their pattern of GM volume, by means of a MVPA. As a preliminary step, a standard univariate analysis was performed, showing effects of educational level on temporal, parietal and frontal regions. These results confirm data from previous studies suggesting that the prolonged cognitive stimulation received through formal education can involve changes in the brain structure in several regions (Foubert-Samier et al., 2012). This finding is consistent with the hypothesis of Brain Maintainance (Nyberg et al., 2012), that suggests the existence of factors (including education) that can have a protective role against brain aging and neurological pathologies

(Nyberg et al., 2012; Rzezak et al., 2015). Brayne and co-workers (Brayne et al., 2010) demonstrated that years of education had no effects on neurodegeneration or vascular pathology, but they acted moderating the effects of these pathologies on the clinical expression of symptoms. Moreover, several studies (e.g., Bennett et al., 2003; Roe et al., 2008) have demonstrated that educational attainment attenuates the cognitive symptoms of Alzheimer Disease (for a review, see Fratiglioni & Wang, 2007).

Importantly, on the basis of the present results, as well as of those from other cross-sectional design studies (e.g., Foubert-Samier et al., 2012), is not possible to infer the causal direction of the relation between educational level and GM volume. So that, we cannot determine whether the increase in GM volume in highly educated subjects is a consequence of formal education, or a preexisting characteristic. Indeed, the highlighted differences between high and low educated subjects could reflect previously existent differences in GM volume or in the progression of brain atrophy (Foubert-Samier et al., 2012). However, from a more general point of view, it has been demonstrated since the sixties (Rosenzweig, 1966; Rosenzweig, Bennett, & Krech, 1964) that neuronal changes occur as a consequence of learning and complex stimulation occurring in an enriched environment (Henriette van Praag, Shubert, Zhao, & Gage, 2005). Moreover, studies on GM density highlighted adolescence as critical periods for the structural brain changes in temporo-parietal cortex (Sowell et al., 2003). Therefore, it is plausible that the cognitive engagement required in formal education lead to changes in brain structure that can be measured in elderly.

On the other hand, there is still debate on which brain regions are more affected by the structural consequences of cognitive engaging activities. In the present study, the MVPA analyses showed performances at chance-level when considering all brain voxels and when focusing on frontal, temporal, occipital and limbic regions. A high level of accuracy was shown only when a parietal ROI-mask was used. This mask (together with the other masks used in the present study) were derived from the work by Foubert-Samier and collaborators (2012) and emerged as being the most sensible areas to effects of educational level on GM volume.

These results indicate that the most informative areas for the discrimination between participants with high vs. low educational level were located in the parietal lobe. A possible role of parietal regions in showing effects of education (i.e., cognitive stimulation) on GM volume was suggested by Draganski and colleagues (2006). These authors found that medical students, after three months of extensive learning, showed an increase in GM volume in posterior and lateral parietal cortex bilaterally, and in posterior portion of right hippocampus. Taken together with our results, these studies suggest that formal education seems to have a role in influencing the GM volume in parietal brain areas.

Therefore, the multivariate analysis indicates that the main structural differences between the brains of participants with low education and subjects with high education seems to be mainly encoded in parietal areas. By focusing on these areas, it is possible to classify subjects as having a low vs. high educational level with an accuracy close to 80%. Aydin and colleagues (Aydin et al., 2007) investigated with a VBM analysis the brain structure of a group of mathematicians, who had been working as academicians and found an increase in GM density in the parietal lobe, and this increase was highly correlated with the time spent as an academician. Moreover, in a study by Serra and co-workers (Serra et al., 2011), authors found that GM volume in parietal areas was reduced in Alzheimer patients with low educational level, if compared with patients with high educational level.

The main strength of these results is that the multivariate analysis was focused only on areas highlighted by a previous study (Foubert-Samier et al., 2012) investigating the relation between education and brain structure, so that, our analyses were driven by strong spatial priors. Therefore, the present study can be considered as a confirmation of previous univariate results obtained by Foubert-Samier and collaborators (2012). Moreover, we went a step further by performing a more subtle investigation in which we analyzed, within the results of previous literature, the most informative areas in discriminating subjects with different educational levels.

In conclusion, education is one of the main components contributing to reserve capacity (Yaakov Stern, 2009), together with occupation and leisure time activities. These variable have been mainly studied in relation with functional differences in individual cognitive functioning, and have been highlighted as variables having an impact on Cognitive Reserve (CR; Stern, 2002, 2009). However, cognitive stimulation (e.g., education) can also contribute to Brain Reserve (BR). BR refers to an increase in redundant neural networks (Yaakov Stern, 2002), also by promoting neurogenesis (J. Brown et al., 2003; Henriette van Praag et al., 2005). Thus, in view of all these results, we suggest that CR is possibly implicated in regional increases in GM volume, so that it can promote several "brain reserves", rather than a general increase in brain plasticity (Serra et al., 2011). Therefore, the present study contributes to the amount of literature suggesting that Brain Reserve and Cognitive Reserve should be considered as "two sides of the same coin".

Chapter 4

NEURAL CORRELATES OF COGNITIVE RESERVE: EDUCATION AND GREY MATTER VOLUME (II)

INTRODUCTION

In the previous study, we performed a Multi-Voxel Pattern Analysis on structural MRI data in order to test whether it was possible to discriminate subjects with low education from subjects with high education by investigating the volume of GM in their brain. This issue was addressed both analyzing all the voxels in the brain (whole-brain analysis) and selecting specific ROIs extracted from the literature. In particular, results from a study conducted on a large dataset (Foubert-Samier et al., 2012) were adopted to guide our analysis. Our results showed that the most informative brain regions in discriminating between high vs. low educated participants were located in the parietal lobe. Therefore, in the present explorative study we selectively focus our attention on the parietal regions in order to investigate in a more precise way where the information about the classification are encoded. Moreover, here we test whether our findings about the role of parietal lobe can be generalized to a different dataset.

MATERIALS AND METHODS

Description of the dataset

The MRI scans used in the present study were obtained from the IXI dataset (Information eXtraction from Images; brain-development.org), a free open-access database of nearly 600 structural MRI images of healthy individuals aged 20 to 86 years, acquired at three different hospitals in London (UK; Hammersmith Hospital, Guy's Hospital and Institute of Psychiatry).

Participants

In order to be consistent with the methods used in the previous study, here we chose only brain scans of subjects older than 60. Moreover, we decided to avoid using data acquired from different scanners, so that we chose only data collected at Guy's Hospital. Therefore, the final ample used in the present study included 109 subjects (64F; see Table 4.1 for descriptive statistics).

In the IXI dataset education was coded according with the educational system in UK, in 5 levels as follows. Level 1 indicated no qualifications; level 2 indicated a ten-years education (Ordinary-levels, GCSEs or CSEs); level 3: A-levels (comparable to high-school diploma); level 4 indicated further education after diploma, such as skill development course (e.g., City & Guilds qualification or National Vocational Qualifications - NVQs) and level 5 indicated University or Polytechnic degree. In order to be consistent with the coding system used in the previous study, we referred the 5 levels explained above to the 5 levels adopted in the previous study (OASIS dataset; see Chapter 3), structured as follows:

- level 1: less than high school graduation;
- level 2: high school graduation;
- level 3: some years at college;
- level 4: college graduation;
- level 5: beyond college education.

Variable	Mean	SD	Range
Sex (female/male)	45/64		
Age (years)	67.5	6.2	60-86
Education level (IXI)	3.5	1.6	1-5
Education level (OASIS)	2.7	1.3	1-4

Table 4.1 Demographic sample characteristics. Descriptive statistics are provided for both educational level coding systems adopted in the OASIS and IXI dataset.

In Table 4.2 the number of subjects for each education level is reported only for the coding system adopted for the next analyses (OASIS coding system). As it can be easily noticed, no subjects explicitly met the criteria to be assigned to level 5, according to OASIS coding system, that is no subjects were reported as having performed education steps after University degree. The reason is that the maximum educational level required in IXI dataset was the University degree. However, it is possible that subjects with level 5 in the IXI coding system had further education after degree.

Number of subjects	**Educational level (OASIS)**				
	1	**2**	**3**	**4**	**5**
F	23	5	14	22	0
M	10	4	7	24	0
Total	33	9	21	46	0

Table 4.2. Number of subjects for each educational level.

MRI Data acquisition parameters

Data used in the present experiment (extracted from the IXI dataset) were acquired with a 1.5T Philips Medical Systems Gyroscan Intera at Guy's Hospital in London (UK). The acquisition parameters were: repetition time/echo time = 9.813 ms, flip angle = 8, echo time = 4.603, number of phase encoding steps = 192, echo train length = 0, reconstruction diameter = 240, voxel size = 1.5x1.5x1.5 mm.

MRI Data Processing

As a first step, all MRI scans were visually inspected to detect artifacts or abnormalities. Then, a VBM analysis (Ashburner & Friston, 2000; Mechelli, Price, Friston, & Ashburner, 2005) was performed by using SPM12 (Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London, UK; http://www.fil.ion.uncl.ac.uk/spm), running under Matlab R2014b. All structural images were first manually aligned along the anterior-posterior commissure (AC-PC) line and set the AC as the origin of the spatial coordinates. Images were then segmented into GM, WM and CSF and imported into a rigidly aligned space (Ashburner & Friston, 2000). For the next steps, only the GM segmented images were used. These images were used to create a study-specific template through a fast diffeomorphic image registration algorithm (DARTEL; Ashburner, 2007), available in SPM12, in order to maximize accuracy and sensitivity (Yassa & Stark, 2009). The last preprocessing step consisted in spatially normalizing into the MNI space the resulting warped GM images, through an affine spatial normalization and applying a modulation in order to ensure that the total amount of gray matter in each voxel was conserved after the registration (Ashburner & Friston, 2000; C. D. Good et al., 2001; Mechelli et al., 2005). Finally, the images were smoothed with an 8 mm full-width at half-maximum (FWHM) Gaussian kernel, commonly adopted in VBM studies (Scarpazza et al., 2013). The resulting smoothed, modulated, normalized data were submitted to the VBM analysis.

Univariate analysis: Voxel-Based Morphometry (VBM)

First, global measures of brain volume were calculated. For each subject we quantified the volume of Grey Matter (GM), White Matter (WM), Cerebrospinal Fluid (CSF) and estimated Total Intracranial Volume (eTIV). Then we explored the relation between these measures and the educational level, that is whether there were differences in brain volumes between different levels of education.

Second, a VBM analysis was performed (with SPM12 software) to focus on differences in regional GM volume. Starting from the results of the previous experiment, we selectively analyzed only voxels in the parietal lobe by adopting a brain-mask created with the Automated Anatomical Labeling (AAL) atlas implemented in Pickatlas software (http://fmri.wfubmc.edu/ software/PickAtlas). A general-linear regression model (GLM) was built in order to test the effect of education level on the amount of GM volume in a voxel-by-voxel fashion. As in the previous study, age and sex were used as covariates of no interest in a regression model (Buckner et al., 2005; Salat et al., 2004), to adjust for possible biases. In addition, we global scaled the data using the estimated Total Intracranial Volume (eTIV; for details on its calculation see Buckner et al., 2004; Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Marcus et al., 2007), in order to account for inter-individual differences in brain size that could affect the results. Statistically significant effects of education on regional GM volume were identified.

Multi-Voxel Pattern Analysis

MVPA was performed by adopting a "searchlight" approach (Kriegeskorte et al., 2006) in order to search across the whole brain for the most informative voxels about the classification between subjects with high vs. lo educational level, in an unbiased fashion. A spherical searchlight with a radius of four voxels around a central voxel was used. Then a Support Vector Machine classifier (SVM; Cortes & Vapnik, 1995) was trained and tested on the classification of interest. SVM algorithms have proved their efficacy for MVPA on neuroimaging data (Pereira et al., 2009). In particular they are adopted for clinical applications on structural MRI data (e.g., Liu et al., 2012).

As a first step, in order to be consistent with the previous study (see Chapter 3) we selected from the whole dataset (N= 109), only subjects with "low" educational levels (level 1 or 2) or with "high" educational level (level 4; no IXI subjects were assigned level 5 with OASIS coding system) were selected (N = 88). Then, participants with an educational level of 1 or 2 (N = 42) were assigned to the "low education" group (LED), while participants with a level of 4 (N = 46) were assigned to the "high education" group (HED). In order to balance the number of subjects belonging to the two groups, 4 subjects were randomly excluded from the HED group. The resulting dataset (N = 84) was half-splitted into two independent subsets, each one including 42 subjects: 21 from the LED group and 21 from the HED group). One subset was used as training-set and the other one was used as independent test-set. A Leave-One-pair-Out scheme was adopted, where the algorithm was trained on N-2 subjects from the training set (40 = 20 LED + 20 HED) and tested on 2 subjects (1 LED + 1 HED) from the the test set. This procedure was repeated 21 times, until each pair of subjects in the test-set was used once to evaluate the classifier performance (see Figure 4.1).

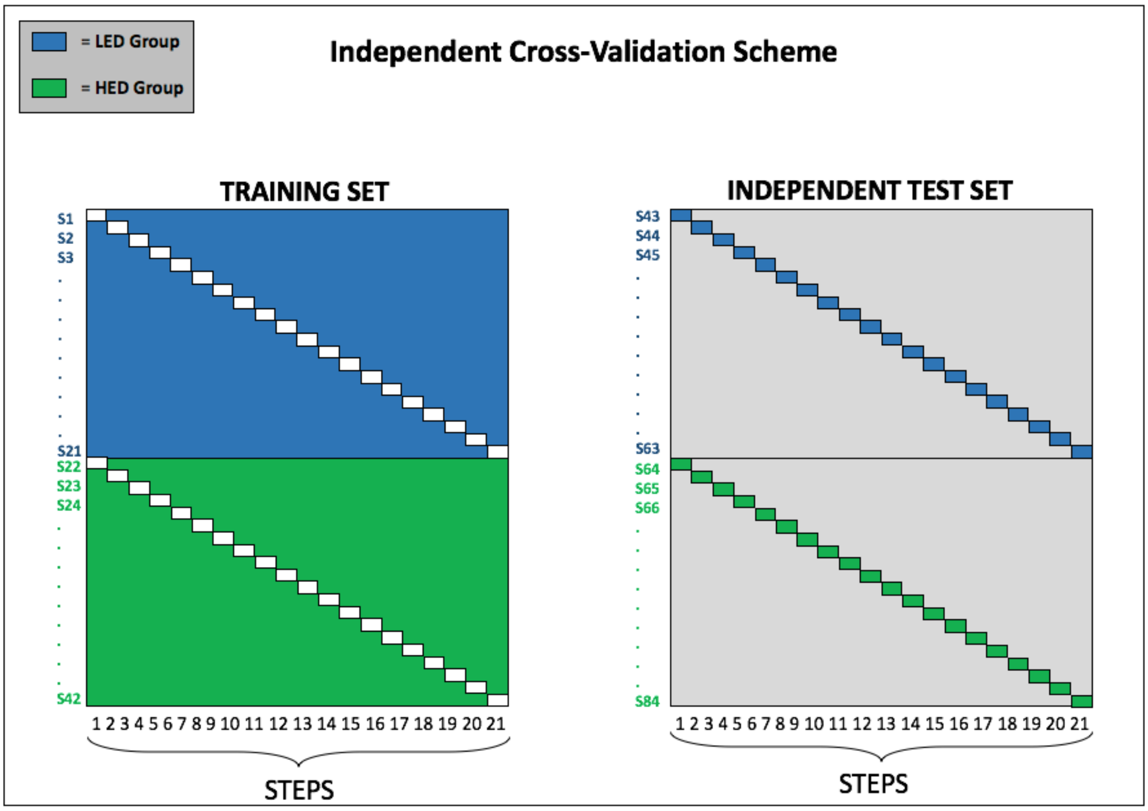

Fig. 4.1. Cross-Validation scheme adopted in this MVPA study. Training set and test set were maintained separate, thus this design is called cross-decoding design. For example, in the 1st step, subjects from S2 to S21 (LED) and from S23 to S42 (HED) were used to train the classifier, while subjects S43 and S64 were used to test the classifier performance. BLUE indicates LED group; GREEN indicates HED group; WHITE indicates unused subjects from the training set; GREY indicates unused subject from the independent test set.

Then, to statistically test the significance of the voxel-by-voxel accuracies reached by the searchlight classifier, we used a binomial test (Pereira et al., 2009) as it has been demonstrated that distributions taken from CV-independent schemes (as the one we adopted) match the binomial distribution when a Leave-One-Out scheme is adopted (Noirhomme et al., 2014). The decoding analyses were performed using The Decoding Toolbox (TDT; Hebart, Görgen, Haynes, & Dubois, 2015; https://sites.google.com/site/tdtdecodingtoolbox/) and LIBSVM (Chang & Lin, 2011; http://www.csie. ntu.edu.tw/~cjlin/libsvm/), running under Matlab R2014b. As input for the multivariate analyses, 8-FWHM smoothed, modulated, normalized images were used.

RESULTS

Global-volume measures

Figure 4.2 shows no significant differences between educational levels in the investigated volumetric measures (GM, WM, CSF, eTIV). This visual descriptive analysis indicates that global measures of brain volume are not correlated with educational level. The subsequent VBM analysis will take into account more subtle differences in regional GM volume.

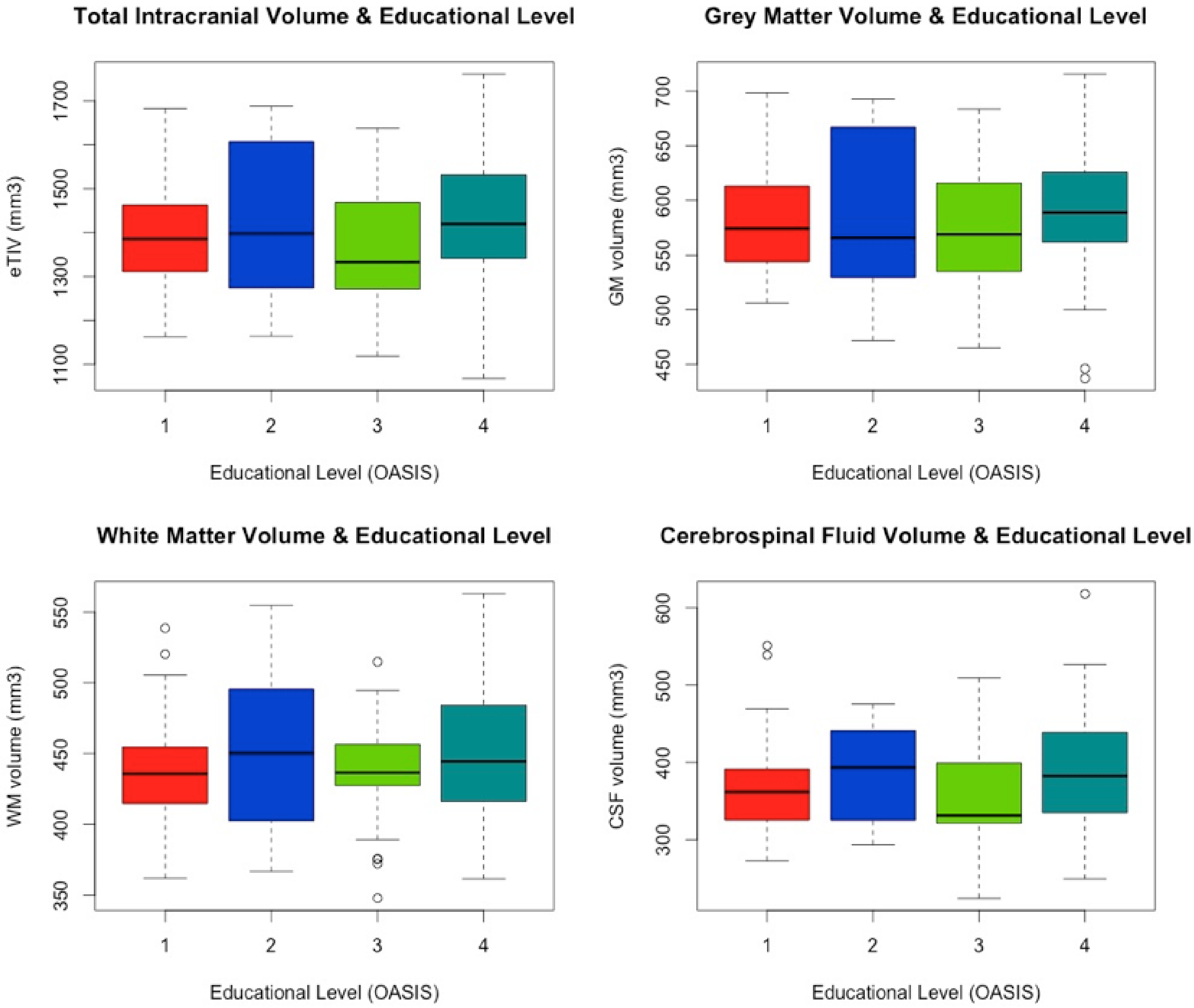

Fig. 4.2. Relation between different measures of brain volume and levels of education. Top left: Total Intracranial Volume (eTIV); top right: Grey Matter (GM); bottom left: White Matter (WM); bottom right: Cerebrospinal Fluid (CSF). All measures are reported in mm^3.

Voxel-Based Morphometry

The regression model built for the VBM analysis showed only one significant cluster in the left supramarginal gyrus (t[105] = 3.87, p<.001) when correcting for the effect of age, sex and eTIV. However no voxels survived the Family-Wise multiple comparisons correction at p<.05. See Table 4.3 and Figure 4.3 for the uncorrected results at p<.001 and minimum cluster-size of 100 voxels.

Peak	**Lat.**	**Cluster size**	**T**	**MNI peak coordinates** (mm)		
				x	**y**	**z**
Supramarginal Gyrus	L	123	3.87	-51	-45	34

Tab. 4.3. Peak-voxel at p<.001 uncorrected (minimum cluster-size = 100). Lat. = lateralization.

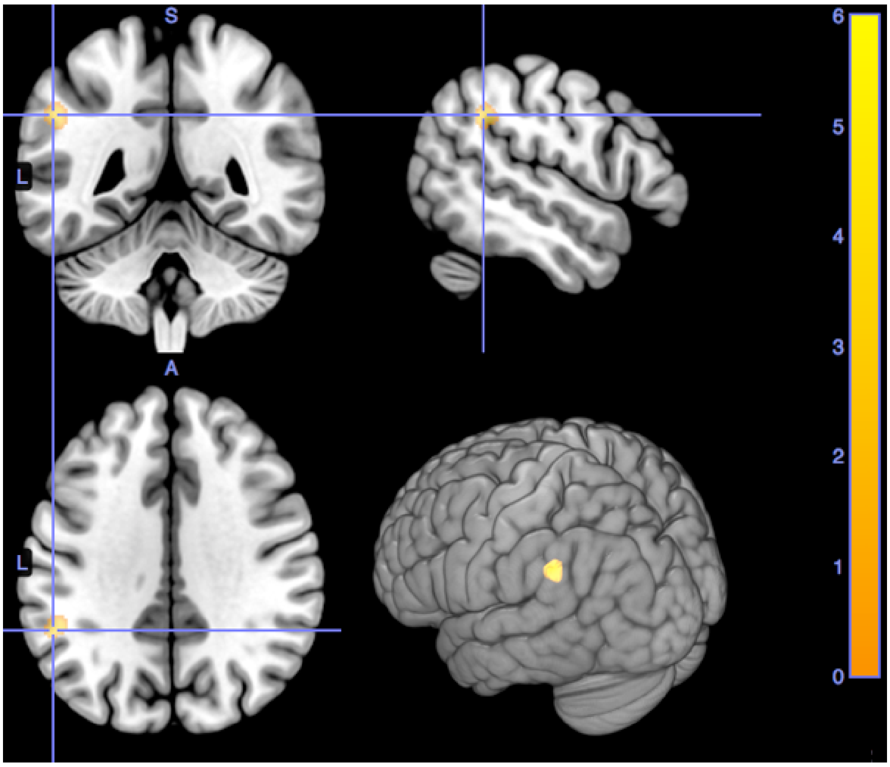

Fig. 4.2. Brain areas showing significant effects of educational level on GM volume, adjusting for age, sex and eTIV. Statistical T maps are thresholded at p<.001 uncorr. (minimum cluster-size = 100). S = superior; L = left; A = Anterior.

Multivariate analysis

Multi-Voxel Pattern analysis was performed investigating the accuracy of a binary SVM classifier in discriminating between LED and HED on the basis of the volume of GM, with a searchlight approach.

The statistical significance of the accuracy values was assessed trough the binomial test and corrected for multiple comparisons at cluster-level using a Monte-Carlo simulation technique implemented in the AlphaSim program (with a cluster-forming threshold p=.0005). Results highlighted a cluster in the left supramarginal gyrus showing accuracies significantly above chance (see Figure 4.3). The voxels in this cluster reached accuracies ranging from 76% to 85% (mean: 77.2%) in classifying brain images as belonging to the LED or the HED group. The most informative voxels highlighted in this multivariate analysis were basically the same emerged from the univariate analysis as the most sensitive to differences of educational level.

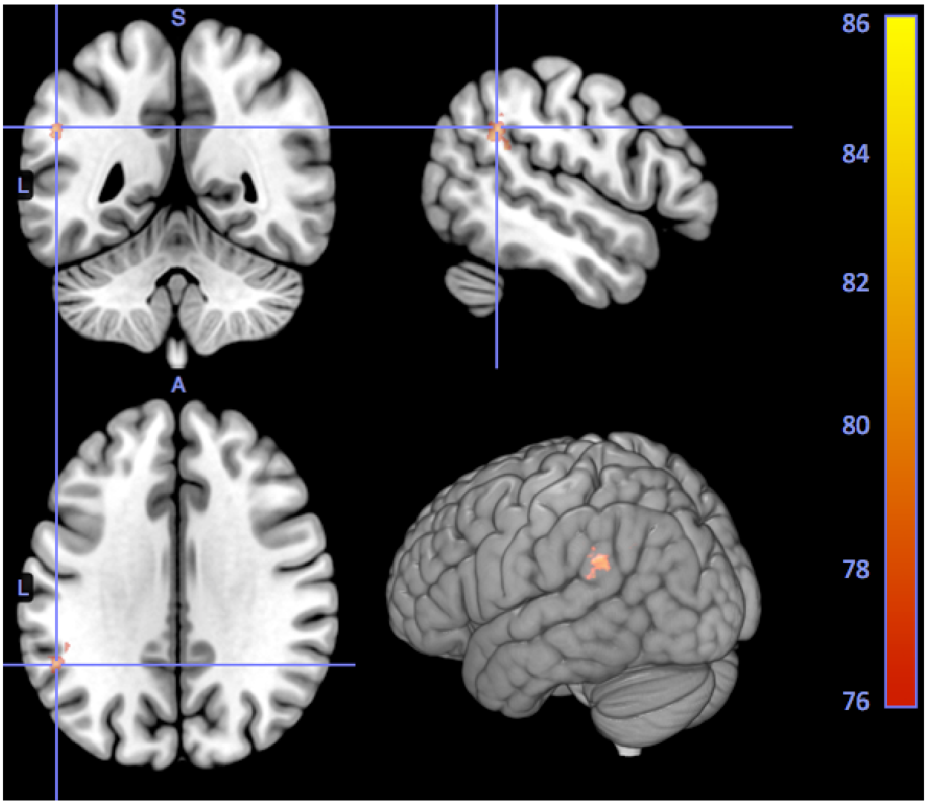

Fig. 4.3 Most informative voxels in the discriminating between LED and HED, highlighted by the searchlight-based MVPA analysis. S = superior; L = lateral; A = anterior.

DISCUSSION

Starting from the findings of a previous work (see Chapter 3), in the present study we investigated in a deeper way the neural basis of Cognitive Reserve (CR) using educational level as proxy of CR. Previous findings highlighted the role of the pattern of GM volume in the parietal lobe in encoding crucial information about the discrimination between high and low educated subjects. Therefore, here we focused on the parietal lobe, performing a searchlight MVPA analysis in order to shed light on which sub-regions are the most informative for the classification of subjects based on educational level, and thus the most sensitive areas to the effects of education on brain structure. We analyzed a sample of healthy elders extracted from a freely available database of MRI images (IXI; brain-development.org). For each subjects in the database a value of educational level was provided. However, in order to be consistent with the previous study, we transformed those values according with the educational level coding system used in the previous study (based on OASIS database coding system).

Uncorrected results of the preliminary VBM analysis showed a cluster in left supramarginal gyrus as significantly related to education level, however no voxels survived multiple comparisons correction. Interestingly, the MVPA analysis showed a significant cluster in the same area as the most informative for the discrimination between LED and HED subjects.

These findings show that two different analysis approaches converged to the same result, suggesting that CR produces specific regional GM changes, rather than a general increase in brain volume.

Our results are consistent with the findings of a recent study (Wook Yoo et al., 2015) on healthy elders, aimed at investigating cognitive reserve from a network perspective, analyzing the relation between the WM network connectivity and education levels. Authors found a role of education in reinforcing network reliability in a sub-network centered in the left supramarginal gyrus. Another study (Serra et al., 2011) demonstrated that AD patients with low educational level had less GM in the bilateral supramarginal gyrus, in the right precuneus and frontal opercular cortex, if compared to patients with high educational level. Ciaramelli and co-workers (2008) suggested that the inferior parietal cortex, including the supramarginal gyrus, is part of a “bottom-up” attentional subsystem that mediates the automatic allocation of attention to task-relevant information. This supports the main idea about the way in which cognitive reserve acts in practice, that is through a more efficient use of functional networks (Yaakov Stern, 2009). Moreover, it has been widely demonstrated that the left supramarginal gyrus has a crucial role in reading (e.g., Sliwinska, Khadilkar, Campbell-Ratcliffe, Quevenco, & Devlin, 2012), in particular in phonological processing (Hartwigsen et al., 2010; Oberhuber et al., 2016). Reading is considered one of the leisure time activities that contributes to cognitive reserve (e.g., Scarmeas & Stern, 2003), and it is taken into account in several studies and considered in composite measures of CR, such as the CRIq (Nucci et al., 2012). Moreover, reading is one of the most stimulated functions in formal education paths.

Taken together, these findings suggest a crucial role of the supramarginal gyrus for the neural basis of CR. Therefore, it seems plausible that this area resulted as the most informative (within the parietal lobe) in discriminating between low and high educated subjects.

The non-significance of univariate VBM results could find an explanation in the lack of sensitivity of the traditional neuroimaging analyses techniques for small effect-sizes. Indeed, it is well known that there are differences between standard voxel-wise analysis and MVPA in sensitivity (Coutanche, 2013; Davis & Poldrack, 2013; Jimura & Poldrack, 2012). These differences are due to the fact that standard analyses investigate the relation between a single-voxel measure (e.g., blood-flow, GM volume) and variables, and thus can show poor sensitivity when mapping the neural basis of experimental variables having a multimodal effect (e.g., a distributed pattern of activity).

Thus, in many cases, the use of a multivariate (e.g., searchlight) technique could be preferable due to its sensitivity in investigating the brain correlates of a condition (e.g., a neurological diagnosis) from a pattern point-of-view.

SECTION B

DETECTING AUTOBIOGRAPHICAL MEMORIES FROM BRAIN ACTIVITY: A MVPA APPLICATION

Introduction

Autobiographical memory (AM) is what people usually mean by using the term “memory”, that is the ability to remember past events from one’s own life. AM is defined as a form of memory that is focused on recalling experienced events integrating them in a unified perspective, in order to create a personal history (Fivush, 2008). The ability to represent ourselves as individuals experiencing events, and to link past, present and future episodic representations into a meaningful framework defining our own lives, is the crucial feature of autobiographical memory (Fivush, 2008). For this reason some authors refer to this kind of memory as *“biography of self”* (e.g., Conway, Singer, & Tagini, 2004). This unique feature makes autobiographical memory a purely human skill, as it is not possible to build an autobiography without being part of a social group (Donald, 2001). Although AM can be considered as a type of episodic memory for information related to the self in both a retrospective and a prospective sense (Schroots, van Dijkum, & Assink, 2004), in this section only the former will be discussed.

Memory systems: a brief overview

In the last twenty years, given the improvement in the ability to study cognitive functions, and in particular memory, there is growing consensus in considering memory not as a unified function but rather as a set of dynamic, integrated systems (Howard Eichenbaum & Cohen, 2004; Fivush, 2008; D. L. . Schacter, Wagner, & Buckner, 2000; Squire, 2004). Memory can be broadly defined as consisting of two main systems: declarative and nondeclarative memory.

Nondeclarative memory refers to the recalling of information with little to no conscious awareness, including procedural knowledge, that is knowledge of how to do things that are well-practiced, such as using a bike, as well as the phenomenon of conditioning and priming (D. L. . Schacter et al., 2000; Squire, 2004). By contrast *declarative memory* is a kind of explicit memory that is available to consciousness, and from a broad point of view, the kind of memory we refer to when speaking about "memory" in general (Fivush, 2008). Indeed, this memory system includes representations of real-life past experiences (e.g., *"I remember when I was 20 and went to Rome for holidays"*). The declarative memory system is not just a unitary warehouse, but is composed by two subsystems: semantic memory and episodic memory (Tulving, 1972). *Semantic* memory refers to general knowledge about the world without specific coordinates about when and where it has been acquired. For example, we know that cows are mammals or that the pope is the leader of the Catholic Church, but we did not learn these information in a specific episode. This is the main difference with *episodic memory*, which refers to specific memories having a link with spatiotemporal coordinates. For example, remembering the first time we have driven a car or the day of our master thesis defense.

Given that episodic memory system includes all memories of specific past events (Fivush, 2008), autobiographical memory can be considered as an episodic memory subsystem, or even as the episodic memory itself.

From Episodic Memory to Autobiographical Memory

Some authors (e.g. Fivush, 2008) state that AM should be considered as a different memory system in respect of episodic memory, arguing that this distinction allows for a more exhaustive understanding not only of the development of these two systems in humans, but also across species. Essentially, we can say that AM involves memory for one's personal past and includes memories that are characterized by both episodic and semantic features (Glisky, 2007).

According to Tulving (Tulving, 2002), two components of episodic memory can be identified, one containing information about what, where, and when of an experience, and the second involving the consciousness of self having experienced the event in the past. This last component, in Tulving's model, involves the so-called "*mental time travel*" (Fivush, 2008; Tulving, 2002). According to some authors (e.g., Fivush, 2008), these components can be considered as separable, as the first appears to be available both across species and human development, while the second, whose main feature is the so-called *"autonoetic awareness*", is typically human.

However, episodic and autobiographical memory are undoubtedly strongly interconnected. Indeed, episodic memory refers to the recollection of a specific event, as well as autobiographical memory, but the latter refers to the retrieval of an episode also including additional information such as the memory of the self experiencing the event (what previously called "*autonoetic awareness")*. Moreover, autobiographical memory has a role in linking independent events together in a coherent personal history (Habermas & Bluck, 2000).

Taken together, these considerations depict a role of autobiographical memory as a system sharing many features with episodic memory. Therefore, in particular for the studies presented in the next chapters, we can speak about "autobiographical episodic memory".

Neuroimaging of Autobiographical Memory

Many methods have been adopted for eliciting AMs in the scanning environment. The main difference between these techniques lies on the degree of control that the experimenter has on several properties of the retrieved memory. That is, some methods cannot be used when the remoteness (or other properties) of the recollected memories is a variable under investigation. In the following, a non-exhaustive list of the most common approaches with a brief description for each one is presented:

- *Generic cues* (Crovitz & Sciffman, 1974): in this method participants are asked to retrieve an AM associated with a provided cue. Several kinds of cues have been used: nouns (Conway et al., 1999), emotional words (Markowitsch, Vandekerckhove, Lanfermann, & Russ, 2003), words referred to a specific semantic field (Muscatell, Addis, & Kensinger, 2009), odors (Masaoka, Sugiyama, Katayama, Kashiwagi, & Homma, 2012), pictures (Burianova & Grady, 2007), or musical clips (Ford, Addis, & Giovanello, 2011). Given that AMs elicited by cues are not necessarily emotional or significant for the participant, they can imply a more time for the retrieval, so that this approach is useful for example when studying memory construction. Moreover, this technique leads to more accurate subjective ratings (Addis, Barense, & Duarte, 2015).

- *pre-scan interview*: this method allows to elicit AMs referred to specific events during the scanner session. The investigated memories are collected before the scanning session (e.g., Addis, Moscovitch, Crawley, & McAndrews, 2004; Maguire, 2001). This technique has the advantage that it is possible to control aspects of the memories retrieved in the scanner (e.g., remoteness, emotional tone or vividness), and that investigated memories can be highly specific and accessible during the fMRI task (Addis et al., 2015).

- *Independent sources*: this is a technique for eliciting AMs by means of cues generated by external sources, such as family or friends (e.g., Rabin & Rosenbaum, 2012). The main advantages of this method is that elicited memories are unrehearsed, and that they can be constrained by collecting more information from the sources (Addis et al., 2015).

- *Prospective method*: in this method participants have to note a sequence of events happened in their lives, that will be used as cues in the scanner (e.g., Cabeza et al., 2004). Here, the main advantage is that it allows to strongly control the encoding phase and test of the retrieval accuracy.

Using these methods, several studies have focused on autobiographical memory's functional brain correlates. Essentially, these studies have highlighted medial and left-lateralized activations when retrieving autobiographical episodic memories (E. A. Maguire, 2001a). Areas constituting the "core" network of AM have been highlighted and seem to include medial and ventrolateral prefrontal cortices, medial and lateral cortices, temporoparietal junction, retrosplenial/posterior cingulate cortex, and the cerebellum (e.g., Addis, McIntosh, Moscovitch, Crawley, & McAndrews, 2004; Cabeza & Nyberg, 2000; for a meta-analysis see Svoboda, McKinnon, & Levine, 2006). These methods have been used to study both recollection and recognition memory. In particular methods based on the use of cues (pictures, words or, as we did, sentences) can be useful to investigate recognition without any explicit instruction.

Aims of this section

As abovementioned, several studies have investigated the neural correlates of AM retrieval, mainly through the use of univariate techniques. Some studies (e.g., Addis, McIntosh, Moscovitch, Crawley, & McAndrews, 2004) adopted multivariate techniques, also investigating the possibility of classifying memories through the use of multivariate analysis methods (e.g., Rissman, Chow, Reggente, & Wagner, 2016; Rissman, Greely, & Wagner, 2010). Moreover few studies have adopted real-life, instead of laboratory-based, stimuli (e.g., Rissman et al., 2016).

In the next chapters (5 and 6), two functional MRI studies using MVPA techniques to decode true subject-specific autobiographical memories will be presented. In the first study, we tested whether it is possible to discriminate between the pattern of brain activity associated with pictures of seen (i.e., previously visited) rooms from that of rooms where participants had never been before. In the second study sentences describing a negative emotional autobiographical memory were used, and the performance of a classifier distinguishing between sentences describing real memories and non-memories was tested.

Chapter 5

“HAVE YOU BEEN THERE BEFORE?”
DECODING ROOMS RECOGNITION IN THE HUMAN BRAIN

INTRODUCTION

When watching a picture of a scene (e.g., a street), an architect can be attracted by the style of the buildings, a photographer can notice the perspective, an artist can be captured by the color balance, and so on. Independently from subjectively relevant aspects which can capture attention, however, all the observers in this example certainly felt whether the picture’s content was familiar, that is, if they experienced before that scene, they recognized it as present in their memory.

Recognition memory represents a critical aspect of our ability to remember (M. W. Brown & Aggleton, 2001) and it includes the ability of judging whether a stimulus had been experienced before, as well as of identifying it (Mandler, 1980). Although different models of recognition memory have been proposed, several authors agree with the point that this process does not necessarily imply the recollection of the stimulus-related information. Therefore, recognizing a scene is independent from recollecting what happened in that scene (e.g., being presented with the image of the street where we live does not imply recalling the episode in which, while walking home one week ago, our smartphone fell down from the pocket). Moreover, the process of recognizing a scene (e.g., a place) can be realized at two different levels (Epstein & Higgins, 2007; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Tversky & Hemenway, 1983): the scene can be identified as belonging to a category (e.g., “a café”) or as referring to a specific place (e.g., “the café where I had breakfast this morning”). The first level refers to semantic categorization, while the second level refers to recognition memory. In other words, in this second level, the place is recognized because it has been previously experienced and thus it is familiar to us. In literature, recollection is described as an effortful and time consuming process, while recognition (or familiarity, considered as the first step of recognition) is an automatic and rapid process. Human beings can recognize natural scenes in less than 150 ms (Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001), also with near-absent attention (Fei-Fei, VanRullen, Koch, & Perona, 2005; Li, VanRullen, Koch, & Perona, 2002). According to some authors, given its speed, this process can be considered a product of the evolution as it allows us to rapidly react in response to novel stimuli (M. W. Brown & Aggleton, 2001).

A great amount of studies have focused on the recognition of complex visual scenes, and on investigating its brain correlates. The parahippocampal cortex, in particular the Parahippocampal Place Area (PPA) and the retrosplenial cortex (RSC) seem to be the most strongly active areas while processing pictures of places (Epstein & Higgins, 2007; Epstein & Kanwisher, 1998; O'Craven & Kanwisher, 2000; Walther, Caddigan, Fei-Fei, & Beck, 2009).

In a work by Cabeza and collaborators (2004), neural correlates of the recognition of pictures of scenes previously encountered in a laboratory setting or in real-world settings were investigated. In particular, the authors used a novel "photo paradigm" in which participants had to judge whether photos were novel, previously experienced in the lab or acquired by themselves in everyday life. The recognition of self-acquired photos (i.e., real-world autobiographical memories), if compared to controlled laboratory condition, highlighted the involvement of medial prefrontal cortex, occipital and parahippocampal regions, as well as hippocampus. Most likely, real-world scenes are encoded in brain-activity patterns, given their complexity level. So that, the most suitable analysis technique for detecting the activity pattern associated with the recognition of previously experienced scenes is the multivariate approach, called Multi-Voxel Pattern Analysis (MVPA). Indeed, this approach has been adopted to deal with the decoding of distributed brain activity processes (e.g., Cox & Savoy, 2003; Haxby et al., 2001), including representations that can potentially exist at a smaller spatial scale than the size of functional voxels (Haynes & Rees, 2006; Kamitani & Tong, 2005).

To our knowledge, a small number of previous studies have tested whether a multivariate approach would be able to accurately classify a brain activation as indicating the response to a previously encountered stimulus or not (Rissman et al., 2016, 2010; Uncapher, Boyd-Meredith, Chow, Rissman, & Wagner, 2015), and none of them dealt with the recognition of a real-world experienced room. Importantly, the majority of the studies focusing on the use of MVPA for detecting memories has focused on memories for information encoded in a laboratory setting rather than real-world derived (Rissman et al., 2016). Determining whether a subject is looking at a picture of a known or unknown scene (e.g., a specific room) can have important implications for forensic uses. Indeed, several authors have discussed the possibility of using MVPA-based memory detection techniques in forensic settings (Bles & Haynes, 2008; Meegan, 2008; Rissman et al., 2016; Schacter & Loftus, 2013; Shen & Jones, 2011).

The aim of the present experiment is to investigate the brain activity pattern evoked by the observation of stimuli depicting previously visited rooms and the possibility to discriminate it from that associated to the processing of unknown-rooms-stimuli. To reach this goal, firstly the (univariate) differences in brain activity between the processing of previously seen vs. unseen rooms were investigated. Subsequently, the performance of a multivariate classifier was evaluated in revealing which brain areas contain more useful information for the discrimination between the brain activity patterns associated to seen and unseen rooms.

MATERIALS AND METHODS

Participants

Thirty healthy participants between the ages of 19 and 31 (mean=24; SD=3.87) took part in the study. All participants had normal or corrected-to-normal vision and were free from any neurological or psychiatric diagnosis. To ensure that anyone of the participants had visited before the rooms used for the experiment, they were enrolled among people not working or living in the area around the Campus Nord of Humbolt Universität zu Berlin, where the rooms used for the experiment were selected. Six participants were excluded from the final sample because of motion artifacts (1), errors in telling which rooms were visited or unknown after the fMRI session (4), sleeping during the task (1). A total of 24 participants was included in the final analysis (11 females). All participants gave informed consent and were remunerated with 22€ for their attendance. The study was approved by the local ethics committee, Humboldt–Universität zu Berlin, Germany.

Stimuli

Eight rooms were selected within the Campus Nord of Humbolt Universität zu Berlin: 1) the hall of the Institute of Vegetative Anatomy at Charité Hospital; 2) the children´s playground at Humboldt Graduate School; 3) the lecture hall at Bernstein Center for Computational Neuroscience Berlin; 4) an office at Berlin Center of Advanced Neuroimaging; 5) the staircase at Dreispitzpassagen Berlin; 6) a room at the basement at Humboldt Graduate School; 7) the conference room at Humboldt Graduate School; 8) the recreation room at Bernstein Center for Computational Neuroscience Berlin. A total of 60 pictures and a videotape were acquired for each room with a Canon 5D Mark II digital camera. Each picture was taken from a different point of view in order to have different representation of rooms' characteristics. The videotape was recorded by positioning the camera approximately 1.6 m from the floor and walking through each room for one minute. All the pictures and video frames were resized in order to match dimensions (900x600) and subjected to luminance normalization through the SHINE Toolbox (Willenbockel et al., 2010), to control for low-level attributes' effect. The videotapes were finally split in sequences with a duration of five-seconds. To avoid auditory information to influence participants performance during the fMRI task, videos were recorded without sound. Taken together, 480 pictures and 96 videos were used as stimuli in the experiment.

Pre-scanner task

A randomized sequence of four out of eight rooms was assigned to each participant. Participants were individually accompanied by an examiner to visit the rooms. Once inside, participants were asked to perform a task (finding five hidden Lego bricks) in order to implicitly ask participants to explore the room, and the time needed to complete this task was recorded. Thus, four of the eight rooms were known ("seen") before the fMRI session, while the remaining for were unknown ("unseen").

In-scanner task

The day after visiting the rooms participants underwent an fMRI session in which they were asked to watch a screen presenting pictures and videos showing the rooms they visited the day before, as well as rooms they had never been before. They were not asked to give any behavioral response while being scanned. The fMRI session was divided into six functional runs. Each run consisted of 32 trials during which a sequence of five pictures (750 ms presentation and 250 ms interval) or one five-seconds video was presented. A eight seconds ITI was inserted, so that a trial was presented each 13 s. The global run duration was 422 s (32 trials × 13 s + 3 volumes before the first trial started). The stimuli were presented in a randomized order across all runs, with the following restrictions: in each run, two videos and two pictures sequence were presented for each room. A maximum of three videos of one condition was presented in a row. Each item (picture or video) appeared once within each run. Stimuli were presented using the Psychtoolbox (http://psychtoolbox.org) running under MATLAB R2014b (The MathWorks, Inc.) and projected onto a screen (1024 × 768 pixel, 60 Hz) from the head-end of the scanner.

Post-scanner task: explicit rooms recognition

After the fMRI session, participants were asked to fill a brief questionnaire in which they had to explicitly confirm which rooms they had seen before and which they watched for the first time during the scanner session. So that, for each room a sequence of pictures was presented and the participants had to answer (*yes/no)* to the question *"have you been to this room already?".* Moreover they had to judge how much confident they were on their answer in a 7-points scale (from 1= "very sure" to 7="very unsure").

fMRI Data Acquisition

Gradient-echo EPI functional MRI volumes were acquired with a Siemens TRIO 3 T scanner with standard head coil (33 slices, TR = 2000 ms, echo time TE = 30 ms, resolution 3 × 3 × 3 mm^3 with 0.75 mm gap, flip angle = 78°, Field of View [FoV] = 192 mm). In each run, 216 images were acquired for each participant. For every subject, six runs of functional MRI were acquired. We also acquired structural MRI data (T1-weighted MPRAGE: 192 axial slices, TR = 1900 ms, TE =2.52 ms, flip angle = 9°, FoV = 256 mm).

fMRI Data Analysis

SPM12 (v. 6685; http://www.fil.ion.ucl.ac.uk/spm/) running under Matlab R2014b (MathWorks, Inc.) was used for fMRI data preprocessing and univariate analysis. The first three volumes of each functional timeserie were discarded to avoid magnetic saturation effects. The images were slice-time corrected with reference to the first recorded slice, motion corrected, coregistered to the anatomical image, and then spatially smoothed with a Gaussian kernel of 6 mm FWHM. Data were high-pass filtered with a cutoff period of 128 s. The images were also segmented in order to obtain normalization parameters for normalizing the accuracy maps resulting from the multivariate analysis. The onsets of the trials (pictures or video) were used to model the fMRI data. A general linear model (Friston et al., 1995) was applied to the data of each run and four event-based HRF-convolved regressors were modeled: pictures of seen rooms (Sp), videos of seen rooms (Sv), pictures of unseen rooms (Up), videos of unseen rooms (Uv). Movement parameters derived from image realignment were included as additional regressors of no interest. At a subject-level a separate contrast was made for each of the four regressors. Then, the obtained contrasts were spatially normalized to the MNI space and entered into a second-level random effects

ANOVA with the stimulus type (Sp, Sv, Up, Uv) as within-subject factor. Activations for the contrasts between seen and unseen stimuli have been investigated ([Sv+Sp]>[Uv+Up]; [Uv+Up]>[Sv+Sp]). Finally, we ran a new first-level analysis on unsmoothed data, modeling one regressor per room in order to prepare data for the subsequent MVPA analysis.

Multi-Voxel Pattern Analysis

Multi-Voxel Pattern Analysis (MVPA) was performed to investigate whether the brain activity pattern elicited by watching at stimuli representing seen rooms can be discriminated from that of unseen rooms.

The decoding analysis between “Seen” vs. “Unseen” rooms was performed considering both pictures and videos, pictures only and videos only. Moreover, we performed a cross-decoding between pictures and videos in which we trained the classifier only on pictures and tested the performance on videos and vice versa. So that, for the “Seen” vs. “Unseen” decoding, four different analyses were realized.

All the decoding analyses followed the same steps. First, a GLM was estimated on unsmoothed data, with eight regressors (rooms) modeled for each run. Second, the 48 estimated GLM parameters (8 regressors × 6 runs) were included in the decoding analysis. A “searchlight” approach (Kriegeskorte et al., 2006) was adopted to search across the whole brain for the voxels that contain more information in an unbiased fashion. A spherical searchlight with a radius of three voxels around a central voxel was adopted and each voxel in the brain served once as the center of the searchlight. The third step was the training and test of a Support Vector Machine classifier (SVM; Cortes & Vapnik, 1995) as it has been demonstrated to be a valid approach for pattern analysis, especially on functional neuroimaging data (Pereira et al., 2009; Schmah et al., 2010). As described in Chapter 2, SVM is a binary classifier that finds the best classifying hyperplane, that maximizes the distance with the closer element of each class. Given a set of features, the SVM classifier is trained on a set of data and it produces a model able to predict with a certain accuracy the class label of the elements in a different dataset (test set). In order to avoid overfitting, the approach known as Leave-One-Out Cross-Validation (LOOCV; see Chapter 2) was used, in which a subset of data is not included in the training set and it is used as test set. For the “Seen” vs. “Unseen” decoding the classifier was trained on GLM estimates from six rooms and a pair of rooms (one seen and one unseen) was left out and used to test the classifier performance. This procedure was repeated with each pair of rooms acting as the test data set once, while the other rooms were used as training data sets. Importantly, although the "Seen" vs. "Unseen" decoding has been realized with room-specific estimated GLM parameters, the leave-one-pair-of-rooms-out cross-validation design allow to test the accuracy of the classifier in discriminating between seen and unseen rooms without the influence of low-level visual differences between rooms. In the rooms decoding GLM estimates from five runs were used as training set, while one run was iteratively left out. The training and testing of the support vector classifier was performed using The Decoding Toolbox (TDT; Hebart, Görgen, Haynes, & Dubois, 2015; https://sites.google.com/site/tdtdecodingtoolbox/) and LIBSVM (Chang & Lin, 2011; http://www.csie. ntu.edu.tw/~cjlin/libsvm/). The result of each decoding analysis was a 3D accuracy map showing a value of accuracy above chance for each voxel in the brain. Each accuracy map was then normalized to a standard stereotaxic space (Montreal Neurological Institute EPI template), and smoothed with a Gaussian kernel of 6 mm FWHM. Finally, a t-test was conducted to statistically test the decoding accuracies for each position in the brain. Decoding results were FWE-corrected for multiple comparison at cluster-level ($p<.05$; cluster-forming threshold = .001).

RESULTS

Functional MRI images of participants were acquired while they were passively viewing pictures or videos of eight rooms, four of which had been previously visited. A second-level within-subject ANOVA was conducted and the differences between brain activity in response to stimuli showing seen and unseen rooms were explored in a univariate analysis. Furthermore, we run a multi-voxel pattern analysis evaluating the performance of a searchlight SVM algorithm in classifying fMRI within-subject activation patterns for seen and unseen stimuli.

Behavioral data analysis

The post-scanning questionnaire revealed that there was no significant difference in the confidence rating between seen and unseen rooms. In other words, participants judged all the eight rooms as previously visited or not with the same degree of confidence. Overall, in 96% of cases (184 out of 192 total ratings) participants chose a score of 1 on the Likert scale, meaning that they were very sure about their answer (about having visited the room or not).

Univariate analysis

We computed an ANOVA with *post-hoc* T-contrasts in order to investigate the differences in BOLD signal between seen and unseen stimuli. All results are reported in Table 5.1 and were FWE corrected at the cluster-level at p<.05 (threshold p<.001). For the contrast between seen and unseen, results showed significant activations (all clusters had p<.001) in Precuneus and Parahippocampal gyrus bilaterally, left cuneus, right angular gyrus and right Superior Parietal Lobule (see Figure 5.1). The reverse contrast highlighted the involvement of frontal and parietal regions. In particular, significant activations were found in left Postcentral gyrys (p<.001), left Medial Frontal gyrus (p<.001), right Frontal Superior gyrus (p=.003).

Contrast	**Region**	**Lat**.	**Cluster size**	**T**	**MNI peak coordinates** (mm)		
					x	**y**	**z**
Seen > Unseen							
	Cuneus	L	216	7.62	-15	-58	17
	Precuneus	R	267	7.28	18	-58	23
	Parahippocampal gyrus	R	94	5.75	27	-40	-13
	Parahippocampal gyrus	L	89	4.92	-21	-43	-10
	Superior Parietal Lobule	R	62	4.44	15	-73	53
	Angolar gyrus	R	22	4.25	39	-79	32
	Precuneus	L	41	4.23	-9	-70	53
Unseen > Seen							
	Postcentral gyrus	L	47	4.55	-57	-11	17
	Medial Frontal gyrus	L	68	4.21	-30	41	17
	Frontal Superior gyrus	R	24	4.25	24	47	44

Table 5.1 Activations for the contrasts Seen > Unseen and Unseen > Seen considering videos and pictures. Abbreviations: Lat., Lateralization; L, left; R, right.

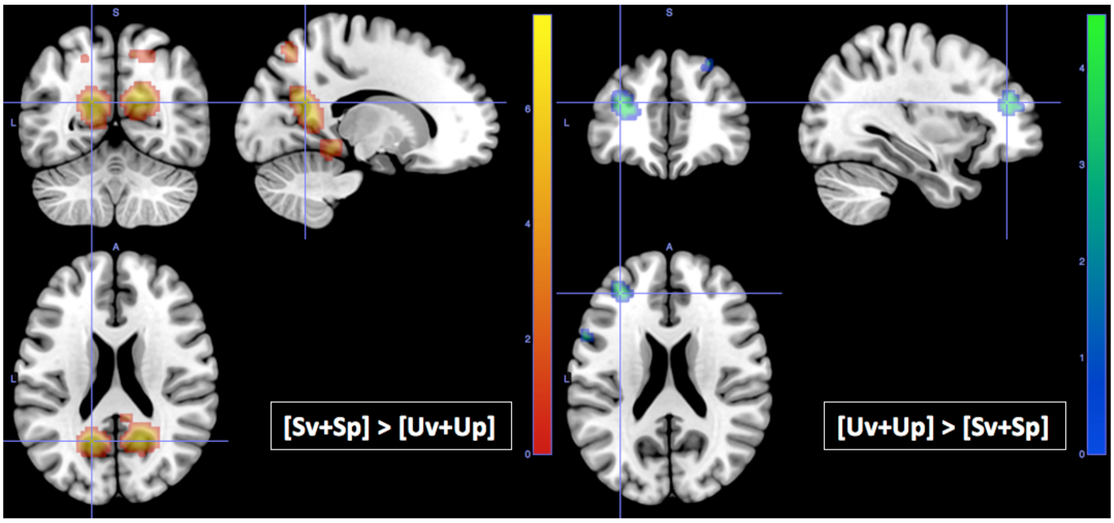

Fig. 5.1 Univariate results for contrast Seen > Unseen (left) and Unseen > Seen (right). Sv = Seen videos; Sp = Seen pictures; Uv = Unseen videos; UP = Unseen pictures; A = anterior; L = left; S = superior.

Multivariate analysis

A series of MVPA decoding analyses were performed in order to identify the areas which were more informative for discriminating between seen and unseen rooms with an accuracy level significantly above chance (FWE cluster-level corrected at p<.05; cluster-forming threshold = 0.001). The decoding performed separately for each participants showed an average maximum accuracy level of 82.2%, ranging from 75.5% to 94.3%. While the maximum accuracy at group level was around 60%. For the group analysis, a one-sample t-test was performed. Searchlight MVPA identified significant clusters in the parietal lobe bilaterally (Figure 5.2). Peak voxel was located in the right Precuneus (MNI: 6, −76, 53; cluster p<.001).

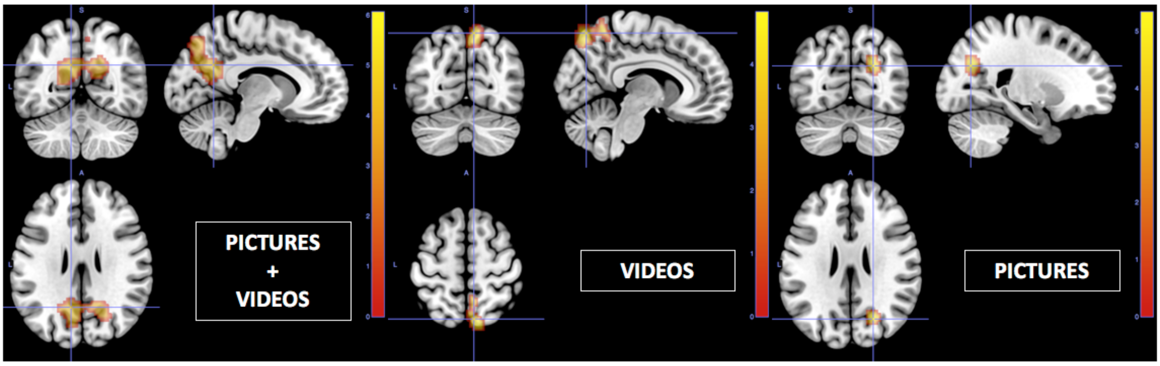

Fig. 5.2 Group-averaged maps of classification accuracy significantly above chance level for seen vs. unseen rooms. Decoding results obtained using both videos and pictures (left), videos only (center) and pictures only (right) are shown. S = superior; L = left; A = anterior.

No strong differences emerged when considering for the decoding pictures- or videos-based activation separately. In both cases right Precuneus (for pictures, MNI: 21, -67, 29; cluster p=.004; for videos, MNI: 12, -73, 59; cluster p<.001) emerged as statistically significant after multiple-comparisons correction (see Fig. 5.3). Finally, cross-decoding was performed between pictures and videos in both ways (i.e. the SVM classifier was trained on pictures and tested on videos and vice versa). Both when training on videos and testing on pictures (peak [MNI]: -3, -70, 50; cluster p<.001), and when training on pictures and testing on videos (peak [MNI]: -12, -79, 50; cluster p=.001), analyses showed a significant cluster surviving the multiple-comparisons correction in the left Precuneus (Fig. 5.3). See Table 5.2 for further details.

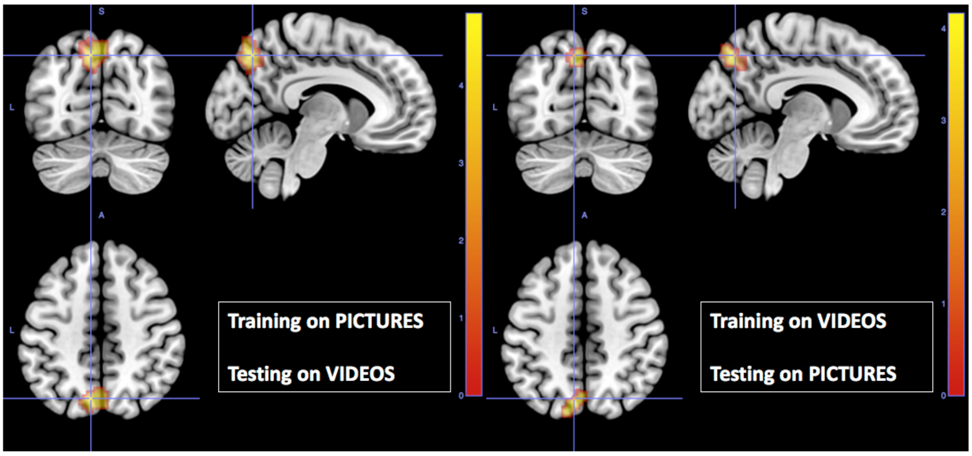

Fig. 5.3 Group-averaged maps of classification accuracy significantly above chance level for seen vs. unseen rooms. Cross-Decoding results obtained training on pictures and testing on videos (left) and viceversa (right) are shown. S = superior; L = left; A = anterior.

Decoding Seen vs. Unseen	**Region**	**Lat.**	**Cluster size**	**T**	**MNI peak coordinates** (mm)		
					x	**y**	**z**
pictures and videos							
	Precuneus	R	698	6.06	6	-76	53
pictures only							
	Precuneus	R	59	5.33	21	-67	29
videos only							
	Precuneus	R	121	3.97	12	-73	59
Cross-Decoding 1							
	Precuneus	L	141	4.03	-3	-70	50
Cross-Decoding 2							
	Precuneus	L	72	4.03	-12	-79	50

Table 5.2 Group-level MVPA results (peak voxels) of decoding using pictures and videos, pictures only, videos only, trining on videos/testing on pictures and training on pictures/testing on videos. Abbreviations: Lat., Lateralization; L, left; R, right. Cross-Decoding 1 = *Training on videos and testing on pictures; Cross-Decoding 2 = Training on pictures and testing on videos.*

DISCUSSION

In this study we aimed to investigate the possibility of discriminating the brain activity patterns related to the processing of visual stimuli showing previously visited rooms, from that of unknown rooms and to highlight the brain correlates of rooms recognition from a multivariate perspective. To address these issues we evaluated the performance of a multivariate classifier in a series of decoding analyses, assessing which brain regions contained more crucial information for the discrimination between seen and unseen rooms. We adopted a searchlight-MVPA approach on regional fMRI activation patterns in response to an implicit recognition task for classification of seen vs. unseen rooms. Furthermore, we performed a "classical" univariate analysis to highlight the differences in brain activity while viewing seen vs. unseen rooms.

In line with previous reports, the univariate analysis revealed higher activity mainly in the precuneus/cuneus (as well as in parietal regions such as superior parietal lobule and angular gyrus), and in the parahippocampal cortex (parahippocampal place area, PPA) for the seen rooms. From a general point of view, some authors proposed for parahippocampal cortex a role in the processing of contextual associations (Aminoff, Kveraga, & Bar, 2013). It is well-known that a particular region within parahippocampal cortex seems to be more active when processing scenes rather than faces or objects, the so-called PPA (Epstein & Kanwisher, 1998). The role of the PPA has been extensively demonstrated in processing relevant information for spatial navigation, such as landmarks (Janzen & van Turennout, 2004; Marchette, Vass, Ryan, & Epstein, 2015). Moreover, in a number of studies PPA has showed greater activity for natural scene categorization (e.g., Walther et al., 2009), visuo-spatial memory (e.g., Cabeza et al., 2004), recognition (e.g., Bar & Aminoff, 2003) and familiarity judgments (Rissman et al., 2016).

On the other hand, a work by Epstein, Harris, Stanley and Kanwisher (1999) apparently contradicted this view, by showing that the activity in the PPA was not related to the level of familiarity of viewed scenes. These authors acquired fMRI images while participants were watching pictures showing scenes from a familiar/unfamiliar environment, landmarks from a familiar/unfamiliar environment, objects or faces. In this experiment, buildings were used as landmarks for the creation of the stimuli. Authors did not found a significantly higher activity in the PPA in response to familiar scenes. By contrast the PPA response was significantly higher for familiar landmarks. One possible explanation for this differential involvement of PPA in familiarity judgment when considering scenes versus landmarks (buildings) has to be searched in the way through which subjects experience a landmark. Activity in the PPA is strong in response to landmarks because participants experience them as part of a scene, rather than detached objects (Epstein et al., 1999; Epstein & Kanwisher, 1998). So that, when viewing a landmark that refers to a previously experienced place (e.g., a room), subjects would probably recognize it as a piece of a more complex scene (Epstein et al., 1999). Moreover, a familiar landmark could drive subjects to "complete" the scene with imagery, thus leading to a greater activation of PPA (O'Craven & Kanwisher, 2000). So that, we argue that in our experiment unknown rooms could be processed as unfamiliar landmarks, and thus as detached objects, rather than scenes. This could account for the stronger activation in PPA while processing pictures of (i.e., pieces of) seen rooms that has been highlighted in the present study.

Moreover, differences in geometric properties between buildings and natural scenes (e.g., a landscape) have to be taken into account. Indeed, from this point of view, rooms are much more similar to buildings than to natural scenes, as they are physically delimited. So that, a picture of scene in a park can be considered a self-contained scene, while a picture taken inside a room is considered as a part of a wider scene, such as a landmark. Therefore, the results of the present experiment (in the univariate contrast seen vs. unseen rooms) could be comparable to the familiar vs. unfamiliar landmarks contrast in Epstein and colleagues' experiment (1999), in which a role of PPA emerged. According to several authors (e.g., Bar & Aminoff, 2003; Davachi, Mitchell, & Wagner, 2003; Eichenbaum, Yonelinas, & Ranganath, 2007; Ranganath et al., 2004) the PPA may act a role in recognition, encoding a global spatial representation of the context in which an item has been encountered. In this view, the PPA activation can be considered as an indirect measure of recognition. In other words, the PPA activation might reveal whether a subject is implicitly locating the picture he is viewing in its own previously experienced spatial context. Moreover, the visual stimuli used in our experiment were not taken from a specific (i.e., subject-specific) point of view, so that the activity of PPA may reflect the view of the same room from different perspectives, already demonstrated in literature (Epstein, Graham, & Downing, 2003). Furthermore, the stronger activity of PPA for seen than for unseen stimuli indicates that in this study PPA did not seemed to respond to complete scene changes, as demonstrated by Epstein and colleagues (2003).

The activation we found in the precuneus is consistent with the literature reporting a role of this area in episodic memory and recognition (e.g., Henson, Rugg, Shallice, Josephs, & Dolan, 1999; for a review see Cavanna & Trimble, 2006). For example, Tulving and colleagues (1994) found an increase in blood flow (measured with PET) in this region, related to the recognition of previously presented sentences. Another work demonstrated that left precuneus is active when specific autobiographical events are retrieved (Addis, McIntosh, et al., 2004).

In the subsequent multivariate analyses the PPA did not seem to be informative for the decoding of seen versus unseen rooms. Overall, multivariate classification analyses demonstrated that it is possible to discriminate between previously seen (i.e., visited) and unseen rooms. The most informative regions for the decoding were observed in the precuneus and involved regions within the retrosplenial cortex (RSC)/ posterior cingulate cortex (PCC). In literature, the RSC is usually active for known positions on an allocentric map (Marchette, Vass, Ryan, & Epstein, 2014; Robertson, Hermann, Mynick, Kravitz, & Kanwisher, 2016; Vass & Epstein, 2013). In other words, RSC is considered to have a so-called *"translational function"* because it helps to switch between egocentric to allocentric coordinates (for a review on RSC functions see Vann, Aggleton, & Maguire, 2009). Burgess and colleagues (Burgess, Becker, King, & O'Keefe, 2001; Byrne, Becker, & Burgess, 2007) proposed a model in which the RSC translates the contextual/spatial information related to an episodic memory from an allocentric representation (located in the hippocampus) to an egocentric one in order to view the remembered scene from a specific point of view. In light of this model, the RSC is supposed to act as a short-term buffer for the representations while the translational process takes place. Although participants had to deal with different viewpoints we found high accuracies on a within subject-basis. This results are consistent with the idea of an important role of the RSC in the translation between allocentric and egocentric coordinates. The difference between the maximum accuracy reached on a subject-level and on a group-level, it might be due to the fact that there could be subject-specific differences in the neural correlates of recognition (familiarity). Each subject can incidentally encode peculiar aspects of the room while visiting it, and the subsequent recognition can be guided by those aspects.

Although activity in the RSC seems to be related to many different processes such as linguistic comprehension and production (Awad, Warren, Scott, Turkheimer, & Wise, 2007), motivational aspects (Small, Zatorre, Dagher, Evans, & Jones-Gotman, 2001) and pain in fibromyalgia (Wik et al., 2006), this region showed a significantly strong role in studies about spatial navigation (Epstein, 2008; Maguire, 2001) and episodic memory (E. A. Maguire, 2001a; Svoboda et al., 2006). In particular RSC seems to show greater activation for the retrieval of recent versus remote autobiographical memories (Gilboa, Winocur, Grady, Hevenor, & Moscovitch, 2004; Oddo et al., 2008; Woodard et al., 2007). Moreover, consistently with our study, some authors (e.g., Addis et al., 2004) demonstrated greater activity in the cuneus for the retrieval of specific autobiographical events.

Overall, we think that one of the strengths of our study is that the incidental memory that has been tested (implicit rooms' recognition) had not been created in a laboratory setting but in a real-world location. Moreover, we used stimuli created from real-world scenes (rooms). A recent meta-analysis has surprisingly demonstrated that functional correlates of memory retrieval (studied with fMRI) for autobiographical memories and for laboratory-encoded memories are not overlapped as much as it can be assumed (McDermott, Szpunar, & Christ, 2009). Rissman and colleagues (2016), starting from this finding, performed an across-studies decoding between laboratory-based and real-world memories. More specifically, these authors evaluated the performance of a multivariate classifier, trained on previously acquired fMRI data from a laboratory-based face recognition task (Rissman et al., 2010), in classifying real-world autobiographical memory states. The reverse analysis (i.e., training on real-world autobiographical memories and testing on laboratory-based memories) was also performed. Results showed a reliable decoding performance in both analysis. In this interesting recent work, Rissman and collaborators (2016) investigated the neural patterns associated to the recall of memories for real-world autobiographical events (after a time interval ranging from 1 to 3 weeks). Participants were asked to wear a necklace-mounted digital camera during their everyday life's activities for 3 weeks. Pictures acquired by this camera were then used as stimuli in the subsequent explicit recalling task inside the fMRI scanner, in which they were asked to make memory judgments about whether photos were showing events from their own lives or not (i.e., images captured by other participants' cameras). Authors found high accuracy in the discrimination between recognized and rejected pictures, as well as between recollection, familiarity and novelty, through the use of a multivoxel pattern analysis. If compared to the study by Rissman and colleagues (2016), the current experiment has an important strength, that is the use of visual stimuli taken from different (and not subject-specific) points of view. This feature, in our opinion, make our study more ecological in particular if referred to possible future forensic applications of this kind of techniques (also taken into account by Rissman et al., 2016) which will be discussed in Chapter 7.

Chapter 6

“DO YOU REMEMBER IT?”
DECODING AUTOBIOGRAPHICAL EMOTIONAL MEMORIES FROM BRAIN ACTIVITY

INTRODUCTION

The neural correlates of autobiographical memory retrieval is represented by a complex network of brain areas that includes medial prefrontal, temporal and retrosplenial cortices and cerebellum (Maguire, 2001a; Piefke & Fink, 2005; Svoboda et al., 2006). A number of factors are involved in autobiographical memory retrieval. Among these, the most studied are the emotional relevance and the relevance for the self.

Despite the relation between emotions and the functions of autobiographical memory is not clear, a broad range of behavioral studies (e.g., D’Argembeau, Comblain, & van der Linden, 2003) have indicated that emotional valence is one of the main components of this memory system (for a review see Holland & Kensinger, 2010) and seems to have a key-role in encoding, storage and retrieval of autobiographical memories (Conway, 1990). It is a common experience that emotionally relevant memories are easier to retrieve. Some authors (e.g., Conway, 1990) argued that when people experience emotional events, the encoding process is enhanced. So that, highly emotional events are more easily accessible than less emotional memories (Buchanan, 2007; Holland & Kensinger, 2010; Robinson, 1980). Moreover, the emotional valence of episodic memories seems to alter brain activity patterns during recollection (for a meta-analysis see Svoboda, McKinnon, & Levine, 2006). According to some authors (e.g., Brewer, 1986), the self influences the encoding and retrieval of events from memory and, on the other hand, autobiographical memories are considered as having a key-role in building the self. Self-referential processing is thus a crucial component of autobiographical memory and it is considered as the “organizer” of other memory systems with lower levels of complexity (Conway & Pleydell-Pearce, 2000; Svoboda et al., 2006).

The great majority of neuroimaging studies about autobiographical memory have adopted standard univariate approaches to functional data analysis. Only few studies have applied multivariate techniques, that allow for the analysis of patterns of activated voxels. A technique that has been utilized in the neuroimaging research of autobiographical memory is the *spatiotemporal partial least square* analysis (ST-PLS), which permits the identification of cross-correlations between the activity of the brain in different time-points related to experimental factors. Addis and co-workers (Addis, McIntosh, et al., 2004) adopted this approach and found different temporal patterns of activity when participants were retrieving general rather than specific autobiographical memories, thus supporting the autobiographical memory hierarchical model proposed by Conway (Conway, 1992; Conway & Pleydell-Pearce, 2000).

More recently, MVPA methods have been adopted for the purpose of classifying true memories from stimuli not-related to any memory content (e.g., Rissman, Chow, Reggente, & Wagner, 2016; Rissman et al., 2010; Uncapher, Boyd-Meredith, Chow, Rissman, & Wagner, 2015).

For instance, Rissman and co-workers (2010) used multivariate methods of fMRI data to detect whether participants were processing previously encountered *versus* new faces.

To our knowledge, none of the studies that used MVPA for memory detection explored the possibility of using verbal statements, describing specific autobiographical information, as stimuli for the decoding. The main advantage of the use of this kind of stimuli to describe autobiographical events is the possibility to customize them to the personal subjective memories. Such a technique can also be applied to the study of emotional autobiographical memories, as the stimuli can be built using the subject's own words and this can foster the re-experience of emotionally relevant aspect of the investigated memories.

The aim of the present fMRI experiment was to investigate whether it is possible to discriminate between sentences describing real autobiographical episodic memories from sentences describing the same memories but using false episodic information (i.e., depicting a piece of the event that actually did not happen; see below for some examples), as well as the level of accuracy that can be reached in detecting each subject's memory. In this experiment, we decided to focus on a highly relevant negative emotional memory in order to maximally stimulate the relevance for the self and thus the ease in the recognition. Moreover, the use of negative emotional autobiographical memories can shed light on what happens in the brain when recognizing a negative event of our life. This can lead to interesting potential implication of these results for the forensic field in future (e.g., an individual suspected for a murder can be investigated in order to determine the presence of memories of the murder). Potential forensic implications will be discussed in Chapter 7. In the current experiment, a specific kind of episodic memory (i.e., the death of a family member) was chosen and the participant-specific episodic memory was investigated. Moreover, we investigated whether different degrees of personal active involvement (e.g., the memory of an important one-by-one conversation with a friend is characterized by a higher personal involvement than the memory of the last time we went at the supermarket) in the retrieved event had an effect on the neural correlates of autobiographical memory recognition, and whether this affects the accuracy in the discrimination of true and false sentences about different aspects of the investigated memory. For this purposes, we first investigated the differences in brain activity between the processing of true memories and non-memories through the use of a standard univariate analysis. Then, we evaluated the performance of a multivariate classifier in revealing which brain areas contained more useful information for discriminating the brain activity pattern associated to true memories from that of non-memories.

MATERIALS AND METHODS

Participants

This study was approved by the local ethics committee, University of Modena and Reggio Emilia, Italy. Six healthy right-handed participants were recruited (age range: 25-49; mean=36; SD=9.86). All participants signed informed consent before participation and were remunerated. Only female participants were recruited as previous studies have shown gender differences in autobiographical memory processing (e.g., Cahill et al., 2001; Piefke & Fink, 2005; Piefke, Weiss, Markowitsch, & Fink, 2005), so that we chose to not introduce gender as variable in our design. All participants had normal or corrected-to-normal vision and were free from any neurological or psychiatric diagnosis.

Preliminary interview and stimuli creation

In order to be consistent with our aims, the fMRI recognition memory task was focused on a personal emotionally relevant negative autobiographical event. A semi-structured interview was carried-out with each participant, during which they were asked to focus on a particular negative personal event: the memory of the recent death of a family member. Then, they were asked to selectively report information related to four categories: a) personal data of the deceased person; b) the moment in which they were told about the death; c) the last time they met the deceased person; d) the funeral. On the basis of the collected information, a set of sentences was built. First, sentences based on real memories were built (e.g., *"When I was told about his death I was alone"*). Then we created sentences describing an alternative event (e.g., *"When I was told about his death I was with my friends"*). See Table 6.1 for further details and for examples of sentences used in the experiment.

Category	**Kind of information**	**Examples of true memory sentences**	**Examples of non-memory sentences**
a	Personal data	*"She was a professor"*	*"Her height was 1.70 m"*
b	News of the death	*"When I was told about his death I was alone"*	*"When I was told about his death I was with my friends"*
c	Last meeting	*"During our last meeting we were alone"*	*"During our last meeting there was also my mother"*
d	Funeral	*"The funeral took place in the morning"*	*"During the funeral it was raining"*

Table 6.1. Examples of true memory and non-memory sentences used in the experiment.

Forty-eight sentences were built for each participant (24 real-memory sentences and 24 non-memory sentences) and each sentence was presented only once. In order to create comparable stimuli, alternative sentences (i.e., real-memory vs. non-memory) about the same piece of information had the same length. Overall, sentences' length ranged between four and seventeen syllables.

For the aims of the current study, the episodic details of interest were those referring to events were the participant had a main role, such as categories b and c (i.e., information about the moment in which they received the news of the death and about the last meeting with the deceased person). Consequently, the whole set of obtained information was divided in episodic details about an event with a personal main-role ("Personal Event") and details about a public event (funeral) or general information about the deceased person categorized as "Not Personal Event".

fMRI task

During the subsequent fMRI experiment, participants were asked to watch a screen presenting sentences about the negative event they previously reported. After the presentation of each sentence, they were asked to give an answer judging whether the sentence was true or false by pressing two buttons with the index or the middle finger of the right hand. The experiment comprised three functional runs, lasting around 8 min each. Each of the three runs consisted of 16 trials in which a warning cue appeared for 500 ms, then a sentence was presented for 5500 ms. The chosen duration of sentence presentation is consistent with electrophysiological studies (e.g., Conway, Pleydell-Pearce, Whitecross, & Sharpe, 2003) showing that the average time needed for the retrieval of autobiographical memories was 5 secs (range: 3 – 9 secs) and fMRI studies (e.g., Addis, McIntosh, Moscovitch, Crawley, & McAndrews, 2004) showing that the autobiographical memory activation pattern reaches its peak between 2 and 8 s. In the present study, after stimulus disappeared, a twelve seconds resting-state interval was inserted, followed by a 2 s time window in which subject were asked to give their response. Finally, 10 seconds of resting-state were acquired before the next trial started. The stimuli were presented in a randomized order across all runs, with the only restriction that in each run half stimuli were true sentences and half false sentences.

fMRI Data Acquisition

Functional MRI volumes were acquired with a Philips Gyroscan Intera 3.0 T MR system equipped with IFIS fMRI system with standard head coil (30 slices, TR = 2000 ms, echo time TE = 35 ms, resolution 3 × 3 × 3 mm^3 with 1 mm gap, flip angle = 78°, Field of View [FoV] = 240 mm). For all participant, 280 images were acquired for each of the three runs. We also acquired structural MRI data (T1-weighted MPRAGE: 192 axial slices, TR = 1900 ms, TE =2.52 ms, flip angle = 9°, FoV = 256 mm).

fMRI Data Analysis

Functional MRI data preprocessing and standard univariate analyses were performed using SPM12 (v. 6685; http://www.fil.ion.ucl.ac.uk/spm/) running under Matlab R2014b (MathWorks, Inc.). The first fifteen EPI volumes of each functional timeserie were discarded. The images were slice-time corrected using the first slice as reference, motion corrected, coregistered to the anatomical image, and finally a spatial smoothing with a Gaussian kernel of 6 mm FWHM was applied. Data were high-pass filtered with a cutoff period of 128 s. A general linear model (Friston et al., 1995) was built on the basis of each trial onset and applied to the data of each run and four event-based HRF-convolved regressors were modeled (2 memory types [PE, NPE] × 2 sentence status [true, false]): true details of a personal event (tPE), false details of a personal event (fPE), true details of a not-personal event (tNPE), false details of a not-personal event (fNPE). Movement parameters derived from the realignment preprocessing step were included as additional regressors of no interest. At a subject-level a separate contrast was made for each of the four regressors. The resulting contrasts were spatially normalized to the MNI space and entered into a second-level random effects ANOVA with the stimulus type (tPE, fPE, tNPE, fNPE) as within-subject factor. The contrasts between true and false memories were investigated considering both memory types ([tPE + tNPE] > [fPE + fNPE]) and selectively considering episodic events with (tPE > fPE) and without (tNPE > fNPE) a main-role of the participant. Moreover, in order to account for a possible inflated false positives rate in our results (Eklund, Nichols, & Knutsson, 2016) due to the small sample size, we repeated the second-level analysis with a nonparametric permutation method (Hayasaka & Nichols, 2003, 2004; T. Nichols & Holmes, 2003), performed using the toolbox SnPM (http://www.nitrc.org/projects/snpm/). Finally, we ran a new first-level GLM analysis on unsmoothed images, modeling one regressor for each stimulus, in order to prepare data for the subsequent MVPA analysis.

Multi-Voxel Pattern Analysis

In the current experiment different memories of the same kind were investigated for each subject, so that we did not expect a common brain pattern for the discrimination between memory-related and non-memory-related stimuli. Therefore, a single-subject Multi-Voxel Pattern Analysis (MVPA) was performed for each participant (i.e., no across-subject decodings were performed). The aim of this analysis was to investigate whether it is possible to accurately discriminate between the brain activity patterns elicited by true-memory sentences and non-memory sentences.

Two "True" vs. "False" decoding analyses were performed for each subject. In the first one, both stimuli referred to memories with and without a personal main-role were considered (PE + NPE), while the second decoding was selectively focused on PE stimuli. The decoding analyses were performed on the basis of the following steps. First, a GLM was estimated on unsmoothed data, with one regressor modeled for each stimulus (as specified above). Second, the estimated GLM parameters were included in each within-subject decoding analysis. A “searchlight” approach (Kriegeskorte et al., 2006) was used to search across the whole brain for the voxels that contain more information in an unbiased fashion. A spherical searchlight with a radius of three voxels was adopted and each voxel in the brain served once as the center of the searchlight sphere. As previously reported, the third step consisted in training and testing a Support Vector Machine classifier (SVM; Cortes & Vapnik, 1995). Following what has previously been introduced in this manuscript, in order to avoid overfitting, a Leave-One-Out Cross-Validation (LOOCV; see Chapter 2) approach was adopted. In the case of the present experiment, the classifier was trained on GLM parameter estimates from two runs and tested on the GLM estimates from the remaining run. This procedure, called Leave-One-Run-Out cross-validation (LORO), was repeated until each run was iteratively used once as the test dataset. Importantly, each run included different stimuli, so that the three runs did not included repeated stimuli. The training and testing of the support vector classifier was performed using The Decoding Toolbox (TDT; Hebart, Görgen, Haynes, & Dubois, 2015; https://sites.google.com/site/tdtdecodingtoolbox/) and LIBSVM (Chang & Lin, 2011; http://www.csie. ntu.edu.tw/~cjlin/libsvm/). Each decoding analysis resulted in a three-dimensional subject-specific accuracy map showing a value of accuracy for each voxel in the brain.

RESULTS

Behavioral data analysis

All stimuli but four (which were not followed by any response) were correctly classified by participants. Participant 5 missed one response, while participant 6 missed three responses. Stimuli not followed by a behavioral response were discarded both from the univariate and multivariate analyses.

Univariate analysis

The brain activity related to stimuli describing true autobiographical memories characterized by a personal main-role (compared to false sentences about the same event; tPE > fPE) revealed significant activations (FWE-corrected for multiple comparisons at cluster level; cluster-forming threshold=.001) in the left anterior cingulate cortex and in the right superior temporal gyrus (see Table 6.2 and Fig. 6.1A). The contrast between true and false sentences without considering the memory type ([tPE + tNPE] > [fPE + fNPE]) and the contrast focused only on NPE stimuli (tNPE > fNPE) did not show any significant voxel surviving multiple comparisons correction.

Contrast	**Region**	**Lat.**	**Cluster size**	**T**	**MNI peak coordinates** (mm)		
					x	**y**	**z**
tPE > fPE							
	Superior Temporal Gyrus	R	35	7.82	63	-49	11
				4.23	66	-46	20
	Anterior Cingulate	L	164	5.61	-3	41	8
				4.87	0	38	23
				4.11	-12	50	11

Table 6.2 FWE-corrected for multiple comparisons at cluster level; cluster-forming threshold=.001

Non-parametric permutation analysis results (FWE-corrected for multiple comparisons at cluster level; cluster-forming threshold=.001) confirmed the significant frontal cluster (left anterior cingulate cortex), while no significant voxels resulted in the superior temporal gyrus. In Figure 6.1A and 6.1B the parametric and non-parametric results are shown.

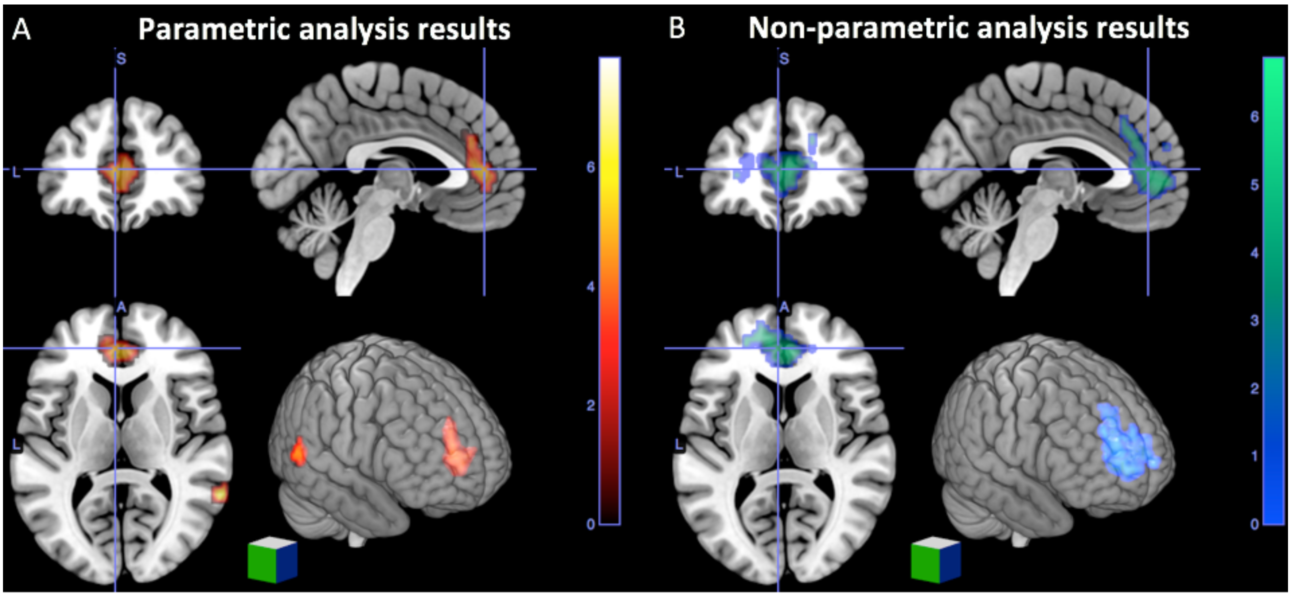

Fig. 6.1 Group-level univariate results for the contrast tPE > fPE. A) parametric results (SPM); B) non-parametric results (SnPM). L = left; S = superior; A = anterior

Multivariate analysis

In the multivariate analysis section, the performance of a linear SVM classifier was examined in discriminating between true sentences describing autobiographical memory and information not memory-related. The classification was based on the fMRI activity pattern investigated through a searchlight approach. Two MVPA decoding analyses were performed in order to test whether single sentences describing real autobiographical memories can be detected from brain activity, where the information about the classification is encoded in the brain and how accurately true autobiographical memories and non-memoriescan be discriminated. The best accuracies reached in each within-subject decoding analysis, are reported in Table 6.2.

As a first step, a GLM with one regressor per stimulus was estimated as previously described. Then, multivariate pattern classification was performed for each participant on the GLM parameter estimates using a LORO cross-validation scheme, to assess whether the response pattern encoded information about the discrimination between true and false memories. Therefore, an SVM algorithm was trained (for further details please see the "Multi-Voxel Pattern Analysis" paragraph above) in discriminating true and false memories without differentiating for the memory type. In order to avoid biased decoding results due to the unbalanced number of stimuli belonging to the different classes across-subjects, a repeated subsampling strategy was adopted. Within the cross-validation design, the more frequent class was repeatedly subsampled for 10 times, running multiple classification iterations, and results were averaged. Results (see Table 6.2) indicated that the most informative voxels showed maximum classification accuracies around 80% (range=76.7-86.5; mean=81.2%; SD=3.65). These values of accuracy indicate the ability of the algorithm in classifying each single item as a true memory or a non-memory. Then, we verified the performance of the machine learning algorithm selectively on PE stimuli. As shown in Table 6.2, the averaged maximum accuracy in classifying each single item increased to a value around 90% (mean=89.1%; SD=2.74) when considering only autobiographical memories in which the participant had a direct and main involvement (PE; see Fig. 6.2).

Participants	Decoding type Single-subject PE + NPE LORO Max Acc.	 Single-subject PE LORO Max Acc.
P1	83.3	86.8
P2	86.5	87.5
P3	80.7	90.5
P4	77.6	91.4
P5	76.7	92.5
P6	82.1	85.8
Mean	81.2	89.1
SD	3.65	2.74

Table 6.3 Best accuracies reached in single-subject searchlight classification analyses. LORO = Leave-One-Run-Out Cross-Validation; SD = Standard Deviation; PE = "Personal Event"; NPE = "Not-Personal Event"; Max acc. = maximum accuracy

In Figure 6.2 the increase in maximum accuracies between the decoding using PE + NPE and the decoding focused only on PE stimuli is shown. The difference in the peak decoding accuracies was statistically significant (t=-3.191; df=5; p=0.024).

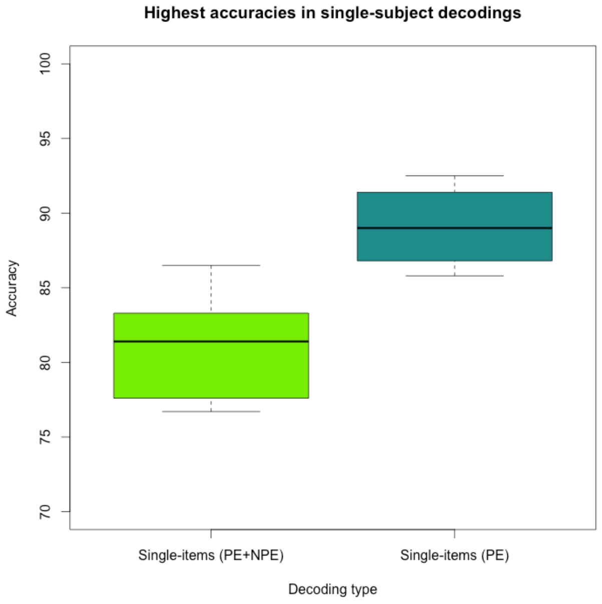

Fig. 6.2 Comparison of single-subject maximum accuracies reached in the two MVPA analyses (on PE+NPE and on PE only). PE = "Personal Event"; NPE = "Not-Personal Event".

As expected, strong between-subjects differences emerged in the patterns of informative voxels identified through the searchlight-based MVPA analysis, as shown in Fig. 6.3.

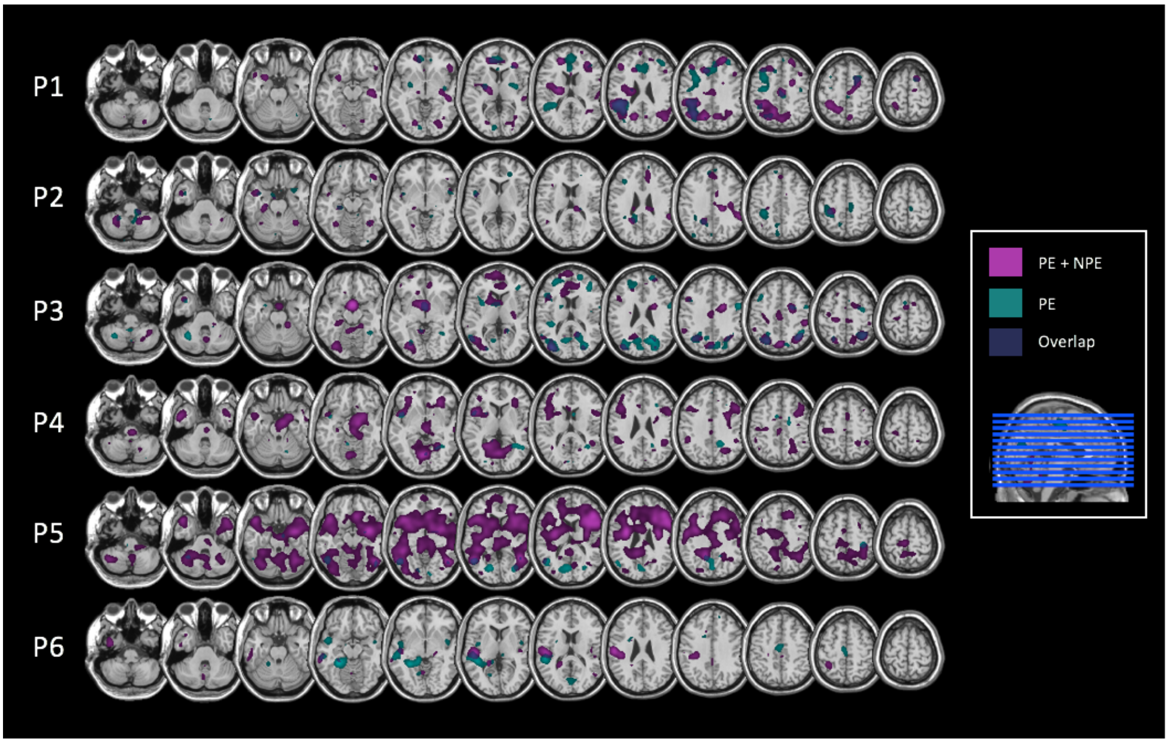

Fig. 6.3. Overlays of normalized single-subjects' accuracy maps only showing accuracies greater than 60%. The pattern of informative voxels is represented with different colors for each participant. As can be easily noticed, the most informative voxels belonged to completely different patterns in different participants. PE = "Personal Event"; NPE = "Not-Personal Event".

The resulting accuracy map obtained for each subject were then entered in a second-level one-sample t-test. Results (FWE-corrected for multiple comparisons at cluster-level with a cluster-forming threshold of 0.001) showed a significant cluster of informative voxels in the left lingual gyrus (see Table 6.3 and Fig. 6.4) for the decoding using PE and NPE stimuli. A significant cluster involving the lingual gyrus and posterior cingulate cortex (PCC) was confirmed also by a non-parametric version of the analysis. By contrast, the decoding analysis performed only on PE stimuli did not show any significant voxel surviving multiple comparisons correction.

Region	Lat.	Cluster size	T	MNI peak coordinates (mm)		
				x	y	z
Lingual Gyrus	L	60	12.27	-21	-61	-4
			11.37	-6	-73	8
			9.96	-18	-70	-4

Table 6.4 Results of the group-level one-sample t test run on decoding accuracy maps (PE+NPE). Results are FWE-corrected for multiple comparisons at cluster level; cluster-forming threshold=.001. PE = “Personal Event”; NPE = “Not-Personal Event”.

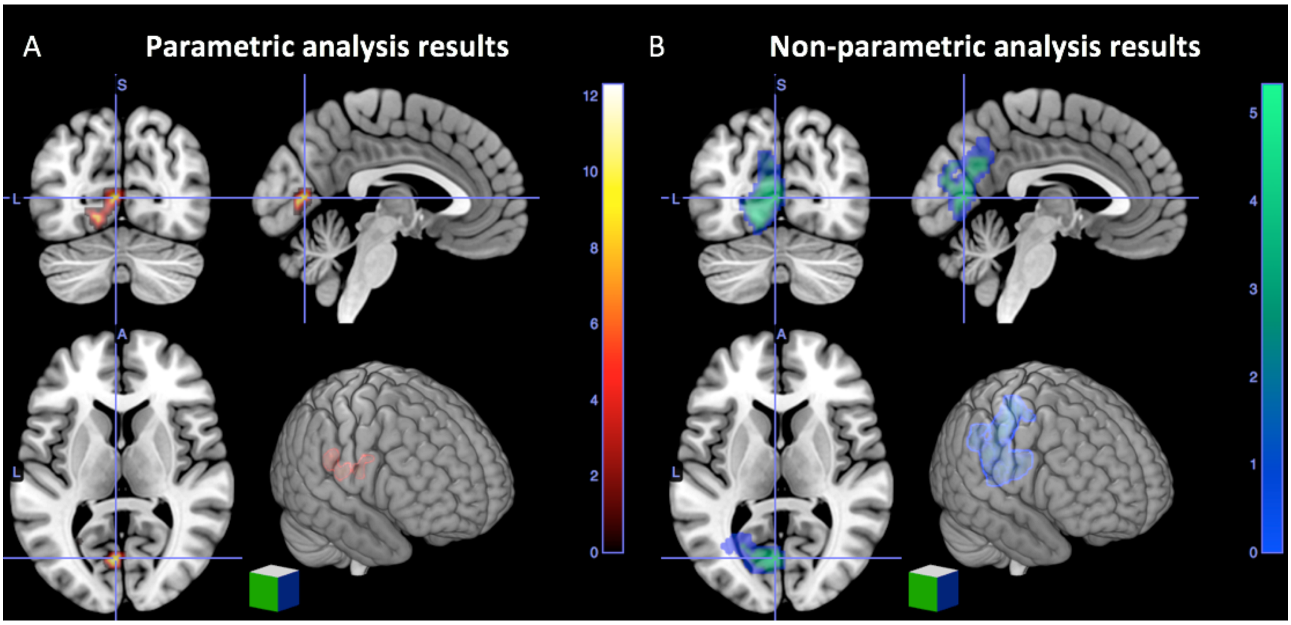

Fig. 6.4 Results of group-level one-sample t-test on decoding accuracy maps (PE+NPE). A) parametric results (SPM); B) non-parametric results (SnPM). L = left; S = superior; A = anterior. Results are FWE-corrected for multiple comparisons at cluster level (cluster-forming threshold=.001). PE = “Personal Event”; NPE = “Not-Personal Event”; S = superior; L = left; A = anterior.

DISCUSSION

The present study mainly aimed at investigating the possibility of classifying single items fMRI scans as referred to true autobiographical emotional episodic memories or not, using MVPA. Preliminarily, a standard univariate analysis was also performed at group-level, in order to test the consistency of the present results with previously reported findings.

Participants were presented with sentences describing true or false autobiographical memories about the death of a family member. Although all sentences were linked to the same personal memory, they were related to four categories based on the kind of information considered, namely: personal data of the deceased, the moment when the participant was told about the death, the last meeting with the deceased and the funeral. These four categories can be grouped according to the degree of personal involvement of the participant: the last meeting with the deceased and the moment when the participant was told about the death can be considered events characterized by a strong involvement, that is, the participants had a main personal role in both these situations (so that stimuli belonging to these categories have been called “Personal Events”, PE). On the contrary, the category about the personal data of the deceased and about the funeral described memories not characterized by a direct main involvement of the participant (“Not-Personal Events”, NPE). This categorization was used both in the univariate and in the multivariate analyses. The fMRI univariate analysis investigated the neural substrates of episodic autobiographical memory recognition and retrieval, showing increased activity in MPFC (i.e., left anterior cingulate) and in right superior temporal gyrus for true memories of event with a personal main involvement (if compared to non-memories of the same kind; contrast tPE>fPE). These results are consistent with previous literature on episodic memory retrieval. Indeed, a number of studies have demonstrated a key-role of MPFC in self-reference (Kircher et al., 2000; Northoff et al., 2006), that is one of the main components of autobiographical memory recollection, as well as time processing (Tranel & Jones, 2006). Moreover, this brain area seems to have a role in emotion processing (Phan, Wager, Taylor, & Liberzon, 2004, 2002), thus activity in MPFC is consistent with the processing of emotionally relevant stimuli as those used in this study. In a previous work, Oddo and colleagues (2008) investigated the neural basis of recent episodic autobiographical memories (compared to semantic memories) focusing only on young female participants. This experiment explored the effect of time on the neural substrates of autobiographical and semantic memory retrieval by investigating memories with different degree of remoteness (i.e., from different life periods). Authors found that MPFC was specifically engaged in the retrieval of recent autobiographical memories and suggested a specific role of MPFC in processing autonoetic, emotional and self-related features of recent autobiographical memories. Furthermore, some studies highlighted a crucial role of left anterior cingulate in the self-processing. For example, in an fMRI study by Kircher and colleagues (2000) the neural systems involved in self-relevant information processing were investigated. Authors compared judgments of self-relevant traits and self-irrelevant traits. Subjects were asked to judge whether psychological trait adjectives were descriptive of themselves. Findings showed activation in the bilateral precuneus, left superior parietal lobule (SPL), left lateral prefrontal cortex and left anterior cingulate, and were interpreted as suggesting the presence of a self-processing brain network.

An increase in the activity of the superior temporal gyrus was also found. Although activity in this region has been associated to the process of episodic recollection (e.g., Kirwan, Bayley, Galván, & Squire, 2008; Yonelinas, Otten, Shaw, & Rugg, 2005), the involvement of this area has not been confirmed with the non-parametric analysis, so that, given the small sample size, it could represent a false positive.

The main aim of the present study was to test the possibility of discriminating the brain activity pattern related to single true autobiographical memories from that related to information not stored in memory. In order to address this point, two MVPA analyses were performed. In the first, both stimuli describing events where the participant had a personal main role (i.e., the last meeting with the deceased and the moment when they were told about the death; PE) and different kind of memories (i.e., semantic information [personal data of the deceased] and an event without a direct main-role of the participant [the funeral]; NPE) were included. The second decoding analysis was selectively focused on PE stimuli. Results showed high performances in classifying fMRI scans related to the processing of single sentences describing true episodic autobiographical memories and non-memories. When considering both PE and NPE stimuli, the maximum average accuracy reached was around 80%, while, surprisingly, when focusing only on PE stimuli, the average maximum classification accuracy increased to about 90%. These results indicate that, in a within-subject decoding analysis, is possible to detect highly informative brain areas about the discrimination between sentences related to real-memories and those not memory-related.

The decoding analysis focused on PE did not show significant voxels at group-level, probably because of the small and unbalanced number of stimuli belonging to categories 2 and 3 (see "Preliminary interview ad stimuli creation" paragraph). Moreover, given the small sample size, we think that these results should be considered in any case promising, and that future studies might be focused only on stimuli describing events with a personal main involvement. Moreover, the accuracy values shown in the multivariate analyses are even more promising if the characteristics of the experimental design are taken into account. Indeed, in the present study there were no repetitions of the same stimulus across the three fMRI runs. This means that each item used for training the classifier was unique. The main advantage of the absence of stimuli redundancy is that such a procedure can go beyond the so-called "repetition suppression". This effect refers to the attenuation of neural responses to repeated stimuli if compared to the response to a single stimulus, and has been demonstrated both in non-human primates (i.e., single-unit recordings; Desimone, 1996) and in humans (i.e., fMRI studies; Henson & Rugg, 2003). This phenomenon seems to reflect both the neuronal adaptation process (i.e., the reduction in neuronal firing rates over time to a repeated stimulus) and attention-related effects of expectation (Larsson & Smith, 2012). Hence, the use of unique stimuli of the same kind (i.e., different information related to the same memory content) can avoid signal loss due to repetition suppression.

As shown in Fig. 6.3, the discrimination maps (i.e., accuracy maps) were strongly different across subjects, indicating that the pattern of most informative voxels for the discrimination between true memories and non-memories involved different brain regions for each subject. This can be due to two reasons. First of all, although the investigated memory was the same (i.e., death of a family member), each participant was tested on his/her own personal emotional memory. Moreover, here we used stimuli (i.e., sentences) that are less structured if compared to visual stimuli such as pictures (as in the previous work; see Chapter 5). Taken together, these aspects might explain the huge amount of between-subjects variability shown in multivariate discrimination patterns. However, it has been shown that a key role in the discrimination between true memories and non-memories is played by the left lingual gyrus. Indeed, as reported in previous studies, the lingual gyrus seems to be involved in the recognition of different kinds of negative emotional stimuli. For example, Taylor and colleagues (1998) in a PET study, found greater activation in this area during recognition of negative images (compared to neutral images). Furthermore, another PET-study (Osaka, Yaoi, Minamoto, & Osaka, 2013) showed increases in activity in the left lingual gyrus during the recognition of negative emotional sentences. More generally, lingual gyrus seems to be involved in emotion processing. Indeed, enhanced activity in this area has been demonstrated while processing unpleasant unmoral sentences (if compared to neutral statements; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002) and words associated with events having a negative emotional meaning (e.g., the word *"assassinate"*; Isenberg et al., 1999).

Overall, the results of previous works have shown a crucial role of lingual gyrus in the recognition of emotionally relevant memory contents. Moreover, this brain region has been indicated as part of a neural network involved in negative emotion processing. Despite the huge between-subject variability, the results of the present multivariate analysis can be interpreted in the same direction, showing (at the group-level both in parametric and in non-parametric analysis) the left lingual gyrus as a significant informative region in classifying whether a single negative emotional sentence describes a real autobiographical event or not. In our opinion, this is the main strength of the present work, as it shows the possibility of accurately detect single emotional autobiographical memories from individual fMRI scans. This aspect could be potentially useful in forensic investigations (see next chapter for a discussion).

However, the present work has some important limitations that should be overcome in further studies. The main limitation is undoubtedly the small sample size. Indeed, given its emotional impact, the particular memory investigated was the factor that mainly affected participants' recruitment. Therefore, the present study does not allow to know whether the efficacy of the method adopted can be generalized. Moreover, although the use of cross-validation procedures in MVPA analyses, the relatively small number of stimuli force us to carefully interpret the present results, as the risk of overfitting (see Chapter 2) must be taken into account.

Furthermore, in future studies, different approaches could be adopted, such as ROI-based analysis instead of a searchlight-based MVPA, in order to optimize the classification accuracy-level rather than focusing on spatial localization of the most informative brain areas.

Chapter 7

GENERAL DISCUSSION AND CONCLUSIONS

Multivariate analysis has been used in four neuroimaging studies in order to test the efficacy of pattern classification techniques on neuroimaging data for two specific applications. As previously discussed, Multi-voxel Pattern Analysis (MVPA) relies on the so-called *reverse inference* mechanism (Poldrack, 2006; Poldrack, 2008; Poldrack, 2007), which substantially inverts the direction of the inference process based on standard GLM analysis techniques, the *forward inference*, intending to infer mental states by looking at brain imaging data. The mental or cognitive states that this technique aims to detect can belong to two categories, namely: “specific states”, which refers to a temporary cognitive state (i.e., related to a task), such as the kind of stimulus a subject is perceiving; or “general states”, referring to stable and not task-related conditions, such as a neurological diagnosis or more broadly, all quantitative structural measures. The idea is that it is possible to detect the presence of a mental state by analyzing brain imaging data patterns. Over the last few years, a number of studies have started testing the effectiveness of these multivariate pattern-based approaches on numerous topics, with fMRI and MRI data analysis. Among all the available multivariate methods for neuroimaging data analysis, linear SVMs are the most commonly adopted, given the relative simplicity in interpreting their results and their generalization ability. Moreover, these algorithms have the ability to overcome the limitations of traditional univariate approaches (i.e., GLM) and of other data-driven techniques.

In the present dissertation, the application of MVPA methods on neuroimaging data has been tested in regards to two different topics (separately addressed in Section A and Section B of the current manuscript), which could potentially have implications in clinical and forensic cognitive neuroscience fields in the near future. In addition to MVPA, a standard univariate analysis has been performed for each study. This has been done for two main reasons: first, to adopt a double approach to each specific research question, thus having a measure of robustness of the results; second, to look at the differences between univariate and multivariate results. However, the aim of the present dissertation is to test the effectiveness of MVPA approach in detecting mental states on critical research questions that could have positive repercussions not only on academic research, but also on real-life settings.

In section A, structural MRI images have been used to investigate the brain correlates of Cognitive Reserve (CR) from a multivariate perspective. In particular, the main goal was to test whether it is possible to discriminate between different levels of CR (using Education as CR proxy) on the basis of the regional amount of Grey Matter (GM). CR refers to the active component of reserve, and can be defined as the ability maximize a performance through the use of alternative brain networks and cognitive strategies in order to better perform a task (Yaakov Stern, 2002, 2009). By contrast, Brain Reserve (BR), refer to the passive component of reserve, defined as individual brain structure differences which can help cope better with pathology (Satz et al., 1993; Yaakov Stern, 2009). Although these two models have been proposed as alternative explanations of the variability in the resilience to brain pathology, some authors suggested that there could be a relation between these concepts. The two studies proposed in Section A of the present manuscript have investigated whether CR can be predicted on the basis of BR measurement, thus assessing the possibility that Cognitive and Brain Reserve theories can describe two sides of the same coin. In the first study (Chapter 3) peculiar patterns of GM volume associated with high and low levels of CR have been investigated. Firstly, a preliminary Voxel-Based Morphometry (VBM) was performed to explore areas associated with higher CR level from a standard univariate perspective. Then, a series of MVPA classifications were performed to discover brain areas mainly encoding information about the discrimination between high and low CR subjects. The first classification was performed considering the whole-brain volume, in order to search all over the brain for patterns of GM discriminating between high and low CR subjects. Then ROI-based analyses were peformed to test the robustness of the results previously obtained in a study by Foubert-Samier and colleagues (2012) on a large cohort. MVPA analyses highlighted areas within the parietal lobe as the most informative for the classification between high and low CR. The second study (Chapter 4) went one step further asking more precisely where in the parietal lobe the most informative clusters are located. To address this point, first a preliminary VBM analysis was initially performed on the parietal lobe only. Then a MVPA searchlight analysis was performed, and the results highlighted a specific cluster in left supramarginal gyrus which seems to encode the most critical information that allows to classify high vs. low CR subjects.

The application of MVPA techniques investigated in chapters 3 and 4, can have useful implementations in clinical practice. Indeed, CR is an important aspect that should be taken into account when assessing patients with cognitive deficits. In particular, in neurodegenerative diseases such as Alzheimer Disease (AD), Cognitive Reserve level can critically influence the progression of the pathology and the efficacy of a cognitive training program (Mondini et al., 2016; Yaakov Stern, 2009; Tucker & Stern, 2011). At the same time, Brain Reserve contributes to protect brain against neurological pathologies. From this perspective, it is extremely useful not only to quantify the CR level, but also to measure the cerebral (protective) changes following CR level, for example in order to make plausible expectations towards a cognitive training efficacy. Moreover, given that inter-individual variations in cognitive reserve can critically affect brain structure, they represent a problem for the application of pattern recognition techniques in clinical setting (Haller et al., 2014), for example for the early diagnosis of AD (Salvatore et al., 2016). So that, the possibility of quantifying the brain correlates of the cognitive reserve level, in a specific individual, could be a crucial preliminary step for the systematic use of automatized diagnostic techniques for neurodegenerative diseases.

In section B, a different application has been investigated. Functional MRI data have been analyzed in order to investigate the neural correlates of autobiographical memories and the possibility to discriminate the pattern of brain activity associated with the processing of information matching to autobiographical memory content. In particular, the main aim of this section was to determine how accurately the presence or absence of specific autobiographical individual memories could be detected based on the analysis of fMRI activity patterns. Some recent papers have tested the efficacy of MVPA techniques as instruments for memory detection that could be potentially useful in forensic field. However, the majority of previous studies on this topic used laboratory-based stimuli (Rissman et al., 2010; Uncapher et al., 2015). Conversely, here the possibility of using real-world event memories has been investigated in two studies addressing different aspects of possible forensic issues.

Many critical issues in the criminal law system depend on proofs based on declarative testimonies. The main purpose of evaluating witness or defendant declarations is to analyze of the dynamics of the crime, i.e. the reconstruction of what most likely happened. However, the introduction of scientific tools in the traditional investigative practice could become extremely useful to deal with some critical questions in the future. Many forensic questions rely on the evaluation of individual memory contents. For this reason, in recent years, a debate about the potential use of brain reading methods as memory detection techniques has emerged within the field of forensic neuroscience and forensic psychology between neuroscientists, psychologists and legal scholars (Bles & Haynes, 2008; Meegan, 2008; Rissman et al., 2016; D. L. Schacter & Loftus, 2013; Shen, Francis X; Jones, 2011). Indeed, MVPA analysis of functional MRI data could potentially be used to reveal covert mental states or personal information without the need of a person's will (Haynes & Rees, 2006). Previous fMRI studies have shown the possibility to accurately determine whether a probe stimulus evokes a sense of recognition or novelty (Rissman et al., 2016, 2010; Uncapher et al., 2015). Therefore, this field of study can be helpful in some specific cases. For instance, one potentially crucial point in evaluating the guilt of an individual suspected of a crime is to determine whether he/she had been in a particular place before (e.g., the crime scene). In Chapter 5 the possibility of predicting whether participants had visited some specific rooms before, has been investigated. In particular, the aim was to determine from the analysis of brain activity whether subjects were looking at pictures showing rooms they had previously visited. A searchlight analysis was performed highlighting which brain areas encoded the most useful information for classifying the rooms as previously experienced or not. The results have proven accurate within-subjects predictions of whether participants were viewing pictures of known rooms. One strength of this experiment is the fact that the good classification performance has been reached using pictures not taken from participants themselves, as in the recent study by Rissman and co-workers (2016), but recorder from different view points. This indicates the robustness as well as the potential applicability of brain reading.

Independently from the guilt condition, another crucial point in a forensic investigation could be to determine whether the accused has memory of the crime (e.g., the homicide), given that an individual having nothing to do with the crime cannot remember it. In order to test a possible application of a MVPA-based memory detector in such a situation, in Chapter 6 a specific memory for a negative emotional event involving the death of a known person, was investigated. The aim of the study was to test whether it is possible to classify sentences as describing a truly happened episode or not by analyzing the pattern of brain activity elicited while reading each sentence. Importantly, the possibility of classifying each single item (i.e., sentence) as a memory or not has been tested, reaching a good level of classification accuracy on a single-subject basis. Moreover, in a group-level analysis, the most informative areas for the classification between memories and non-memories have been investigated, despite the small sample size.

Taken together with the previous findings (e.g., Rissman et al., 2016), the results of these two experiments suggest that the analysis of brain scans for memory detection can potentially be a useful tool for forensic uses. However, the present results should not be assumed to demonstrate an immediate practical applicability of these MVPA-based memory detection techniques in forensic contexts. Indeed, there are many aspects that need to be systematically investigated in further studies. For example, the participants in the two experiments were not asked or incentivized to intentionally alter their performance, and thus they were not motivated in trying to fake the classifier. By contrast, in real-life judicial and forensic situations, the motivation in lying, malingering, and not cooperating is high. Furthermore, some individuals could be tested on genuine false memories, and thus could unintentionally fake the classifier. Whether false memories can be distinguished from real memories is one of the main issues that should be addressed in the near future.

Therefore, at the moment, the use of these techniques in real-world forensic settings is not plausible, due to several limitations, and to the claimed possibility of using covert countermeasures to alter classification results (e.g., Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011) or the possibility to willfully manipulate mental states under incentives (Uncapher et al., 2015). Indeed, when focusing on single specific memories (in particular in a forensic setting) a subject could have advantages in voluntarily trying to conceal that information, so that he could try to perform some different covert cognitive activity (e.g., counting backward) in order to alter the brain signal related to the recognition of target stimuli. To date, the susceptibility of fMRI-based memory detection to this kind of alteration has been claimed, but methods for overcoming it have not been systematically discussed, yet. Moreover, from the legal perspective the analysis of covert brain responses for determining whether a subject possess a specific memory can have some ethical implications, such as issues related to rights against self-incrimination. This topic and related issues are dealt with by the recently emerged field of neuroethics, which aims to address ethical and legal issues related to the relationship between neuroscience and law. However, this emerging approach has shown to be potentially useful in future forensic applications, once some methodological, technical and ethical issues have been addressed (Haynes & Rees, 2006). Despite these important issues that limit real-world applications of these fMRI-based techniques, what has to be taken into account is that the studies presented in this thesis, together with other recent studies, have shown the possibility of tailoring the neuroimaging investigation to the individual personal memories, and that accurate performances can also be achieved at single-subject level.

In conclusion, Multi-Voxel Pattern Analysis (MVPA) has changed the way of thinking about how neuroimaging data can address research questions. In addition to providing a more powerful and sensitive approach, these methods have impacted on the kind of research questions that can be asked to neuroimaging data, in a way that seems more in line with the main questions in cognitive neuroscience. The use of multivariate techniques can lead to a deeper comprehension of how representations are spatially distributed in the brain, as well as the link between brain and behavior.

Going beyond the specific results obtained in the four studies presented in this thesis, the main point is the ductility of MVPA methods in answering to different research questions in a reliable way.

Taken together, these findings contribute to show the wide range of potential MVPA applications in the neuroimaging research. Moreover, thanks to the possibility of making predictions at single-trial or single-subject level, these methods can potentially have practical applications in real-life settings. Pattern-based multivariate analyses could be used to decode covert information in the brain that is voluntarily concealed by an individual, or to identify preclinical stages of neurodegenerative diseases. Therefore, it could be possible to decode covert specific mental states or general neurological conditions without the need for collaboration. This could lead to ethical issues (e.g., forensic use), but also to considerable improvements in clinical practice.

References

Addis, D. R., Barense, M., & Duarte, A. (2015). *The Wiley handbook on the cognitive neuroscience of memory*. John Wiley & Sons.

Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: A partial least squares approach. *Neuroimage*, *23*(4), 1460–1471. http://doi.org/10.1016/j.neuroimage.2004.08.007

Addis, D. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus*, *14*(6), 752–762. http://doi.org/10.1002/hipo.10215

Allefeld, C., & Haynes, J. D. (2014). Searchlight-based multi-voxel pattern analysis of fMRI by cross-validated MANOVA. *NeuroImage*, *89*, 345–357. http://doi.org/10.1016/j.neuroimage.2013.11.043

Aminoff, E. M., Kveraga, K., & Bar, M. (2013). The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2013.06.009

Arenaza-Urquijo, E. M., Landeau, B., La Joie, R., Mevel, K., Mézenge, F., Perrotin, A., … Chételat, G. (2013). Relationships between years of education and gray matter volume, metabolism and functional connectivity in healthy elders. *NeuroImage*, *83*, 450–457. http://doi.org/10.1016/j.neuroimage.2013.06.053

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, *38*(1), 95–113. http://doi.org/10.1016/j.neuroimage.2007.07.007

Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry—The Methods. *NeuroImage*, *11*(6), 805–821. http://doi.org/10.1006/nimg.2000.0582

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2007). A common system for the comprehension and production of narrative speech. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(43), 11455–11464. http://doi.org/10.1523/JNEUROSCI.5257-06.2007

Aydin, K., Ucar, a, Oguz, K. K., Okur, O. O., Agayev, a, Unal, Z., … Ozturk, C. (2007). Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study. *AJNR. American Journal of Neuroradiology*, *28*(10), 1859–64.

http://doi.org/10.3174/ajnr.A0696

Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, *38*(2), 347–58. http://doi.org/S0896627303001673 [pii]

Bartrés-Faz, D., & Arenaza-Urquijo, E. M. (2011). Structural and functional imaging correlates of cognitive and brain reserve hypotheses in healthy and pathological aging. *Brain Topography*, *24*(3–4), 340–357. http://doi.org/10.1007/s10548-011-0195-9

Bartrés-Faz, D., Solé-Padullés, C., Junqué, C., Rami, L., Bosch, B., Bargalló, N., … Molinuevo, J. L. (2009). Interactions of cognitive reserve with regional brain anatomy and brain function during a working memory task in healthy elders. *Biological Psychology*, *80*(2), 256–259. http://doi.org/10.1016/j.biopsycho.2008.10.005

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: Emerging concepts in cognitive reserve. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2013.08.012

Beauregard, M., Lévesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(18), RC165. http://doi.org/20015619 [pii]

Bennett, D. A., Wilson, R. S., Schneider, J. A., Evans, D. A., Mendes de Leon, C. F., Arnold, S. E., … Bienias, J. L. (2003). Education modifies the relation of AD pathology to level of cognitive function in older persons. *Neurology*, *60*, 1909–1915. http://doi.org/10.1212/01.WNL.0000069923.64550.9F

Bles, M., & Haynes, J.-D. (2008). Detecting concealed information using brain-imaging technology. *Neurocase*, *14*(1), 82–92. http://doi.org/10.1080/13554790801992784

Bosch, B., Bartrés-Faz, D., Rami, L., Arenaza-Urquijo, E. M., Fernàndez-Espejo, D., Junqué, C., … Molinuevo, J. L. (2010). Cognitive reserve modulates task-induced activations and deactivations in healthy elders, amnestic mild cognitive impairment and mild Alzheimer's disease. *Cortex*, *46*(4), 451–461. http://doi.org/10.1016/j.cortex.2009.05.006

Bousquet, O., Boucheron, S., & Lugosi, G. (2004). Introduction to Statistical Learning Theory. In *Advanced Lectures on Machine Learning* (Vol. 3176, pp. 169–207).

http://doi.org/10.1007/978-3-540-28650-9_8

Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., & May, A. (2008). Training-induced brain structure changes in the elderly. *J Neurosci*, *28*(28), 7031–7035. http://doi.org/10.1523/JNEUROSCI.0742-08.2008

Bray, S., Chang, C., & Hoeft, F. (2009). Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. *Frontiers in Human Neuroscience*, *3*(October), 32. http://doi.org/10.3389/neuro.09.032.2009

Brayne, C., Ince, P. G., Keage, H. A. D., McKeith, I. G., Matthews, F. E., Polvikoski, T., & Sulkava, R. (2010). Education, the brain and dementia: Neuroprotection or compensation? *Brain*, *133*(8), 2210–2216. http://doi.org/10.1093/brain/awq185

Brown, J., Cooper-Kuhn, C. M., Kempermann, G., Van Praag, H., Winkler, J., Gage, F. H., & Kuhn, H. G. (2003). Enriched environment and physical activity stimulate hippocampal but not olfactory bulb neurogenesis. *European Journal of Neuroscience*, *17*(10), 2042–2046. http://doi.org/10.1046/j.1460-9568.2003.02647.x

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nature Reviews. Neuroscience*, *2*(1), 51–61. http://doi.org/10.1038/35049064

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin*, *133*(5), 761–79. http://doi.org/10.1037/0033-2909.133.5.761

Buckner, R. L., Head, D., Parker, J., Fotenos, A. F., Marcus, D., Morris, J. C., & Snyder, A. Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: Reliability and validation against manual measurement of total intracranial volume. *NeuroImage*, *23*(2), 724–738. http://doi.org/10.1016/j.neuroimage.2004.06.018

Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., … Mintun, M. A. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: Evidence for a relationship between default activity, amyloid, and memory. *Journal of Neuroscience*, *25*(34), 7709–7717. http://doi.org/10.1523/jneurosci.2177-05.2005

Burgess, N., Becker, S., King, J. A., & O'Keefe, J. (2001). Memory for events and their spatial context: models and experiments. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1413), 1493–503. http://doi.org/10.1098/rstb.2001.0948

Burianova, H., & Grady, C. L. (2007). Common and unique neural activations in autobiographical, episodic, and semantic retrieval. *Journal of Cognitive Neuroscience*, *19*(9), 1520–34. http://doi.org/10.1162/jocn.2007.19.9.1520

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review*, *114*(2), 340–375. http://doi.org/10.1037/0033-295X.114.2.340

Cabeza, R., & Nyberg, L. (2000). Imaging cognition: An empirical review of PET studies with normal subjects. *J Cogn Neurosci*, *9*(1), 1–26.

Cabeza, R., Prince, S. E., Daselaar, S. M., Greenberg, D. L., Budde, M., Dolcos, F., … Rubin, D. C. (2004). Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm. *Journal of Cognitive Neuroscience*, *16*(9), 1583–1594. http://doi.org/10.1162/0898929042568578

Cahill, L., Haier, R. J., White, N. S., Fallon, J., Kilpatrick, L., Lawrence, C., … Alkire, M. T. (2001). Sex-related difference in amygdala activity during emotionally influenced memory storage. *Neurobiology of Learning and Memory*, *75*(1), 1–9. http://doi.org/10.1006/nlme.2000.3999

Calhoun, V. D., Adali, T., Hansen, L. K., Larsen, J., & Pekar, J. J. (2003). ICA of Functional MRI Data: An Overview. In *4th International Symposium on Independent Component Analysis and Blind Signal Separation* (pp. 281–288). http://doi.org/10.1.1.3.7473

Canli, T., & Amin, Z. (2002). Neuroimaging of emotion and personality: Scientific evidence and ethical considerations. *Brain and Cognition*, *50*(3), 414–431. http://doi.org/10.1016/S0278-2626(02)00517-1

Carlson, T. a, Schrater, P., & He, S. (2003). Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience*, *15*(5), 704–17. http://doi.org/10.1162/089892903322307429

Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, *461*(7266), 983–986. http://doi.org/10.1038/nature08461

Carroll, M. K., Cecchi, G. A., Rish, I., Garg, R., & Rao, A. R. (2009). Prediction and interpretation of distributed neural activity with sparse models. *NeuroImage*, *44*(1), 112–122. http://doi.org/10.1016/j.neuroimage.2008.08.020

Casanova, R., Hsu, F. C., & Espeland, M. A. (2012). Classification of Structural MRI Images in Alzheimer's Disease from the Perspective of Ill-Posed Problems. *PLoS ONE*, *7*(10). http://doi.org/10.1371/journal.pone.0044877

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*. http://doi.org/10.1093/brain/awl004

Chang, C., & Lin, C. (2011). LIBSVM : A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*, 1–39. http://doi.org/10.1145/1961189.1961199

Childress, A. R., Mozley, P. D., McElgin, W., Fitzgerald, J., Reivich, M., & O'Brien, C. P. (1999). Limbic activation during cue-induced cocaine craving. *American Journal of Psychiatry*, *156*(1), 11–18. http://doi.org/10.1046/j.1523-1747.2000.00951.x

Ciaramelli, E., Grady, C. L., & Moscovitch, M. (2008). Top-down and bottom-up attention to memory: A hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia*. http://doi.org/10.1016/j.neuropsychologia.2008.03.022

Coffey, C. E., Saxton, J. A., Ratcliff, G., Bryan, R. N., & Lucke, J. F. (1999). Relation of education to brain size in normal aging: implications for the reserve hypothesis. *Neurology*, *53*(1), 189–196. http://doi.org/10.1212/WNL.54.5.1207

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., … Kramer, A. F. (2006). Aerobic Exercise Training Increases Brain Volume in Aging Humans. *Journal of Gerontology*, *61A*(11), 5.

Conway, M. A. (1990). Autobiographical memory: An introduction. *Autobiographical Memory: An Introduction.*

Conway, M. A. (1992). A structural model of autobiographical memory. In *Theoretical perspectives on autobiographical memory* (pp. 167–193). Springer.

Conway, M. A., Pleydell-Pearce, C. W., Whitecross, S. E., & Sharpe, H. (2003). Neurophysiological correlates of memory for experienced and imagined events. *Neuropsychologia*, *41*(3), 334–340. http://doi.org/10.1016/S0028-3932(02)00165-3

Conway, M. A., Singer, J. A., & Tagini, A. (2004). the Self and Autobiographical Memory: Correspondence and Coherence. *Social Cognition*, *22*(5), 491–529. http://doi.org/10.1521/soco.22.5.491.50768

Conway, M. A., Turk, D. J., Miller, S. L., Logan, J., Nebes, R. D., Meltzer, C. C., & Becker, J. T. (1999). A positron emission tomography (PET) study of autobiographical memory retrieval. *Memory*, *7*, 679–702. http://doi.org/10.1080/096582199387805

Conway, M., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. http://doi.org/10.1037//0033-295X

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*(3), 273–297. http://doi.org/10.1023/A:1022627411411

Costafreda, S. G., Chu, C., Ashburner, J., & Fu, C. H. Y. (2009). Prognostic and diagnostic potential of the structural neuroanatomy of depression. *PLoS ONE*, *4*(7). http://doi.org/10.1371/journal.pone.0006353

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cognitive, Affective & Behavioral Neuroscience*, *13*(3), 667–73. http://doi.org/10.3758/s13415-013-0186-2

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, *19*(2), 261–270. http://doi.org/10.1016/S1053-8119(03)00049-1

Crovitz, H. F., & Sciffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, *4*(5B), 517–518. http://doi.org/10.3758/BF03334277

Cummins, R. A., Walsh, R. N., Budtz-Olsen, O. E., Kostantinos, T., & Horsfall, C. R. (1973). Environmentally-induced changes in the brains of elderly rats. *Nature*, *243*(5409), 516–518. http://doi.org/10.1038/243516a0

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, *15*(12), 806–813. http://doi.org/10.1111/j.0956-7976.2004.00760.x

D'Argembeau, A., Comblain, C., & van der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Applied Cognitive Psychology*, *17*(3), 281–294. http://doi.org/10.1002/acp.856

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 2157–62. http://doi.org/10.1073/pnas.0337195100

Davatzikos, C., Fan, Y., Wu, X., Shen, D., & Resnick, S. M. (2009). Detection of prodromal Alzheimer's disease via pattern classification of MRI. *Neurobiology of Aging*, *29*(4), 514–523. http://doi.org/10.1016/j.neurobiolaging.2006.11.010.Detection

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., … Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, *28*(3), 663–668. http://doi.org/10.1016/j.neuroimage.2005.08.009

Davis, T., & Poldrack, R. A. (2013). Measuring neural representations with fMRI: Practices and pitfalls. *Annals of the New York Academy of Sciences*, *1296*(1), 108–134. http://doi.org/10.1111/nyas.12156

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., & Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *NeuroImage*, *43*(1), 44–58. http://doi.org/10.1016/j.neuroimage.2008.06.037

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. U. S. A.*, *93*(24), 13494–13499. http://doi.org/10.1073/pnas.93.24.13494

Di Bono, M. G. (2009). *Beyond Mind Reading : Advanced Machine Learning Techniques* *for fMRI Data Analysis. PhD thesis. Department of Developmental Psychology and Socialization, University of Padua (Italy).*

Di Bono, M. G., & Zorzi, M. (2008). Decoding Cognitive States from fMRI Data Using Support Vector Regression. *Developmental Psychology*, *6*(2), 189–201.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron*, *73*(3), 415–434. http://doi.org/10.1016/j.neuron.2012.01.010

Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. WW Norton & Company.

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., & May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning 20. *J Neurosci.*, *26*(1529–2401 (Electronic)), 6314–6317.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern classification. 2nd. *Edition. New York*.

Dufouil, C., Alpérovitch, a, & Tzourio, C. (2003). Influence of education on the relationship between white matter lesions and cognition. *Neurology*, *60*(5), 831–6. http://doi.org/10.1212/01.WNL.0000049456.33231.96

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap. Monographs on Statistics and Applied Probability, No. 57. Chapman and Hall, London, 436 p. Monographs on Statistics and Applied Probability* (Vol. 57). http://doi.org/10.1111/1467-9639.00050

Eichenbaum, H., & Cohen, N. J. (2004). *From conditioning to conscious recollection: Memory systems of the brain*. Oxford University Press on Demand.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–52. http://doi.org/10.1146/annurev.neuro.30.051606.094328

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 201602413. http://doi.org/10.1073/pnas.1602413113

Engvig, A., Fjell, A. M., Westlye, L. T., Moberget, T., Sundseth, O., Larsen, V. A., &

Walhovd, K. B. (2010). Effects of memory training on cortical thickness in the elderly. *NeuroImage*, *52*(4), 1667–1676. http://doi.org/10.1016/j.neuroimage.2010.05.041

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2008.07.004

Epstein, R. A., & Higgins, J. S. (2007). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex*, *17*(7), 1680–1693. http://doi.org/10.1093/cercor/bhl079

Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron*, *37*(5), 865–876. http://doi.org/10.1016/S0896-6273(03)00117-X

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, *23*(1), 115–125. http://doi.org/10.1016/S0896-6273(00)80758-8

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. http://doi.org/10.1038/33402

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., … Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3017–22. http://doi.org/10.1073/pnas.1015950108

Etzel, J. A. (2015). MVPA Permutation Schemes: Permutation Testing for the Group Level. In *Proceedings - 2015 International Workshop on Pattern Recognition in NeuroImaging, PRNI 2015* (pp. 65–68). http://doi.org/10.1109/PRNI.2015.29

Etzel, J. A., & Braver, T. S. (2013). MVPA permutation schemes: Permutation testing in the land of cross-validation. In *Proceedings - 2013 3rd International Workshop on Pattern Recognition in Neuroimaging, PRNI 2013* (pp. 140–143). http://doi.org/10.1109/PRNI.2013.44

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2013.03.041

Fay, T. B., Yeates, K. O., Taylor, H. G., Bangert, B., Dietrich, A., Nuss, K. E., … Wright, M. (2010). Cognitive reserve as a moderator of postconcussive symptoms in children with complicated and uncomplicated mild traumatic brain injury. *Journal of the International Neuropsychological Society*, *16*(1), 94–105. http://doi.org/10.1017/S1355617709991007

Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2005). Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. *Visual Cognition*, *12*(6), 893–924. http://doi.org/10.1080/13506280444000571

Fitzpatrick, A. L., Kuller, L. H., Ives, D. G., Lopez, O. L., Jagust, W., Breitner, J. C. S., … Dulberg, C. (2004). Incidence and Prevalence of Dementia in the Cardiovascular Health Study. *Journal of the American Geriatrics Society*, *52*(2), 195–204. http://doi.org/10.1111/j.1532-5415.2004.52058.x

Fivush, R. (2008). The Development of Autobiographical Memory. In *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 55–68). http://doi.org/10.1002/9780470713679.ch4

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. http://doi.org/10.1016/0022-3956(75)90026-6

Ford, J. H., Addis, D. R., & Giovanello, K. S. (2011). Differential neural activity during search of specific and general autobiographical memories elicited by musical cues. *Neuropsychologia*, *49*(9), 2514–2526. http://doi.org/10.1016/j.neuropsychologia.2011.04.032

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology*, *64*(6), 1032–1039. http://doi.org/10.1212/01.WNL.0000154530.72969.11

Foubert-Samier, A., Catheline, G., Amieva, H., Dilharreguy, B., Helmer, C., Allard, M., & Dartigues, J. F. (2012). Education, occupation, leisure activities, and brain reserve: A population-based study. *Neurobiology of Aging*, *33*(2), 423.e15-423.e25. http://doi.org/10.1016/j.neurobiolaging.2010.09.023

Fratiglioni, L., & Wang, H.-X. (2007). Brain Reserve Hypothesis in Dementia. *Journal of Alzheimer's Disease*, *12*(September 2007), 11–22.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1). Springer series in statistics Springer, Berlin.

Friston, K. J., Holmes, a. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. http://doi.org/10.1002/hbm.460020402

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage*, *55*(1), 312–319. http://doi.org/10.1016/j.neuroimage.2010.11.025

Gardner, J. L. (2010). Is cortical vasculature functionally organized? *NeuroImage*, *49*(3), 1953–1956. http://doi.org/10.1016/j.neuroimage.2009.07.004

Garibotto, V., Borroni, B., Kalbe, E., Herholz, K., Salmon, E., Holtoff, V., … Perani, D. (2008). Education and occupation as proxies for reserve in aMCI converters and AD: FDG-PET evidence. *Neurology*, *71*(17), 1342–1349. http://doi.org/10.1212/01.wnl.0000327670.62378.c0

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245. http://doi.org/23/27/9240 [pii]

Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J., & Moscovitch, M. (2004). Remembering our past: Functional neuroanatomy of recollection of recent and very remote personal events. *Cerebral Cortex*, *14*(11), 1214–1225. http://doi.org/10.1093/cercor/bhh082

Glisky, E. L. (2007). Chapter 1 Changes in Cognitive Function in Human Aging. *Brain Aging: Models, Methods, and Mechanisms.*, 1–11.

Golland, P., & Fischl, B. (2003). Permutation tests for classification: towards statistical significance in image-based studies. *Information Processing in Medical Imaging : Proceedings of the ... Conference*, *18*(August), 330–341. http://doi.org/10.1007/978-3-540-45087-0_28

Golland, P., Liang, F., Mukherjee, S., & Panchenko, D. (2000). Permutation tests for classification. *Learning Theory*, *1*(August), 0–21.

http://doi.org/10.1007/11503415_34

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*(1), 21–36. http://doi.org/10.1006/nimg.2001.0786

Good, P. (2013). *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media.

Gothelf, D., Hoeft, F., Ueno, T., Sugiura, L., Lee, A. D., Thompson, P., & Reiss, A. L. (2011). Developmental changes in multivariate neuroanatomical patterns that predict risk for psychosis in 22q11.2 deletion syndrome. *Journal of Psychiatric Research*, *45*(3), 322–331. http://doi.org/10.1016/j.jpsychires.2010.07.008

Greenwood, P. M., & Parasuraman, R. (2010). Neuronal and cognitive plasticity: A neurocognitive framework for ameliorating cognitive aging. *Frontiers in Aging Neuroscience*, *2*. http://doi.org/10.3389/fnagi.2010.00150

Grosenick, L., Greer, S., & Knutson, B. (2008). Interpretable classifiers for FMRI improve prediction of purchases. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *16*(6), 539–538. http://doi.org/10.1109/TNSRE.2008.926701

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)*, *3*(3), 1157–1182. http://doi.org/10.1016/j.aca.2011.07.027

Habermas, T., & Bluck, S. (2000). Getting a life: the emergence of the life story in adolescence. *Psychological Bulletin*, *126*(5), 748–769. http://doi.org/10.1037/0033-2909.126.5.748

Haller, S., Lovblad, K. O., Giannakopoulos, P., & Van De Ville, D. (2014). Multivariate pattern recognition for diagnosis and prognosis in clinical neuroimaging: State of the Art, current challenges and future trends. *Brain Topography*, *27*(3), 329–337. http://doi.org/10.1007/s10548-014-0360-z

Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, *21*(4), 1639–1651. http://doi.org/10.1016/j.neuroimage.2003.11.029

Hansen, L. K., Larsen, J., Nielsen, F. a, Strother, S. C., Rostrup, E., Savoy, R., … Paulson, O. B. (1999). Generalizable patterns in neuroimaging: how many principal components? *NeuroImage*, *9*, 534–544. http://doi.org/10.1006/nimg.1998.0425

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proc Natl Acad Sci U S A*, *107*(38), 16494–16499. http://doi.org/10.1073/pnas.1008121107

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(September), 2425–2430. http://doi.org/10.1126/science.1063736

Hayasaka, S., & Nichols, T. E. (2003). Validating cluster size inference: Random field and permutation methods. *NeuroImage*, *20*(4), 2343–2356. http://doi.org/10.1016/j.neuroimage.2003.08.003

Hayasaka, S., & Nichols, T. E. (2004). Combining voxel intensity and cluster extent with permutation test framework. *NeuroImage*, *23*(1), 54–63. http://doi.org/10.1016/j.neuroimage.2004.04.035

Haynes, J. D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron*, *87*(2), 257–270. http://doi.org/10.1016/j.neuron.2015.05.025

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature reviews.Neuroscience*, *7*(7), 523–534. http://doi.org/10.1038/nrn1931

Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading Hidden Intentions in the Human Brain. *Current Biology*, *17*(4), 323–328. http://doi.org/10.1016/j.cub.2006.11.072

Hebart, M. N., Görgen, K., Haynes, J.-D., & Dubois, J. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, *8*(January), 1–18. http://doi.org/10.3389/fninf.2014.00088

Heekeren, H. R., Marrett, S., Ruff, D. A., Bandettini, P. A., & Ungerleider, L. G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(26), 10023–8. http://doi.org/10.1073/pnas.0603949103

Henson, R. (2006). Forward inference using functional neuroimaging: Dissociations versus associations. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2005.12.005

Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270. http://doi.org/10.1016/S0028-3932(02)00159-8

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *J Neurosci*, *19*(10), 3962–3972.

Hoeft, F., Walter, E., Lightbody, A. A., Hazlett, H. C., Chang, C., Piven, J., & Reiss, A. L. (2011). Neuroanatomical Differences in Toddler Boys With Fragile X Syndrome and Idiopathic Autism. *Archives of General Psychiatry*, *68*(3), 295. http://doi.org/10.1001/archgenpsychiatry.2010.153

Holland, A. C., & Kensinger, E. A. (2010). Emotion and autobiographical memory. *Physics of Life Reviews*. http://doi.org/10.1016/j.plrev.2010.01.006

Ince, P. G. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. *Lancet*, *357*(9251), 169–175. http://doi.org/10.1016/S0140-6736(00)03589-3

Isenberg, N., Silbersweig, D., Engelien, a, Emmerich, S., Malavade, K., Beattie, B., … Stern, E. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(18), 10456–10459. http://doi.org/10.1073/pnas.96.18.10456

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 4–37. http://doi.org/10.1109/34.824819

Janzen, G., & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, *7*(6), 673–677. http://doi.org/10.1038/nn1257

Jimura, K., & Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia*, *50*(4), 544–552. http://doi.org/10.1016/j.neuropsychologia.2011.11.007

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685. http://doi.org/10.1038/nn1444

Katzman, R. (1993). Education and the prevalence of dementia and Alzheimer's disease. *Neurology*, *43*(1), 13–20. http://doi.org/10.1212/WNL.43.1_Part_1.13

Katzman, R., Aronson, M., Fuld, P., Kawas, C., Brown, T., Morgenstern, H., … Ooi, W. L. (1989). Development of dementing illnesses in an 80-year-old volunteer cohort. *Annals of Neurology*, *25*(4), 317–324.

Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., … David, A. S. (2000). Towards a functional neuroanatomy of self processing: Effects of faces and words. *Cognitive Brain Research*, *10*(1–2), 133–144. http://doi.org/10.1016/S0926-6410(00)00036-7

Kirwan, C. B., Bayley, P. J., Galván, V. V, & Squire, L. R. (2008). Detailed recollection of remote autobiographical memory after damage to the medial temporal lobe. *PNAS*, *105*(7), 2676–2680. http://doi.org/10.1073/pnas.0712155105

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection 2 Methods for Accuracy Estimation. *Proc. of IJCAI'95*, 1137–1145.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 3863–3868. http://doi.org/10.1073/pnas.0600244103

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage*, *26*(2), 317–329. http://doi.org/10.1016/j.neuroimage.2005.01.048

Larsson, J., & Smith, A. T. (2012). FMRI repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex*. http://doi.org/10.1093/cercor/bhr119

Lemm, S., Blankertz, B., Dickhaus, T., Muller, K. R., & Müller, K.-R. (2011). Introduction to machine learning for brain imaging. *NeuroImage*, *56*(2), 387–399. http://doi.org/10.1016/j.neuroimage.2010.11.004

León, I., García-García, J., & Roldán-Tapia, L. (2014). Estimating cognitive reserve in healthy adults using the Cognitive Reserve Scale. *PloS One*, *9*(7), e102632. http://doi.org/10.1371/journal.pone.0102632

Leon, I., Garcia, J., & Roldan-Tapia, L. (2011). Development of the scale of cognitive reserve in Spanish population: a pilot study. *Rev Neurol*, *52*(11), 653–660.

Letenneur, L., Gilleron, V., Commenges, D., Helmer, C., Orgogozo, J. M., & Dartigues, J. F. (1999). Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project. *Journal of Neurology, Neurosurgery, and Psychiatry*, *66*(2), 177–183. http://doi.org/10.1136/jnnp.66.2.177

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9596–601. http://doi.org/10.1073/pnas.092277599

Liu, F., Guo, W., Yu, D., Gao, Q., Gao, K., Xue, Z., … Chen, H. (2012). Classification of different therapeutic responses of major depressive disorder with multivariate pattern analysis method based on structural MR scans. *PLoS ONE*, *7*(7). http://doi.org/10.1371/journal.pone.0040968

Liu, Y., Julkunen, V., Paajanen, T., Westman, E., Wahlund, L. O., Aitken, A., … Soininen, H. (2012). Education increases reserve against Alzheimer's disease-Evidence from structural MRI analysis. *Neuroradiology*, *54*(9), 929–938. http://doi.org/10.1007/s00234-012-1005-0

Logothetis, N. K., & Wandell, B. a. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, *66*, 735–769. http://doi.org/10.1146/annurev.physiol.66.082602.092845

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, *136*(4), 659–676. http://doi.org/10.1037/a0020080

Maguire, E. A. (2001a). Neuroimaging studies of autobiographical event memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1413), 1441–51. http://doi.org/10.1098/rstb.2001.0944

Maguire, E. A. (2001b). The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*(3), 225–238. http://doi.org/10.1111/1467-9450.00233

Maguire, E. a, Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4398–403. http://doi.org/10.1073/pnas.070039597

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. http://doi.org/10.1037/0033-295X.87.3.252

Manly, J. J., Byrd, D. a, Touradji, P., & Stern, Y. (2004). Acculturation, Reading Level, and Neuropsychological Test Performance Among African American Elders. *Applied Neuropsychology*, *11*(1), 37–46. http://doi.org/10.1207/s15324826an1101_5

Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience*, *17*(11), 1598–1606. http://doi.org/10.1038/nn.3834

Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2015). Outside Looking In: Landmark Generalization in the Human Navigational System. *The Journal of Neuroscience*, *35*(44), 14896–908. http://doi.org/10.1523/JNEUROSCI.2270-15.2015

Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, *19*(9), 1498–1507. http://doi.org/10.1162/jocn.2007.19.9.1498

Markowitsch, H. J., Vandekerckhove, M. M. P., Lanfermann, H., & Russ, M. O. (2003). Engagement of lateral and medial prefrontal areas in the ecphory of sad and happy autobiographical memories. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *39*(4–5), 643–665. http://doi.org/10.1016/S0010-9452(08)70858-X

Masaoka, Y., Sugiyama, H., Katayama, A., Kashiwagi, M., & Homma, I. (2012). Slow breathing and emotions associated with odor-induced autobiographical memories.

Chemical Senses, *37*(4), 379–388. http://doi.org/10.1093/chemse/bjr120

McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, *44*(2), 379–387. http://doi.org/10.1016/j.neuron.2004.09.019

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, *47*(11), 2290–2298. http://doi.org/10.1016/j.neuropsychologia.2008.12.025

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: methods and applications. *Current Medical Imaging Reviews*, *1*(2), 105–113. http://doi.org/Doi 10.2174/1573405054038726

Meegan, D. V. (2008). Neuroimaging Techniques for Memory Detection: Scientific, Ethical, and Legal Issues. *American Journal of Bioethics*, *8*(1), 9–20. http://doi.org/10.1080/15265160701842007

Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the cognitive reserve hypothesis: A systematic review with meta-analyses and qualitative analyses. *PLoS ONE*, *7*(6). http://doi.org/10.1371/journal.pone.0038268

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *NeuroImage*, *53*(1), 103–118. http://doi.org/10.1016/j.neuroimage.2010.05.051

Moll, J., de Oliveira-Souza, R., Bramati, I. E., & Grafman, J. (2002). Functional Networks in Emotional Moral and Nonmoral Social Judgments. *NeuroImage*, *16*(3), 696–703. http://doi.org/10.1006/nimg.2002.1118

Mondini, S., Madella, I., Zangrossi, A., Bigolin, A., Tomasi, C., Michieletto, M., … Mapelli, D. (2016). Cognitive Reserve in Dementia: Implications for Cognitive Training. *Frontiers in Aging Neuroscience*, *8*(April), 84. http://doi.org/10.3389/fnagi.2016.00084

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*, *43*(11), 2412–2414. http://doi.org/10.1212/WNL.43.11.2412-a

Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., & Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Archives of Neurology*, *58*(3), 397–405. http://doi.org/10.1001/archneur.58.3.397

Mourão-Miranda, J., Bokde, A. L. W., Born, C., Hampel, H., & Stetter, M. (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. *NeuroImage*, *28*(4), 980–995. http://doi.org/10.1016/j.neuroimage.2005.06.070

Mugler, J. P., & Brookeman, J. R. (1990). Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magnetic Resonance in Medicine*, *15*(1), 152–157. http://doi.org/10.1002/mrm.1910150117

Muscatell, K. A., Addis, D. R., & Kensinger, E. A. (2009). Self-involvement modulates the effective connectivity of the autobiographical memory network. *Social Cognitive and Affective Neuroscience*, *5*(1), 68–76. http://doi.org/10.1093/scan/nsp043

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. http://doi.org/10.1002/hbm.1058

Nichols, T., & Holmes, A. (2003). Nonparametric Permutation Tests for Functional Neuroimaging. *Human Brain Function: Second Edition*, *25*(August 1999), 887–910. http://doi.org/10.1016/B978-012264841-0/50048-2

Noirhomme, Q., Lesenfants, D., Gomez, F., Soddu, A., Schrouff, J., Garraux, G., … Laureys, S. (2014). Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions. *NeuroImage: Clinical*, *4*, 687–694. http://doi.org/10.1016/j.nicl.2014.04.004

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2006.07.005

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain-A meta-analysis of imaging studies on the self. *NeuroImage*, *31*(1), 440–457. http://doi.org/10.1016/j.neuroimage.2005.12.002

Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging Clinical and*

Experimental Research, *24*(3), 218–26. http://doi.org/10.3275/7800

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, *16*(5), 292–305. http://doi.org/10.1016/j.tics.2012.04.005

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–23. http://doi.org/10.1162/08989290051137549

O'Toole, A. J., Jiang, F., Abdi, H., Pénard, N., Dunlop, J. P., & Parent, M. A. (2007). Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. *Journal of Cognitive Neuroscience*, *19*(11), 1735–1752.

Oberhuber, M., Hope, T. M. H., Seghier, M. L., Jones, O. P., Prejawa, S., Green, D. W., & Price, C. J. (2016). Four Functionally Distinct Regions in the Left Supramarginal Gyrus Support Word Processing. *Cerebral Cortex*.

Oddo, S., Lux, S., Weiss, P., Schwab, A., Welzer, H., Markowitsch, H., & Fink, G. (2008). Specific role of medial prefrontal cortex in retrieving recent autobiographical memories: An fMRI study of young female subjects. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*.

OED. (2016). Oxford English Dictionary Online. Retrieved from http://dictionary.oed.com/

Ogawa, S., Lee, T.-M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Ojala, M., & Garriga, G. C. (2009). Permutation tests for studying classifier performance. In *Proceedings - IEEE International Conference on Data Mining, ICDM* (pp. 908–913). http://doi.org/10.1109/ICDM.2009.108

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience and Biobehavioral Reviews*. http://doi.org/10.1016/j.neubiorev.2012.01.004

Osaka, M., Yaoi, K., Minamoto, T., & Osaka, N. (2013). When do negative and positive emotions modulate working memory performance? *Scientific Reports*, *3*, 1375. http://doi.org/10.1038/srep01375

Ott, A., Breteler, M. M., van Harskamp, F., Claus, J. J., van der Cammen, T. J., Grobbee, D. E., & Hofman, A. (1995). Prevalence of Alzheimer's disease and vascular dementia: association with education. The Rotterdam study. *BMJ (Clinical Research Ed.)*, *310*(6985), 970–3. http://doi.org/10.1136/bmj.310.6985.970

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, *60*, 173–196. http://doi.org/10.1146/annurev.psych.59.103006.093656

Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Annals of the New York Academy of Sciences*, *930*(1), 315–29. http://doi.org/10.1111/j.1749-6632.2001.tb05741.x

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annu Rev Neurosci*, *28*, 377–401. http://doi.org/10.1146/annurev.neuro.27.070203.144216

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Academic press.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2008.11.007

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–8. http://doi.org/10.1001/archneur.56.3.303

Phan, K. L., Wager, T. D., Taylor, S. F., & Liberzon, I. (2004). Functional neuroimaging studies of human emotions. *CNS Spectrums*, *9*(4), 258–266. http://doi.org/10.1017/S1092852900009196

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*(2), 331–48. http://doi.org/10.1006/nimg.2002.1087

Piefke, M., & Fink, G. R. (2005). Recollections of one's own past: The effects of aging and gender on the neural mechanisms of episodic autobiographical memory. In

Anatomy and Embryology (Vol. 210, pp. 497–512). http://doi.org/10.1007/s00429-005-0038-0

Piefke, M., Weiss, P. H., Markowitsch, H. J., & Fink, G. R. (2005). Gender differences in the functional neuroanatomy of emotional episodic autobiographical memory. *Human Brain Mapping*, *24*(4), 313–324. http://doi.org/10.1002/hbm.20092

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2005.12.004

Poldrack, R. A. (2008). The role of fMRI in Cognitive Neuroscience: where do we stand? *Current Opinion in Neurobiology*. http://doi.org/10.1016/j.conb.2008.07.006

Poletti, M., Emre, M., & Bonuccelli, U. (2011). Mild cognitive impairment and cognitive reserve in Parkinson's disease. *Parkinsonism & Related Disorders*, *17*(8), 579–586. http://doi.org/10.1016/j.parkreldis.2011.03.013

Poppenk, J., & Norman, K. A. (2012). Mechanisms supporting superior source memory for familiar items: A multi-voxel pattern analysis study. *Neuropsychologia*, *50*(13), 3015–3026. http://doi.org/10.1016/j.neuropsychologia.2012.07.010

Pridmore, S., Chambers, A., & McArthur, M. (2005). Neuroimaging in psychopathy. *Australian and New Zealand Journal of Psychiatry*. http://doi.org/10.1111/j.1440-1614.2005.01679.x

Rabin, J. S., & Rosenbaum, R. S. (2012). Familiarity modulates the functional relationship between theory of mind and autobiographical memory. *NeuroImage*, *62*(1), 520–529. http://doi.org/10.1016/j.neuroimage.2012.05.002

Raine, A., Meloy, J. R., Bihrle, S., Stoddard, J., LaCasse, L., & Buchsbaum, M. S. (1998). Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers. *Behavioral Sciences and the Law*, *16*(3), 319–332. http://doi.org/10.1002/(SICI)1099-0798(199822)16:3<319::AID-BSL311>3.0.CO;2-G

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*, *42*(1), 2–13. http://doi.org/10.1016/j.neuropsychologia.2003.07.006

Rentz, D. M., Locascio, J. J., Becker, J. A., Moran, E. K., Eng, E., Buckner, R. L., … Johnson, K. A. (2010). Cognition, reserve, and amyloid deposition in normal aging. *Annals of Neurology*, *67*(3), 353–364. http://doi.org/10.1002/ana.21904

Rissman, J., Chow, T. E., Reggente, N., & Wagner, A. D. (2016). Decoding fMRI Signatures of Real-world Autobiographical Memory Retrieval. *Journal of Cognitive Neuroscience*, *28*(4), 1–17. http://doi.org/10.1162/jocn_a_00920

Rissman, J., Greely, H. T., & Wagner, A. D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences*, *107*(21), 9849–9854. http://doi.org/10.1073/pnas.1001028107

Robertson, C. E., Hermann, K. L., Mynick, A., Kravitz, D. J., & Kanwisher, N. (2016). Neural Representations Integrate the Current Field of View with the Remembered 360° Panorama in Scene-Selective Cortex. *Current Biology*, *26*(18), 2463–2468.

Robinson, J. A. (1980). Affect and retrieval of personal memories. *Motivation and Emotion*, *4*(2), 149–174. http://doi.org/10.1007/BF00995196

Roe, C. M., Mintun, M. A., D’Angelo, G., Xiong, C., Grant, E. A., & Morris, J. C. (2008). Alzheimer disease and cognitive reserve: variation of education effect with carbon 11-labeled Pittsburgh Compound B uptake. *Archives of Neurology*, *65*(11), 1467–71. http://doi.org/10.1001/archneur.65.11.1467

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. http://doi.org/10.1016/0010-0285(76)90013-X

Rosenzweig, M. R. (1966). Environmental complexity, cerebral change, and behavior. *The American Psychologist*, *21*(4), 321–332. http://doi.org/10.1037/h0023555

Rosenzweig, M. R., Bennett, E. L., & Krech, D. (1964). Cerebral effects of environmental complexity and training among adult rats. *Journal of Comparative and Physiological Psychology*, *57*(3), 438–439. http://doi.org/10.1037/h0046387

Rzezak, P., Squarzoni, P., Duran, F. L., De Toledo Ferraz Alves, T., Tamashiro-Duran, J., Bottino, C. M., … Busatto, G. F. (2015). Relationship between brain age-related reduction in gray matter and educational attainment. *PLoS ONE*, *10*(10), 1–15. http://doi.org/10.1371/journal.pone.0140945

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S. R., Busa, E., … Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral Cortex*, *14*(7), 721–730. http://doi.org/10.1093/cercor/bhh032

Salvatore, C., Battista, P., & Castiglioni, I. (2016). Frontiers for the early diagnosis of AD by means of MRI brain imaging and support vector machines. *Current Alzheimer Research*, *13*(5), 509–533.

Sato, J. R., Mourão-Miranda, J., Morais Martin, M. d G., Amaro, E., Morettin, P. A., & Brammer, M. J. (2008). The impact of functional connectivity changes on support vector machines mapping of fMRI data. *Journal of Neuroscience Methods*, *172*(1), 94–104. http://doi.org/10.1016/j.jneumeth.2008.04.008

Satz, P., Benson, D. F., Crinella, F., D ’elia, L., Earnest, K., Evans, G., … Jolyn, L. (1993). Brain Reserve Capacity on Symptom Onset After Brain Injury: A Formulation and Review of Evidence for Threshold Theory. *Neuropsychology*, *7*(3), 273–295. http://doi.org/10.1037/0894-4105.7.3.273

Scarmeas, N., Levy, G., Tang, M. X., Manly, J., & Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer’s disease. *Neurology*, *57*(12), 2236–42. http://doi.org/10.1212/WNL.57.12.2236

Scarmeas, N., & Stern, Y. (2003). Cognitive Reserve and Lifestyle. *Journal of Clinical and Experimental Neuropsychology*. http://doi.org/10.1076/jcen.25.5.625.14576

Scarpazza, C., Sartori, G., De Simone, M. S., & Mechelli, A. (2013). When the single matters more than the group: Very high false positive rates in single case Voxel Based Morphometry. *NeuroImage*, *70*, 175–188. http://doi.org/10.1016/j.neuroimage.2012.12.045

Schacter, D. L. ., Wagner, A. D. ., & Buckner, R. L. (2000). Memory Systems of 1999. *The Oxford Handbook of Memory*, 627–643.

Schacter, D. L., & Loftus, E. F. (2013). Memory and law: What can cognitive neuroscience contribute? *Nature Neuroscience*, *16*(2), 119–23. http://doi.org/10.1038/nn.3294

Schmah, T., Yourganov, G., Zemel, R. S., Hinton, G. E., Small, S. L., & Strother, S. C. (2010). Comparing classification methods for longitudinal fMRI studies. *Neural Computation*, *22*(11), 2729–2762. http://doi.org/10.1162/NECO_a_00024

Schroots, J. J. F., van Dijkum, C., & Assink, M. H. J. (2004). Autobiographical Memory From a Life Span Perspective. *The International Journal of Aging and Human Development*, *58*(1), 69–85. http://doi.org/10.2190/7A1A-8HCE-0FD9-7CTX

Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., … Mourão-Miranda, J. (2013). PRoNTo: Pattern recognition for neuroimaging toolbox. *Neuroinformatics*, *11*(3), 319–337. http://doi.org/10.1007/s12021-013-9178-1

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I. M., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex*, *48*(8), 991–996. http://doi.org/10.1016/j.cortex.2011.04.009

Seider, T. R., Fieo, R. A., O'Shea, A., Porges, E. C., Woods, A. J., & Cohen, R. A. (2016). Cognitively Engaging Activity Is Associated with Greater Cortical and Subcortical Volumes. *Frontiers in Aging Neuroscience*, *8*(May), 1–10. http://doi.org/10.3389/fnagi.2016.00094

Serra, L., Cercignani, M., Petrosini, L., Basile, B., Perri, R., Fadda, L., … Bozzali, M. (2011). Neuroanatomical correlates of cognitive reserve in Alzheimer disease. *Rejuvenation Research*, *14*(2), 143–51. http://doi.org/10.1089/rej.2010.1103

Shen, Francis X; Jones, O. D. (2011). Brain Scans as Evidence: Truths, Proofs, Lies, and Lessons. *62 Mercer L. Rev. 861*, *352*, 1–17. http://doi.org/10.2139/ssrn.1736288

Sliwinska, M. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., & Devlin, J. T. (2012). Early and sustained supramarginal gyrus contributions to phonological processing. *Frontiers in Psychology*, *3*(MAY). http://doi.org/10.3389/fpsyg.2012.00161

Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C., & Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate. *Brain*, *124*(9), 1720–1733. http://doi.org/10.1093/brain/124.9.1720

Sole-Padulles, C., Bartres-Faz, D., Junque, C., Vendrell, P., Rami, L., Clemente, I. C., … Molinuevo, J. L. (2009). Brain structure and function related to cognitive reserve variables in normal aging, mild cognitive impairment and Alzheimer's disease. *Neurobiol Aging*, *30*(7), 1114–1124. http://doi.org/10.1016/j.neurobiolaging.2007.10.008

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 309–315. http://doi.org/10.1038/nn1008

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. In *Neurobiology of Learning and Memory* (Vol. 82, pp. 171–177). http://doi.org/10.1016/j.nlm.2004.06.005

Staff, R. T., Murray, A. D., Deary, I. J., & Whalley, L. J. (2004). What provides cerebral reserve? *Brain*. http://doi.org/10.1093/brain/awh144

Steffener, J., Habeck, C., O'Shea, D., Razlighi, Q., Bherer, L., & Stern, Y. (2016). Differences between chronological and brain age are related to education and self-reported physical activity. *Neurobiology of Aging*, *40*, 138–144. http://doi.org/10.1016/j.neurobiolaging.2016.01.014

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage*, *65*, 69–82. http://doi.org/10.1016/j.neuroimage.2012.09.063

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society : JINS*, *8*(3), 448–460. http://doi.org/10.1017/S1355617702813248

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*(10), 2015–2028. http://doi.org/10.1016/j.neuropsychologia.2009.03.004

Stern, Y., Alexander, G. E., Prohovnik, I., & Mayeux, R. (1992). Inverse relationship between education and parietotemporal perfusion deficit in Alzheimer's disease. *Annals of Neurology*, *32*(3), 371–5. http://doi.org/10.1002/ana.410320311

Stern, Y., Alexander, G. E., Prohovnik, I., Stricks, L., Link, B., Lennon, M. C., & Mayeux, R. (1995). Relationship between lifetime occupation and parietal flow: implications for a reserve against Alzheimer's disease pathology. *Neurology*, *45*(1), 55–60. http://doi.org/10.1212/WNL.45.1.55

Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., & Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA : The Journal of the American Medical Association*, *271*(13), 1004–1010.

http://doi.org/10.1001/jama.271.13.1004

Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2006). Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. *Neurology*, *67*(3), 467–473. http://doi.org/10.1212/01.wnl.0000228231.26111.6e

Strother, S., La Conte, S., Kai Hansen, L., Anderson, J., Zhang, J., Pulapura, S., & Rottenberg, D. (2004). Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. In *NeuroImage* (Vol. 23). http://doi.org/10.1016/j.neuroimage.2004.07.022

Sumowski, J. F., Chiaravalloti, N., Wylie, G., & Deluca, J. (2009). Cognitive reserve moderates the negative effect of brain atrophy on cognitive efficiency in multiple sclerosis. *Journal of the International Neuropsychological Society : JINS*, *15*(4), 606–12. http://doi.org/10.1017/S1355617709090912

Sumowski, J. F., & Leavitt, V. M. (2013). Cognitive reserve in multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *19*(9), 1122–1127. http://doi.org/10.1177/1352458513498834

Sumowski, J. F., Rocca, M. A., Leavitt, V. M., Riccitelli, G., Sandry, J., Deluca, J., … Filippi, M. (2016). Searching for the neural basis of reserve against memory decline: Intellectual enrichment linked to larger hippocampal volume in multiple sclerosis. *European Journal of Neurology*, *23*(1), 39–44. http://doi.org/10.1111/ene.12662

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*. http://doi.org/10.1016/j.neuropsychologia.2006.05.023

Taylor, S. F., Liberzon, I., Fig, L. M., Decker, L. R., Minoshima, S., & Koeppe, R. A. (1998). The effect of emotional content on visual recognition memory: a PET activation study. *NeuroImage*, *8*(2), 188–97. http://doi.org/10.1006/nimg.1998.0356

Thomason, M. E., Burrows, B. E., Gabrieli, J. D. E., & Glover, G. H. (2005). Breath holding reveals differences in fMRI BOLD signal in children and adults. *NeuroImage*, *25*(3), 824–837. http://doi.org/10.1016/j.neuroimage.2004.12.026

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. http://doi.org/10.1038/381520a0

Tranel, D., & Jones, R. (2006). Knowing “what” and knowing “when.” *Journal of*

Clinical and Experimental Neuropsychology, 28, 43–66.

Tucker, A. M., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research*, *8*(4), 354–60. http://doi.org/10.2174/1567211212225912050

Tulving, E. (1972). Episodic and semantic memory 1. *Organization of Memory. London: Academic*, *381*(4), 382–404.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annu Rev Psychol*, *53*, 1–25. http://doi.org/10.1146/annurev.psych.53.100901.135114

Tulving, E., Kapur, S., Markowitsch, H. J., Craik, F. I. M., Habibt, R., & Houle, S. (1994). Neuroanatomical correlates of retrieval in episodic memory: Auditory sentence recognition. *Psychology*, *91*, 2012–2015. http://doi.org/10.1073/pnas.91.6.2012

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology*, *15*(1), 121–149. http://doi.org/10.1016/0010-0285(83)90006-3

Uncapher, M. R., Boyd-Meredith, J. T., Chow, T. E., Rissman, J., & Wagner, a. D. (2015). Goal-Directed Modulation of Neural Memory Patterns: Implications for fMRI-Based Memory Detection. *Journal of Neuroscience*, *35*(22), 8531–8545. http://doi.org/10.1523/JNEUROSCI.5145-14.2015

Ung, H., Brown, J. E., Johnson, K. A., Younger, J., Hush, J., & Mackey, S. (2014). Multivariate classification of structural MRI data detects chronic low back pain. *Cerebral Cortex*, *24*(4), 1037–1044. http://doi.org/10.1093/cercor/bhs378

Uttal, W. R. (2015). *Macroneural theories in cognitive neuroscience*. Psychology Press.

Valenzuela, M. J., & Sachdev, P. (2006a). Brain reserve and cognitive decline: a non-parametric systematic review. *Psychological Medicine*, *36*(8), 1065–73. http://doi.org/10.1017/S0033291706007744

Valenzuela, M. J., & Sachdev, P. (2006b). Brain reserve and dementia: a systematic review. *Psychological Medicine*, *36*(4), 441–54. http://doi.org/10.1017/S0033291705006264

Valenzuela, M. J., Sachdev, P., Wen, W., Chen, X., & Brodaty, H. (2008). Lifespan mental activity predicts diminished rate of hippocampal atrophy. *PLoS ONE*, *3*(7), 1–6. http://doi.org/10.1371/journal.pone.0002598 van Praag, H., Kempermann, G., & Gage, F. H. (2000). Neural consequences of environmental enrichment. *Nat Rev Neurosci*, *1*(3), 191–198. http://doi.org/10.1038/35044558

van Praag, H., Shubert, T., Zhao, C., & Gage, F. H. (2005). Exercise enhances learning and hippocampal neurogenesis in aged mice. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(38), 8680–5. http://doi.org/10.1523/JNEUROSCI.1731-05.2005

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews. Neuroscience*, *10*(11), 792–802. http://doi.org/10.1038/nrn2733

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, *30*(6), 655–668. http://doi.org/10.1068/p3029

Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. *Springer*. http://doi.org/10.1109/TNN.1997.641482

Vapnik, V. N. (1998). *Statistical Learning Theory*. *John Wiley & Sons*. http://doi.org/10.2307/1271368

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*. http://doi.org/10.1109/72.788640

Vapnik, V. N., & Chervonenkis, A. Y. (1974). Pattern recognition theory. Nauka, Moscow.

Vass, L. K., & Epstein, R. A. (2013). Abstract Representations of Location and Facing Direction in the Human Brain. *Journal of Neuroscience*, *33*(14), 6133–6142. http://doi.org/Doi 10.1523/Jneurosci.3873-12.2013

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural Scene Categories Revealed in Distributed Patterns of Activity in the Human Brain. *Journal of Neuroscience*, *29*(34), 10573–10581. http://doi.org/10.1523/JNEUROSCI.0559-09.2009

Watanabe, S. (1969). Knowing and Guessing: a Quantitative Study of Inference and Information.

Weygandt, M., Hackmack, K., Pfüller, C., Bellmann-Strobl, J., Paul, F., Zipp, F., &

Haynes, J. D. (2011). MRI pattern recognition in multiple sclerosis normal-appearing brain areas. *PLoS ONE*, *6*(6). http://doi.org/10.1371/journal.pone.0021138

Wik, G., Fischer, H., Finer, B., Bragee, B., Kristianson, M., & Fredrikson, M. (2006). Retrospenial cortical deactivation during painful stimulation of fibromyalgic patients. *The International Journal of Neuroscience*, *116*(1), 1–8. http://doi.org/10.1080/00207450690962208

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods*, *42*(3), 671–684. http://doi.org/10.3758/brm.42.3.671

Wilson, R. S. (2009). Elderly women with larger social networks are.

Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. http://doi.org/10.1109/4235.585893

Woodard, J. L., Seidenberg, M., Nielson, K. a, Miller, S. K., Franczak, M., Antuono, P., … Rao, S. M. (2007). Temporally graded activation of neocortical regions in response to memories of different ages. *Journal of Cognitive Neuroscience*, *19*(7), 1113–1124. http://doi.org/10.1162/jocn.2007.19.7.1113

Wook Yoo, S., Han, C. E., Shin, J. S., Won Seo, S., Na, D. L., Kaiser, M., … Seong, J.-K. (2015). A Network Flow-based Analysis of Cognitive Reserve in Normal Ageing and Alzheimer's Disease. *Scientific Reports*, *5*, 10057. http://doi.org/10.1038/srep10057

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI Time-Series Revisited—Again. *NeuroImage*, *2*(3), 173–181. http://doi.org/10.1006/nimg.1995.1023

Yamashita, O., Sato, M., Yoshioka, T., Tong, F., & Kamitani, Y. (2008). New classification algorithm using sparse estimation for fMRI decoding. *NeuroImage*, 1414–1429.

Yassa, M. A., & Stark, C. E. L. (2009). A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *NeuroImage*, *44*(2), 319–327. http://doi.org/10.1016/j.neuroimage.2008.09.016

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *J Neurosci*, *25*(11), 3002–3008. http://doi.org/10.1523/JNEUROSCI.5295-04.2005