Kent Academic Repository
Full text document (pdf)

Citation for published version

Ihalainen, Riku   (2022) EEG-based effective and functional connectivity for differentiating and predicting altered states of consciousness.   Doctor of Philosophy (PhD) thesis, University of Kent,.

DOI

Link to record in KAR

https://kar.kent.ac.uk/97273/

Document Version

UNSPECIFIED

EEG-based effective and functional connectivity for differentiating and predicting altered states of consciousness

Riku Ihalainen

March 2022

Abstract

How does the brain sustain consciousness? In this thesis, and in the work leading up to it, we provide new computational evidence for the importance of the posterior hot zone on one hand, and for long-distance frontoparietal connectivity on the other, in explaining the contrast between loss of consciousness and in maintaining conscious responsiveness.

We adopt a factorial approach in our study, crossing two altered states of consciousness with two analytical methods for measuring changes in brain associated with these altered states. Specifically, we study healthy controls under propofol-anaesthesia and patients suffering from disorders of consciousness (DoC), employing functional and effective electroencephalographic (EEG) connectivity, thereby forming a 2-by-2 study design.

We first demonstrate the power of functional EEG connectivity for predicting anaesthetic states in the healthy brain, by building a single multivariate regression model combining phase-lag brain connectivity and behaviour- and power-based dependent measures. We show that baseline alpha- and beta-connectivity, as measured prior to an anaesthetic induction, can predict both behaviour- and power-based measures during the induction and peak unresponsiveness, specifically as measured from the posterior electrodes. Next, we study patients suffering from DoC and show that the alpha-band functional connectivity over the left hemisphere, and graph-theoretic network centrality on the right, significantly predict the patient's clinical

diagnosis. Our findings suggest a dissociation between mean spectral connectivity and network properties.

Building on these findings, we then turn to dynamic causal modelling (DCM) to estimate modulations in effective brain connectivity due to anaesthesia, in and between the default mode network (DMN), the salience network (SAL), and the central executive network (CEN). Advancing current understanding of anaesthetic-induced LOC, we show evidence for a selective breakdown in the posterior hot zone and in medial feedforward frontoparietal connectivity within the DMN, and of parietal inter-network connectivity linking DMN and CEN. In a novel DCM-based out-of-sample cross-validation, we establish the predictive validity of our models, specifically highlighting frontoparietal connectivity as a generalisable predictor of states of consciousness. Importantly, we demonstrate a generalisation of this predictive power in an unseen dataset from the post-anaesthetic recovery state.

Finally, we again use DCM to investigate changes in the effective connectivity between DoC patients and healthy controls within the DMN. Specifically, we show that the key difference between healthy controls or conscious patients and completely unresponsive patients is a reduction in left-hemispheric backward frontoparietal connectivity. Finally, with out-of-sample cross-validation, we show that left-hemispheric frontoparietal connectivity can not only distinguish patient groups from each other, it can also generalise to an unseen data subset collected from seemingly unresponsive patients who show evidence of consciousness when assessed with functional

neuroimaging. This suggests that effective EEG connectivity can be used to identify covertly aware patients who seem behaviourally unresponsive.

Overall, this thesis provides novel insights into the brain dynamics underlying transitions between altered states of consciousness and highlights the value of tracking these dynamics in a clinical context. DCM, though computationally more expensive, can accurately predict states of consciousness and provide causal explanations of the brain dynamics that cannot be inferred from functional connectivity alone. Functional connectivity, though correlational, is still an accurate predictive tool of altered states of consciousness. With clinically challenging, ambiguous cases like potentially covertly aware patients, we propose that the causal explanations and accurate predictions of DCM modelling could outweigh the computational complexity.

Acknowledgements

The path towards a PhD thesis is often described as a solitary endeavour. Mine, however, was everything but, and I do wish to acknowledge all those who helped me along the way.

First and foremost, I want to thank my supervisor Dr. Srivas Chennu for his endless support, motivation, guidance, and for always having the time when advice was needed. His constructive and meticulous supervising has been everything I could have ever hoped. In the same breath, I wish to thank my secondary supervisor Prof. Howard Bowman for his continuous support, advice, and feedback throughout the past years. Thank you to everyone I interacted with at University of Kent, particularly Konstantinos Patlatzoglou for all the insightful discussions, Ismail Anwar, George Langroudi, Yuzhou Lin, Lam Pham, Tristan Strange, Ayah Helal, James Brookhouse (and those I forgot) for peer companionship and especially for your friendship. Thank you to my panel Dr. Rogerio de Lemos, Dr. Palaniappan Ramaswamy and Prof. Sally Fincher, and to the School of Computing at University of Kent for their support in the form of a research scholarship. I also want to thank Dr. Daniel Bor from the Department of Psychology at University of Cambridge and Dr. Caroline Li from the School of Computing at University of Kent for agreeing to examine this thesis. I look forward to the discussions.

Thank you to our collaborators, without whom this research would have been impossible to finish. Particularly, I am deeply grateful to Dr. Olivia Gosseries, Dr. Jitka Annen, Rajanikant Panda and Prof. Steven Laureys and all others in the Coma Science Group at University-Hospital of Liège. Thank you to Dr. Katie Warnaby and

List of Publications

Journal Papers

- Ihalainen, R., Gosseries, O., Van de Steen, F., Raimondo, F., Panda, R., Bonhomme, V., Marinazzo, D., Bowman, H., Laureys, S. & Chennu, S. (2021). How hot is the hot zone? Computational modelling clarifies the role of parietal and frontoparietal connectivity during anaesthetic-induced loss of consciousness. *NeuroImage*, *231*, 117841.

Conference Posters & Talks

- Ihalainen, R., Annen, J., Thibaut, A., Chatelle, C., Cassol, H., Martens, G., Schnakers, C., Gosseries, O., Laureys, S. & Chennu, S. (2018). Poster presentation: Hemispheric electroencelphalographic connectivity is correlated with behavior and brain metabolism in disorder of consciousness. *Association for the Scientific Study of Consciousness conference (ASSC22),* in Krakow, Poland.

- Ihalainen, R., Gosseries, O., Raimondo, F., Bonhomme, V., Laureys, S. & Chennu, S. (2019). Flash Talks and Poster presentation: Computational modelling of posterior hot zone in anaesthesia-induced loss of consciousness. *Studying, Measuring, and Altering Consciousness (Luminous Project),* in Merton College, University of Oxford, England.

Unrelated work

- Ihalainen, R., Kotsaridis, G., Vivas, A. B., & Paraskevopoulos, E. (2022). The effect of musical training on the processing of audiovisual correspondences: Evidence from a reaction time task. *Manuscript submitted for publication.*

Table of Contents

1. Introduction

This chapter provides the general background into the overarching theme in this thesis, namely the study of consciousness. We introduce what we mean by consciousness and highlight the theoretical and practical reasons for studying it. We end the chapter by outlining the specific approach chosen in this research, and by providing the general organisation of this thesis into parts and chapters.

1.1   What is consciousness and why study it?

Consciousness poses arguably one of the most baffling problems in human knowledge; understanding it has been proclaimed as one of the major unsolved problems in biology and the ultimate intellectual challenge (Chalmers, 1995; Dehaene & Changeux, 2004; Koch, 2004). We are not aware of the electrochemical signalling occurring in billions of synapses at any given time in our life. Rather, there is a subjective element associated with such brain activation that we are – or can be – aware of; a phenomenal property of 'what it is like', for example, to see something or to hear something, or to *be* something (Nagel, 1974). That is, a system or a process can be argued to be conscious if there is something that *it is like to be* that system or process.

In this thesis consciousness is therefore generally defined as (having) a qualitative subjective experience (of any kind). This refers to the so-called 'qualia', for example, to the experience of colours (e.g. the 'redness' of the

colour red. See *qualia*, e.g. Damasio, 1999; Koch, 2019; Ramachandran & Hirstein, 1997, but also Dennett, 1988, 2014). In the literature, the subjective experience and the corresponding physical process of 'feeling/experiencing' is typically segregated. A common example to illustrate this segregation describes a neuroscientist living in a fully black-and-white environment, who works towards understanding the perception of colour. As the years pass, she learns to know the complete physical truth about colour perception – the physics, chemistry, neuroscience etc. – but lacks the perceptual experience of colour. Eventually, she leaves the room and experiences colour for the first time; the central question of the thought-experiment is whether any new knowledge was gained by the subjective experience (Jackson, 1982).

This distinction between the physical correlates of experience and the phenomenal experience itself – and further, why should there be an experience at all – relate to a problem Chalmers famously termed as 'the hard problem of consciousness' (in contrast to 'the easy problems', e.g. explaining how the brain integrates information, categorises and discriminates stimuli, or focuses attention; Chalmers, 1995, 1996). That is, how does experience arise out of presumably non-sentient matter? Whether this gap can be fulfilled is a debate not settled. For example, Chalmers argues that the gap cannot be fulfilled, and thus, rejects the idea of physicalism of consciousness (ibid.). Other philosophers have argued that explaining consciousness in neurobiological terms is not only attainable (Churchland, 2005), but that once we fully understand how consciousness emerges, then we understand why consciousness is produced and there will be nothing else to explain (Dennett, 2005).

13

A number of neuroscientists have taken a more pragmatic approach to the hard problem and towards the problem of consciousness overall. For example, 'the real problem of consciousness' asks how to account for the properties and phenomenology of consciousness in terms of biological mechanism. That is, how can we distinguish different aspects of consciousness and map their first-person subjective descriptions (i.e. phenomenological properties) onto underlying biophysical mechanisms (Seth, 2016). Another view considers consciousness as a highly limited mental workspace enabling complex processing, comparisons, and combinations of otherwise disparate forms of information (Bor, 2012; Bor & Seth, 2012). In providing such novelty into problem solving, more advanced strategies to reach biological goals, that may otherwise be challenging or even impossible to reach, can be formed. Hence, under this view, consciousness may serve an evolutionary role by providing innovative solutions to complex problems.

Regardless of how interesting the philosophical endeavours trying to explain why consciousness exists in the universe at all are (i.e. the hard problem), it is by no means clear that we can make any significant headway in explaining consciousness by trying to answers such problems directly. Moreover, it is not obvious that by just simply thinking about it, one can make significant progress in explaining consciousness, although this of course can guide our thinking, models and theories. Rather, studying and explaining the properties of consciousness, and their respective neuronal basis, via empirical research and experimentation, might get you somewhere. Hence, in this thesis, we leave the philosophical side to actual philosophers and align more with 'the real problem',

focusing on what can we tell about how consciousness is correlated with the processes in the brain – especially changes in functional and effective/directed connectivity at the network level.

To this end, several cortical and cortico-subcortical network-level mechanisms have been proposed to explain human consciousness and its loss. For example, networks involving subcortical structures such as the thalamus (e.g. Fridman, Beattie, Broft, Laureys, & Schif, 2014; Schiff, 2010) and/or the upper parts of the brainstem (Parvizi & Damasio, 2003; Solms, 2019; Solms & Friston, 2018) have been suggested to play a pivotal role in consciousness. Of the cortical mechanisms, two in particular have received an increasing amount of interest and supporting evidence. On the one hand, empirical studies have suggested an association between loss of consciousness and disruptions of within- and between-network connectivity in large-scale frontoparietal networks (Bor & Seth, 2012; Laureys & Schiff, 2012). On the other, temporo-parieto-occipital areas – colloquially named as 'the posterior hot zone' – have been suggested to mediate changes in consciousness during sleep (Siclari et al., 2017; Lee et al., 2019), and in patients with brain damage (Vanhaudenhuyse et al., 2010; Wu et al., 2015). However, whether the neural correlates of consciousness (NCC) have an anterior contribution (Bor & Seth, 2012; Chennu et al., 2014; Chennu, O'Connor, Adapa, Menon, & Bekinschtein, 2016; Del Cul et al., 2009; Laureys & Schiff, 2012) or are predominantly posterior (Koch, Massimini, Boly, & Tononi, 2016; Koch et al., 2016b; Siclari et al., 2017), is still debated.

In addition to increasing our understanding of the biological and physical origins of consciousness by exploring the NCC (Koch et al., 2016a; Rees, Kreiman, & Koch, 2002), applying rigorous scientific methods to study consciousness is important also for more practical reasons. For example, in recent years, a number of neuroimaging technologies have been proposed to aid in assessment of consciousness after brain injury, and consequently, to help addressing the diagnostic and prognostic challenges in such patients. These include positron emission topography (PET; Stender et al., 2014; Thibaut et al., 2012), magnetic resonance imaging (MRI; Demertzi et al., 2015), and electroencephalogram (EEG; Chennu et al., 2014, 2017; King et al., 2013; Lehembre et al., 2012; Sitt et al., 2014). Similarly, the evidence base for the use of neurotechnology in guiding assessments of the depth of general anaesthesia has been increasingly built in recent years (Chennu, O'Connor, Adapa, Menon, & Bekinschtein, 2016; Cimenser et al., 2011; Mhuircheartaigh, Warnaby, Rogers, Jbabdi, & Tracey, 2013; Purdon et al., 2013; Warnaby, Sleigh, Hight, Jbabdi, & Tracey, 2017).

The hope is that such research may lead both to better theoretical understanding of the physiological basis of consciousness, and to new and improved ways to assess individual states of consciousness, and to reach diagnosis and prognosis of such states. Such research may have far-reaching consequences; for example, patients in what earlier was called a 'vegetative state' (nowadays termed unresponsive wakefulness syndrome, UWS, Laureys et al., 2010) by definition lack of the subjective experience of pain (Jennett & Plum, 1972). This has led to situations where, depending on the national policies, small

surgeries may be performed without anaesthesia and physicians may be allowed to withhold or withdraw life-sustaining treatment like artificial respiration, resuscitation and artificial nutrition and hydration (Buckley et al., 2004; Demertzi et al., 2013; Kuehlmeyer et al., 2014). Notwithstanding these possibly critical consequences, the issue of subjective pain experience in UWS patients has not been settled. Not only have studies indicated that the rate of misdiagnosis of UWS borders 40% (van Erp et al., 2015; Schnakers et al., 2009), brain imaging studies have concurrently suggested residual pain-related activity indicating that many of these patients may possess components of awareness (e.g. Boly et al., 2008; Chatelle et al., 2014; Kassubek et al., 2003; Laureys et al., 2002).

Hence, we do not only focus on illuminating the changes in the underlying brain dynamics due to transitions in the state of consciousness at the group level, but also on whether the information about connectivity can be used in the clinical context to, for example, predict individual states of consciousness. To this end, we estimate both functional connectivity (dwPLI; Vinck, Oostenveld, Van Wingerden, Battaglia, & Pennartz, 2011) and effective connectivity (dynamic causal modeling for cross-spectral densities, DCM CSD; Friston et al., 2012; Moran et al., 2009) in transitions in consciousness due to propofol-anaesthesia and disorders of consciousness (DoC). By adopting this 2-by-2 study design (figure 1.1), we are able to gain novel insights into network-level differences and more accurate methods for predicting individual states of consciousness at both, the level of electrodes and neural sources.
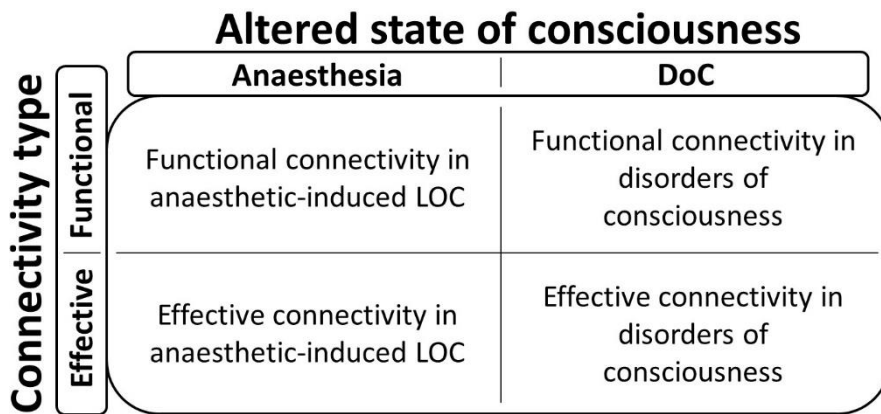
Figure 1.1. Illustration of the 2-by-2 design used in this thesis. We estimate both functional (FC) and effective connectivity (EC) in altered states of consciousness induced by propofol-anaesthesia (loss of consciousness, LOC) and disorders of consciousness.

1.2 Organisation of the thesis, aims, research questions, and hypotheses

This thesis starts from the general background and introduction in this chapter. Chapter 2 provides a review of the methods and techniques used in this thesis, and continues with an overview of the relevant theories and prior research on the topic of consciousness.

The next four chapters form the main body of this thesis and consists of four research chapters exploring the brain dynamics in altered states of consciousness due to general anaesthesia and DoC. It begins in Chapter 3, by exploring a previously proposed EEG marker for anaesthetic-induced loss of consciousness (LOC), namely the slow-wave activity saturation (SWAS), and its possible correlation with pre-anaesthesia resting brain connectivity (top-left cell in figure 1). In particular, we combine resting state EEG recorded from healthy volunteers

under propofol-induced sedation with individual measurements of propofol-concentrations in blood plasma, alongside behavioural measures of responsiveness. We assess the participants' spectral power and functional connectivity during a pre-anaesthetic baseline period and correlate them with individual variation in propofol-concentrations, time needed to reach SWAS, level of SWA-power, and time of loss of behavioural responsiveness (LOBR). In doing so, we hypothesise and provide evidence for SWAS as a distinct and individualised index for propofol-induced loss of consciousness, that is linked with pre-anaesthesia baseline brain connectivity. This research was conducted in collaboration with Katie Warnaby (FMRIB Centre, University of Oxford) and Jostein Holmgren (Anaesthesia Neuroimaging Group, University of Oxford).

In chapter 4, we explore the level of granularity with which functional connectivity can distinguish and predict DoC states (cell 2 in figure 1). In particular, we are interested in the possible hemispheric differences in EEG brain connectivity and hypothesise a specific role for left-hemisphere in distinguishing the DoC states, especially UWS from minimally conscious states (MCS- and MCS+. See section 2.6.3 for details). Linking bedside EEG data with behaviour-based diagnostic labels, we provide evidence for left-hemisphere-specific differences in connectivity and in doing so lay the groundwork for further investigation of brain connectivity in DoC at the level of neural sources (chapter 6). For this work, we collaborated with the Coma Science Group in University-Hospital of Liège with Jitka Annen, Aurore Thibaut, Camille Chatelle, Helena Cassol, Geraldine Martens, Caroline Schnakers (University of California), Olivia Gosseries, and Steven Laureys.

In chapter 5, we explore the changes in effective connectivity following propofol-induced LOC (cell 3 in figure 1). In particular, we assess the differences in intra- and inter-network effective connectivity in three key resting state networks using dynamic causal modelling (DCM) to explain cross-spectral densities from EEG data. In doing so we contribute to the frontal vs. posterior debate of neural correlates of consciousness by providing novel computational evidence supporting a selective breakdown of posterior parietal and medial feedforward frontoparietal connectivity within the default mode network (DMN) and of parietal inter-network connectivity linking the DMN and the central executive network. Going further, we establish the generalised predictive validity of our models using a novel DCM-based cross-validation, by predicting unseen data from the post-anaesthetic recovery state. This research was a collaboration with Frederik Van De Steen and Daniele Marinazzo (Ghent University) and with the Coma Science Group and GIGA consciousness in University-Hospital of Liège with Olivia Gosseries, Federico Raimondo, Rajanikant Panda, Vincent Bonhomme, and Steven Laureys.

Chapter 6 continues our work exploring the brain connectivity in DoC. We assess differences in effective connectivity between cortico-cortical regions of the DMN between DoC patients (UWS and MCS+) and healthy controls using DCM for EEG (cell 4 in figure 1). Following the methodology defined in chapter 5, we first provide evidence for reduction in left frontoparietal connectivity and test the prospective performance of DCM-connectivity within DMN in classifying states of consciousness in DoC based on two connectivity subsets: frontoparietal and posterior connections. We follow this by adopting a data-

driven approach to the classification problem by investigating the predictive performance of single connections, and provide evidence for a key difference in left top-down frontoparietal connectivity in distinguishing the UWS patients from MCS+ and healthy controls. Finally, we demonstrate that our DCM models generalise to a more difficult classification problem: in a leave-one-*state*-out cross-validation paradigm, we train the models on UWS patients with a confirmed PET negative diagnosis on the one hand, and either healthy controls or MCS+ patients on the other. The models are then tested on datasets from 'covertly aware' UWS patients with a PET positive diagnosis. With this generalisation, we provide evidence for the hypothesis that if our modelled effects are valid, and if the sustained PET metabolism reflects covert consciousness in the UWS PET+ patients, our model should classify these patients as healthy controls/MCS+ rather than UWS PET-. In this research, we collaborated with the Coma Science Group in the University-Hospital of Liège.

Chapter 7 concludes this thesis. We first provide an overview of the main findings of the research chapters. We then combine the main conclusions, evaluate the robustness of the results, and reflect the contribution of this thesis to the current research. Last, we provide potential follow-up directions to advance the examination of the neural dynamics of the altered consciousness explored in this thesis.

In the next chapter, we introduce the relevant methods, alongside with the relevant prior research and theories of consciousness.

2. Literature review

In this chapter, we will provide an overview on the specific methods used in this thesis to examine the difference in network-level neural dynamics in anaesthesia and disorders of consciousness. First, we will provide an overview of human electrophysiology and the concept of resting state brain connectivity. We will then cover the necessary steps for electroencephalography (EEG) analysis used in this thesis. In particular, we introduce the preprocessing pipeline (Chennu et al., 2017; Chennu 2018), spectral frequency analysis, functional connectivity, and network measures based on graph/information theory. We will then provide a brief outline of the theoretical basis of dynamic causal modelling and parametric empirical Bayes, which we used to evaluate effective connectivity in our models. We will then turn back to study of consciousness, providing an overview of the leading theories of consciousness relevant to this thesis, namely the Global (Neuronal) Workspace Theory (GNWT; Baars, 1997; Dehaene & Changeux, 2011; Dehaene, Changeux, & Naccache, 2011) and Integrated Information Theory (IIT; Oizumi, Albantakis, & Tononi, 2014; Tononi, 2004, 2012; Tononi, Boly, Massimini, & Koch, 2016). Last, we will provide an overview of the previous research on anaesthetic- and brain damage-induced altered states of consciousness, focusing particularly on common EEG markers, and on functional and effective connectivity.

2.1 Human electrophysiology and resting state brain connectivity

In recent decades, neuroimaging has consolidated its place as the predominant technique in behavioural and cognitive neuroscience. These techniques include, for example, functional magnetic resonance imaging (fMRI), EEG, magnetoencephalography (MEG), and positron emission topography (PET). Each of the modalities have been used in the context of studying consciousness, and each of them have their particular set of strengths and weaknesses. For example, fMRI has a very high spatial resolution (typically around 3-4 mm voxel size, although much smaller voxel sizes can be achieved with higher field magnets), whereas its temporal resolution is restricted by the hemodynamic response time (commonly around 3 seconds, with a peak occurring 5-6 seconds after stimulus; Glover, 2011).

In contrast, EEG is commonly attributed with excellent temporal resolution (<100ms) but with poorer spatial resolution (reasonably accurate at the centimetre scale; Burle et al., 2015; Cohen, 2017). Furthermore, due to a number of complicating factors – for example spatial blurring and signal mixing due to different volume conduction in the tissues of the head – determining exactly which neural events caused the observed signal is not possible without imposing a priori constraints on the models (Grech et al., 2008; Michel et al., 2004). It is also generally believed that activity of deep brain structures is not registered in scalp activity (hence, the focus on cortical mechanisms in this thesis; However, see for example Krishnaswamy et al., 2017; Seeber et al., 2019 for scalp recordings of deeper brain structures). Nevertheless, due to its portability, cost effectiveness, and accessibility, EEG is a particularly attractive solution for

studying brain function in transitions of consciousness, for example with patients suffering from brain damage. We go through some of the key literature providing evidence for the usefulness of EEG in the context consciousness studies later in this chapter.

The EEG signal is thought to arise from post-synaptic potentials from the deep layers of pyramidal cells (Cohen, 2017) and neuronal oscillations cover a broad frequency spectrum. In the literature, the canonical bands are commonly defined as 0.5-4 Hz (delta), 4-8 (theta), 8-13 (alpha), and >13 Hz (beta; Teplan, 2002). There is a long history relating EEG oscillations and cognitive functions together (Karakaş, 2020; Klimesch, 1999; Nyhus & Curran, 2010; Harmony, 2013) and in laboratory settings, ongoing EEG signals have been time-locked to the occurrence of task-related stimuli (the event-related potential, ERP; (Luck & Kappenman, 2011; Luck, 2014). Commonly, ERPs are contrasted with 'baseline' signals recorded at rest (i.e. when no stimulus is presented). This resting state activation corresponds with specific brain networks whose activity is suspended during goal-directed behaviours, and as such, it has been argued to provide a reliable baseline measure against which task-elicited activation can be compared (Gusnard and Raichle 2001; Raichle et al. 2001, but see also Klein, 2014).

The identification of such resting state networks (RSN) has led to an explosion of interest in studying the underlying brain dynamics without the presence of any particular stimulus (van Diessen et al., 2015). The underlying idea behind these networks is that cortical areas are to some extent specialised for some aspects of sensory, motor, and cognitive processing and that this

specialisation is anatomically segregated within the cortex. This is in contrast to functional localisation which implies that function can be localised to a specific cortical area. Nonetheless, the cortical infrastructure supporting cognitive functions may involve many of these specialised areas unified and mediated by functional integration. This integration manifests in neural oscillations which reflect the communication between the different cortical areas. This integration can in turn be characterised in terms of connectivity (Razi & Friston, 2016).

There are three commonly studied types of brain connectivity: structural, functional, and effective connectivity. Structural connectivity refers to the anatomical connections in the brain – to the white matter – that has been referred to as the connectome (Sporns, Tononi, & Kötter, 2005). We define functional connectivity in this thesis following Friston (2011) and Razi & Friston (2016) as statistical dependencies between distinct neurophysiological events. These dependencies can be undirected (as in correlation or coherence) or directed (as in Granger causality (GC) and transfer entropy; Granger, 1969; Schreiber, 2000). In contrast, causal/effective connectivity is seen as causal influence (in a control theory sense) of one neural population over another (Stephan et al., 2010). DCM can also be characterised as a framework that enables estimation and testing of specific models (and/or hypotheses) from a relatively large model-space directly. This feature is even further strengthened in a relatively recent extension to the DCM procedure – namely parametric empirical Bayes (PEB; Friston et al., 2016) – as Bayesian model reduction (BMR) as part of PEB enables the estimation of significantly larger model spaces efficiently (see section 2.5.2 for further details about PEB and BMR).

In this thesis, we are especially interested in estimating the causal connectivity – and to test its potential predictive power – in three well-known resting state networks (RSN; default mode, salience, and central executive networks) in addition to a large model comprising of the three RSNs with additionally hypothesised inter-RSN connections. Hence, to estimate effective connectivity in these models, DCM is used in chapters 5 and 6. The estimated DCM models are in addition used to predict states of anaesthesia and disorders of consciousness (DoC).

However, DCM is a more complex and computationally expensive method requiring more time and effort to master in comparison to, for example, estimations of undirected functional connectivity. If accurate prediction preformance can be obtained with undirected functional connectivity, performing computationally expensive analyses for classifications would be resources wasted. Thus, in chapters 3 and 4 we use functional connectivity estimates alongside linear regression to predict states of anaesthesia and DoC, respectively.

Next, we provide an overview of the analytical steps involved in transforming raw human EEG signals recorded from the scalp into the functional and effective connectivity measures used in this thesis.

2.2   Electroencephalography – preprocessing

In this thesis, we used EEG to capture the cortical dynamics to assess and predict altered states of consciousness. The raw EEG signal is known to contain inherent artefacts (e.g. noise artefacts, eye blinks, movement artefacts). Hence, a

data preprocessing pipeline is typically applied to remove these confounds. While there is no universally accepted standard for the preprocessing pipeline, the steps therein typically include downsampling, filtering, epoching, noise line and artefact removal, and re-referencing (Kim, 2018). Below, we present a short description of the data analysis pipeline used in this thesis. All chapter-specific details are further described in the corresponding chapters.

We obtained raw EEG data as captured with 256-channel high-density EEG (EGI; chapters 4, 5, and 6) with Net Amps amplifier and both raw and preprocessed EEG datasets as captured with a 32-channel EEG cap (BrainCap MR, Easycap GmbH; chapter 3) with an MR-compatible amplifier system (BrainAmp MRplus, Brain Products GmbH). The preprocessing steps for the raw EEG signals in chapters 4, 5, and 6 followed the data analysis pipeline provided in Chennu et al. (2017) and Chennu (2018; *MOHAWK* v.1.2. Available from: https://github.com/srivaschennu/MOHAWK). In chapter 3, we preprocessed a 32-channel EEG dataset acquired during anaesthesia using the same pipeline as above. This dataset was recorded after participants reached the peak dose of propofol (fully anaesthetised for 10 minutes, peak-unresponsiveness). In addition, in chapter 3, we analysed a preprocessed dataset collected during the propofol induction period.

In brief, the pipeline utilises functionality provided by EEGLAB (Delorme & Makeig, 2004), FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011), and the Brain Connectivity Toolbox (Rubinov & Sporns, 2010), implemented in MATLAB2017a (The Mathworks Inc., 2017). From the high-density recordings,

we first removed the EEG channels on the face, neck, and near the eyes to minimise the influence of muscular and ocular artefacts, consequently retaining 173 channels for further processing (32 channels in chapter 3). The data were then downsampled to 250 Hz. As most electromyographic noise was observed in high frequencies, the data were bandpass-filtered within a range of 0.5 – 45 Hz, encompassing the canonical delta (0 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 13 Hz), beta (13 – 30 Hz), and gamma (30 – 45 Hz) bands (except in chapter 3, where the induction dataset was already bandpass-filtered between 0.5 – 30 Hz). The recordings were then epoched into 10-second long segments – an epoch length previously applied and seen fit in the context of altered states of consciousness due to propofol-anaesthesia (Chennu et al., 2016) and disorders of consciousness (Chennu et al., 2014; 2017). The time points within each epoch were baseline-corrected relative to the mean voltage over the entire epoch. Artefacts were rejected using a quasi-automated procedure that flags abnormally noisy channels and epochs by calculating their variance, which then are confirmed by visual inspection. Next, the infomax independent component analysis algorithm (Bell & Sejnowski, 1995) was used to identify and remove components of activity from potentially non-neural origins, followed by interpolation of the removed (noisy) channels using a spherical spline interpolation. Lastly, the data were re-referenced to the common average and the first 60 epochs (first 10 minutes) were retained for further analysis.

2.3   Frequency analysis

EEG data allows a spectral decomposition of the signal into different
frequencies quantifying the amount of oscillatory activity at each frequency. At
each point in time, the signal's amplitude (square root of power) and its phase
can describe the oscillations at a particular frequency (figure 2.1).
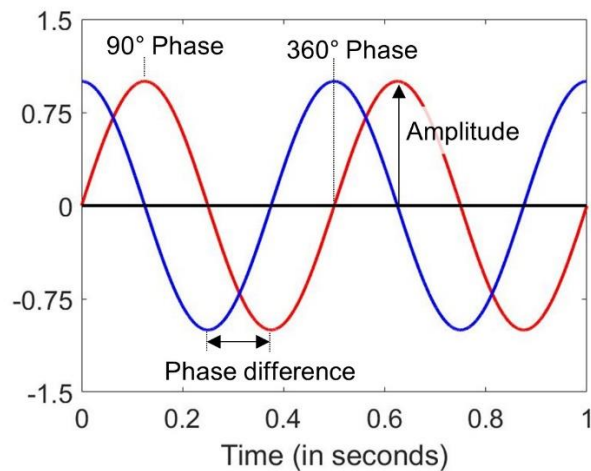
Figure 2.1. Illustration of amplitude and phase for two sine waves (2 Hz)
with 90° difference in phase. Essentially, phase of a function quantifies the
fraction of the cycle covered up to t (here, time), whereas amplitude quantifies
the amount of a specific frequency included in the signal.

In signal analysis, the Fourier transform is a common method for calculating
the power and phase spectra: with a continuous signal $x(t)$ the continuous Fourier
transform (CFT) into different frequencies f is defined as

$$X(\omega) = \sum_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt,$$

29

where $e^{-j\omega t}$ are the complex exponentials, and ω is the angular frequency corresponding to the linear frequency f ($\omega = 2\pi f$). The square magnitude of $X(\omega)$ ($|X(\omega)|^2$) is the power spectrum.

With EEG, however, the data are recorded only for a limited time interval, whereas in above, it is assumed that the signal is continuous in time and infinite in length. Consequently, the discrete Fourier transform (DFT) is used instead. DFT is defined as

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N},$$

where $k = 0, \ldots, N - 1$.

The traditional Fourier analysis has a number of shortcomings with regard to analysing EEG signals: the time-frequency resolution is relatively poor, they cannot differentiate oscillations from artefacts, and are generally designed for stationary and regular signals (Bruns, 2004; van Vugt, Sederberg, & Kahana, 2007). To overcome these limitations, it has been shown that by obtaining and averaging multiple independent frequency decompositions from the same sample by segmenting the signal, the estimation bias of oscillation detection can be reduced (i.e. 'tapering' of the data; ibid.). These methods use a certain number K different orthogonal windows (taper functions) that – when averaged – reduce the variance in the estimate by providing an independent estimate of the signal. This multitaper method has been shown to provide more reliable power estimations for noisy data.

Accordingly, in chapters 3 and 4, this multitaper method – with 5 Slepian tapers – was used to estimate the spectral power in the canonical frequency bands 0.5-4 Hz (delta), 4-8 (theta), 8-13 (alpha), and >13 Hz (beta). The continuous resting data were segmented into 10-second long windows – an epoch length previously applied and seen fit in the context of altered states of consciousness due to propofol-anaesthesia (Chennu et al., 2016) and disorders of consciousness (Chennu et al., 2014; 2017). In chapter 3, multitaper frequency transformation was implemented. Spectral power values between 0.5 and 30 Hz were calculated within bins of 0.1 Hz, alongside with cross-spectrum between the time-frequency decompositions at frequency bins of 0.06 Hz (with spectral smoothing of 0.3). For the cross-spectrum, we used zero-padding to increase the frequency resolution (length determined by rounding up the maximum trial length up to the next power of 2. See FieldTrip, cfg.pad; Oostenveld et al., 2011). In chapter 4, spectral and cross-spectral estimates were calculated at bins of 0.1 Hz between 0.5 and 45 Hz. In addition, in chapter 3, we estimated and predicted various measures based on slow-wave activity saturation (SWAS; 0.5 – 1.5 Hz). SWAS is a characteristic waveform seen in both anaesthesia and sleep (Murphy et al., 2011) and has been suggested as an important manifestation of perception loss indicating individual's depth of anaesthesia (Mhuircheartaigh et al., 2013; Warnaby et al., 2017).

2.4 Functional connectivity

Connectivity can be used to characterise the non-trivial, mathematical relationship between two signals, for example, between EEG signals from two electrodes. Quantifying systematic phase-coupling, i.e. systematic differences in phase (lags and leads), is an example of this, which has been hypothesised to reflect a flexible communication structure between brain regions (Engel, Fries, & Singer, 2001; Fries, 2005). One traditional example of measuring this phase-synchronisation is spectral coherence, which indicates linear correlations between two signals (Adey, Walter, & Hendrix, 1961; Sakkalis, 2011). However, it intermingles phase and amplitude correlations, and hence, increases with amplitude covariance, leaving the relative importance of phase covariance unclear (Lachaux et al., 1999). This has led to development of phase-synchronisation methods that use only the relative phase between signals, such as the phase-locking value (PLV; ibid).

PLV produces a value between 0 and 1 indicating the amount of synchronisation between the phases of the signals. However, it is well known that indexing phase-synchronisation can be complicated by the presence of a common reference, the presence of noise sources, sample-size bias, and by volume-conduction of source activity. In the case of EEG, volume conduction of source activity and the use of a common reference can spuriously inflate phase-synchronisation indices (Vinck et al., 2011). To overcome these problems, the use of the imaginary part of the coherence (ImC; Nolte et al., 2004) and ImC's improved version, the phase-lag index (PLI; Stam, Nolte, & Daffertshofer, 2007), have been suggested. PLI estimates the extent to which phase leads or lags are

32

non-equiprobable (irrespective of the magnitudes of the phase angle differences). By averaging the activity detected at $0°$ or $180°$ phase differences between the signals, PLI effectively removes the largest contributions of volume-conducted activity (assuming that volume conduction results in either identical ($0°$) or opposite ($180°$) phases; Stam et al., 2007).

Finally, Vinck et al. (2011) suggested weighting of the signs of the observed contributions of phase angle differences (leads and lags) by the magnitude of the imaginary part of the cross-spectrum (the absolute magnitude). By weighting the PLI in this way (wPLI), they demonstrated reduced sensitivity to noise and increased sensitivity to detect changes in the phase-synchronisation. Vinck et al. (2011) further introduced a debiased estimator of the wPLI (dwPLI) to correct for sample-size bias present in both, PLI and ImC.

It is worth noting though, that down-weighting identical and opposite phases may not only improve the robustness to noise and sample-size bias, but lead to underestimation of true connectivity patterns at small time lags (Cohen, 2015). Nevertheless, dwPLI has previously been utilised and tested specifically in the context of anaesthetic-driven unconsciousness (e.g. Chennu et al., 2016; Kim et al., 2016), as well as in the context of disorders of consciousness (Chennu et al., 2014; 2017). Hence, taking into consideration the advantages of dwPLI over the other phase-synchronisation measures, we use dwPLI as the measure of functional connectivity in chapters 3 and 4.

Therefore, the cross-spectrum between every pair of electrodes was used to calculate the dwPLI values. Within each frequency band, dwPLI values at the

peak frequency of the oscillatory signals in the average spectrum across channels was used to represent the connectivity between the channel pairs. For all subjects, these dwPLI values across all channel pairs were used to construct symmetric 32 x 32 (chapter 3) and 173 x 173 (chapter 4) connectivity matrices for the canonical frequency bands.

2.4.1 Brain-network analysis

Brain connectivity estimated as above can be used to construct a map of connections that can be modelled as a graph. This allows us to apply well-established methods from network science to characterise the properties of these EEG-derived brain networks, and measure how alterations in consciousness modulates network properties. A prominent example of a well-studied network property is its small-worldness (Watts & Strogatz, 1998). A related example, network centrality, which we use in chapter 4, indexes the presence of inter-modular connectivity hubs. This measure, as captured by the participation coefficient (Guimera & Amaral, 2005), has been previously shown to correlate with the level of awareness and PET metabolism in DoC patients (Chennu et al., 2014; Chennu et al., 2017). Essentially, participation coefficient captures the distribution of a node's edges among the communities of a graph. A node with its edges entirely restricted to its community has a participation coefficient of zero, while a node with its edges evenly distributed among all communities has a maximal participation coefficient approaching to one ("connector hubs").

To quantify participation coefficients, each dwPLI connectivity matrix constructed as described earlier was proportionally thresholded, retaining between 90% - 10% of the largest dwPLI values. After applying the thresholds, the matrices were binarised by setting the non-zero values to one. This procedure was repeated at each value within the threshold (in steps of 2.5%) and these binarised matrices were modelled as a network with channels as nodes and the zero-values as edges. Finally, using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010), we calculated the participation coefficients for all frequency bands, using the global average over the thresholded range of connection densities. Following the analysis pipeline (Chennu et al., 2017; Chennu, 2018) – we calculated the standard deviation of participation coefficients across network nodes. Larger standard deviations would indicate a diversity of participation coefficients, with some nodes – with high participation coefficients – serving as hubs, and other nodes – with low participation coefficients – in the periphery of the network. To relate these network properties to behavioural and clinical measures, we subjected these standard deviations to multiple linear regression analyses.

2.5 Effective Connectivity – Dynamic causal modeling: General case

During the past decade, dynamic causal modeling (DCM) has become one of the most predominant ways to characterise causal[1] – or effective – connectivity within distributed neuronal networks; it is a generic approach to infer hidden (unobserved) neuronal states from measured brain activity (Daunizeau, David, & Stephan, 2011; Razi & Friston, 2016). Originally introduced for fMRI data (Friston, Harrison, & Penny, 2003), DCM has since been extended for event-related electromagnetic responses (David et al., 2006), and induced/steady-state (cross-spectral) responses (Friston et al., 2012; Moran et al., 2009). In general, DCMs link the specific features of the measured data to the modelled neuronal dynamics using a generative model; they combine a hemodynamic (fMRI) or an electromagnetic forward model (M/EEG) with differential equations describing the (hidden) neuronal dynamics (neural state equations). The posterior densities for the model parameters (probabilities describing the uncertainty about a parameter θ after observing data) in addition to an approximation of the log model evidence can then be obtained using Bayesian inversion. In doing so, DCMs enable us to draw inferences on the model parameters, such as the

---

[1] It is important to note what is meant by 'causality' here; DCMs are causal in a control-theory sense (Friston, 2009). In DCMs, causal interactions among hidden state variables (i.e. time-varying properties of neural populations that cannot be observed directly) are expressed in terms of differential equations. These equations describe how the present state of one neuronal population causes change (i.e. dynamics) in another, and importantly, how these interactions are modulated by external manipulations or endogenous brain activity (Stephan et al., 2010).

synaptic time constants, conduction delays – and importantly – connection
strengths, in addition to the most likely model (figure 2.2).

We next turn to the specific case used in this thesis (DCM for cross-spectral
densities, (Friston et al., 2012; Moran et al., 2009) and provide an overview of
parametric empirical Bayes – a relatively recent enhancement of DCM that can
be used to infer the commonalities and differences across subjects (Friston et al.,
2016). Henceforth, we restrict our elaboration to DCM for EEG.
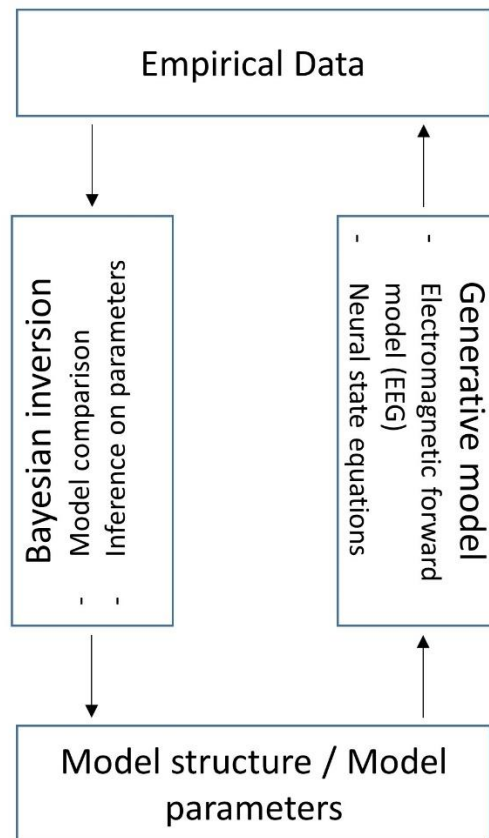
Figure 2.2. General framework for dynamic causal modeling. The modelled
neuronal dynamics are linked with specific features of the measured data using a
generative model. With Bayesian inversion, the posterior densities for each

parameter, in addition to an approximation of the log model evidence, are obtained. This enables inference on, for example, connection strengths and the most likely model structure given the observed data.

2.5.1 DCM for cross-spectral densities (CSD)

In DCM for CSD, the observed cross-spectral densities in the EEG data are explained by a generative model that combines a biologically plausible neural mass model with an electrophysiological forward model mapping the underlying neural states to the observed data. Each node in the proposed DCM models – that is, each electromagnetic source – consists of three neural subpopulations, each loosely associated with a specific cortical layer; pyramidal cells, inhibitory interneurons and spiny stellate cells (ERP model; Moran, Pinotsis & Friston, 2013). DCM does not simply estimate the activity at a particular source at a particular point in time – instead, the idea is to model the source activity over time, in terms of interacting inhibitory and excitatory populations of neurons.

The subpopulations within each node are connected to each other via *intrinsic* connections, while nodes are connected to each other via *extrinsic* connections. Three types of extrinsic connections are defined, each differing in terms of their origin and target layers/subpopulation: forward connections targeting spiny stellate cells in the granular layer, backward connections targeting pyramidal cells and inhibitory interneurons in both supra- and infragranular layers, and lateral connections targeting all subpopulations. This laminar specificity in the extrinsic cortical connections partly reflects the hierarchical organisation in the

brain. Generally speaking, the backward connections are thought to have more inhibitory and largely modulatory effect in the nodes they target (top-down connections), while forward connections are viewed as having a strong driving effect (bottom-up; Salin & Bullier, 1995; Sherman & Guillery, 1998).

The dynamics of hidden states in each node are described by second-order differential equations (state equations), which depend on both, the parametrised intrinsic and extrinsic connection strengths. This enables the computation of the linear mapping from the endogenous neuronal fluctuations to the EEG sensor spectral densities. In turn, this enables us to model differences in the spectra as a result of changes in the underlying neurophysiological parameters describing the intrinsic and extrinsic connectivity of coupled neuronal populations.

2.5.2 Parametric empirical Bayes

In DCM, a variational Bayesian scheme called Variational Laplace is used to approximate the conditional or posterior density over the parameters given by the model inversion process, by maximizing a lower bound (the negative free energy) on the log-evidence (Friston et al., 2007). The parametric empirical Bayes (PEB) framework is a relatively recent extension to the DCM procedure used, for example, to infer the commonalities and differences across subjects (Friston et al., 2016). Briefly, the subject-specific parameters of interest (here, effective connectivity between nodes in a DCM model) are taken to the group-level and modelled using a General Linear Model (GLM), partitioning the between-subject variability into designed effects and unexplained random effects

captured by the covariance component. The focus is on using Bayesian model reduction (BMR) – a particularly efficient form of Bayesian model selection (BMS) – to enable inversion of multiple models of a single dataset and a single hierarchical Bayesian model of multiple datasets that conveys both the estimated connection strengths and their uncertainty (posterior covariance). As such, it is argued that hypotheses about commonalities and differences across subjects can be tested with more precise parameter estimates than with traditional frequentist comparisons (Friston et al., 2016).

A particular advantage of PEB is that as part of the BMR process – when no strong a priori hypotheses about the model structure exist, as in the present study – a greedy search can be used to compare the negative free energies for the reduced models, iteratively discarding parameters that do not contribute to the free energy (originally 'post-hoc DCM analysis', Friston & Penny, 2011; Rosa, Friston & Penny, 2012). The procedure stops when discarding any parameters starts to decrease the negative free energy, returning the model that most effectively trades-off goodness of fit and model complexity in explaining the data. Last, a Bayesian Model Average (BMA) is calculated over the best 256 models weighted by their model evidence (from the final iteration of the greedy search). For each connection, a posterior probability for the connection being present vs. absent is calculated by comparing evidence from all the models in which the parameter is switched on versus all the models in which it is switched off.

2.5.3 DCM-based leave-one-out cross-validation

By iteratively fitting a multivariate linear model to provide the posterior predictive density over the connectivity changes (as described in detail in Friston et al., 2016), the posterior belief of an explanatory variable can be evaluated for a left-out participant (leave-one-subject-out cross-validation, LOSOCV). In other words, the covariates for the left-out participant (test data) can be predicted based on a PEB model fitted to all but one subject (training data).

To apply this approach, a group-level PEB model is first fitted to the training data to model the parameters (in this example, the connection strength with two regressors – the group mean and difference, e.g. $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$). To predict the value of a particular regressor for the left-out third participant, given their estimated connection strength, a separate PEB model is then fitted to the test data. Here, the prior expectation for the value to be predicted is set to zero and the prior variance (uncertainty) is set to a multiple of the regressor's variance in the training data (for a formal definition of PEB modelling, see Zeidman et al., 2019).

A GLM can then be constructed for the left-out participant

$$\theta_i^{(1)} = [\theta^{(2)}]^T [X_i]^T + \varepsilon^{(2)}$$

where $\theta_i^{(1)}$ is the estimated connection strength for the left-out participant, i, $\theta^{(2)}$ are the group-level PEB parameters estimated from the training data, and X the design matrix with the two regressors.

Given that we know both $\theta_i^{(1)}$ and $\theta^{(2)}$, and the first element in the left-out participants' design matrix X_i (modelled with 1 for the group mean), we can then

infer for the missing element of the design matrix X_i (see Friston et al., 2016 and Zeidman et al., 2019 for details). This procedure, repeated for each participant, generates probabilities of state affiliation, which can then be used to calculate the Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values, illustrating the diagnostic ability of the covariates.

In chapters 5 and 6, DCM-based LOSOCV is used as a validation of our modelling framework; we investigate which sets of connections are predictive of the state of consciousness (chapter 5) and of the DoC group (chapter 6) in unseen data.

2.6   The scientific study of consciousness

The recent application of methods such as neuroimaging in studying consciousness have led to theories of consciousness that generally aim to offer more detailed accounts of its nature and features, than the more long-lived metaphysical theories, which aim to locate consciousness in the general ontological scheme of the universe. Physicalist theories that equate neural events with phenomenology (i.e. identity theories) provide a useful framework, for example, in the clinical context, where assessing states of consciousness in DoC patients or in patients under anaesthesia is pivotal. In recent decades, most modern physicalist theories have developed over and beyond of merely claiming that particular brain areas might give rise to conscious experience. Rather, they

focus on global patterns of underlying neuronal processing that drives consciousness forward.

Here, we provide an overview on two such theories that are relevant to the research in this thesis, namely Global (Neuronal) Workspace Theory (GNWT; Baars, 1988, 1997; Dehaene, Kerszberg, & Changeux, 1998; Dehaene & Changeux, 2011; Dehaene et al., 2011) and Integrated Information Theory (IIT; Oizumi, Albantakis, & Tononi, 2014; Tononi, 2004, 2012; Tononi, Boly, Massimini, & Koch, 2016). We will then introduce disorders of consciousness and anaesthesia as altered states of consciousness and provide a general overview on the neuroimaging results aiming to distinguish these states from normal wakefulness.

2.6.1 Current leading theories – Global Neuronal Workspace Theory

The original thesis of the Global Workspace Theory (GWT) was proposed by Baars (1988); it is a psychological construct suggesting that perceptual content (and information more generally) is processed by intrinsically unconscious, specialised, and localised modules and only becomes conscious when widely broadcasted to other processors across the brain (forming the global workspace). Here, broadcasting implies that the information is widely available to many local modules, which is hypothesised to constitute conscious experience. These local modules involve processors related to the past (memory), present (sensory input, attention), and future (values systems, motor plans, verbal report; (Mashour, Roelfsema, Changeux, & Dehaene, 2020).

43

GNWT (Dehaene et al., 1998; Dehaene & Changeux, 2011; Dehaene et al., 2011) is a further extension of the original cognitive thesis and suggests a defined brain network as the neural instantiation of the GWT. It suggests a second, reciprocally connected computational space consisting of widely distributed excitatory neurons with long-range axons, acting upon the relevant neuronal processor modules. This distributed network is suggested to receive bottom-up information from and transmit top-down information to any of the various processor modules, thereby selecting and broadcasting information. GNWT is not a localizationist approach to consciousness; rather the neuronal workspace is posited to act as a "router" associated with millions of neurons widely distributed across many brain regions (Mashour et al., 2020). The prefrontal cortex has been posited to play a key role in the global workspace because of the density of the neurons therein thought to be critical to the broadcasting process. Other areas crucial to the GNWT include the anterior temporal cortex, inferior parietal cortex, anterior and posterior cingulate cortex and precuneus.

2.6.2 Integrated Information Theory and the posterior hot zone

Unlike theories that start from neural events and try to describe their effects on conscious experience, IIT (Oizumi et al., 2014; Tononi, 2004, 2012; Tononi et al., 2016) attempts to provide a framework capable of explaining why some systems are conscious, why they feel the particular way they do, and what would it take for other systems to become conscious. Rather than starting from physical principles and arriving at consciousness, IIT aims to build an explanatory bridge

between the brain activity and subjective experience starting from the phenomenology itself (i.e. from consciousness), and by reasoning a set of essential properties (axioms) that a physical substrate would have to have in order to account for conscious experiences (postulates). The latest version of IIT proposes five axioms, suggesting that any conscious experience:

- exists to itself (each experience is intrinsically actual and real)

- is structured composition of multiple elements that are experienced at the same time (e.g. colour, shape, identity of the object)

- is specific (in a sense of being composed of specific phenomenal distinctions, thereby differentiating itself from other conscious experiences)

- is irreducible to the sum of each experience's components (integrated)

- is definite and hence, exclusive of other simultaneous conscious experiences.

Based on these five axioms, the properties of the physical substrate that are required to account for these regularities in conscious experience are proposed:

- to account for the intrinsic existence, a system must have cause-effect power upon itself

- the system must be structured as a composition of various subsets of elements constituting the system itself, that also have cause-effect power within the system

- the cause-effect structure of the system must be integrated and not reducible to the effects of its independent components. Moreover, this intrinsic irreducibility can be measured as integrated information (Φ), quantifying to what extent the cause-effect structure changes if the system is partitioned into components

- the system must specify a particular cause-effect structure differing from other possible ones which is specified over a single set of components (definite): the set over which it is maximally irreducible (Φ^{max})

Hence, IIT suggests that consciousness is a function of system's capacity for information integration; altered states such as DoC or anaesthesia are characterised by relatively lower information integration and differentiation. Although Φ^{max} is not computable for the human brain, several proxy measures inspired by the IIT have been developed (see for example Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008). Experimental evidence supporting IIT's characterisation of altered states of consciousness (i.e. a decrease in integration and an increase in modularity) has been observed in brain activity during sleep (Massimini et al., 2005; Tagliazucchi et al., 2013), anaesthesia (Casali et al., 2013; Kim et al., 2018; Monti et al., 2013), and in DoC (Achard et al., 2012; Bodart et al., 2017, 2018; Lutkenhoff, Johnson, Casarotto, Massimini, & Monti, 2020).

An obvious current question is how GNWT and IIT relate to each other. IIT itself is not tied down to any particular brain regions – rather than providing the neural correlates, it offers an explanation of the underlying physical mechanism

from which conscious experiences are argued to arise. Hence, IIT and GNWT are not mutually exclusive; in fact, both theories propose similarly, although not identically, that broad simultaneous activation of cortical areas is essential for conscious experiences (Frigato, 2021). Nonetheless, a good deal of converging evidence have suggested a key role for areas in the posterior cortex in sleep (Lee et al., 2019; Siclari et al., 2017), general anaesthesia (Alkire, Hudetz, & Tononi, 2008; Gaskell et al., 2017), and in DoC patients (Vanhaudenhuyse et al., 2010; Wu et al., 2015). Such evidence have led Tononi and collaborators to consider the posterior hot zone (temporo-parieto-occipital areas, figure 2.3) to be essential for consciousness (Koch et al., 2016a; Boly et al., 2017). Rather than essential for consciousness, activity in frontal areas are viewed to be involved with activities subsequent to conscious experience (such as selective attention, working memory, task reporting, and task monitoring).
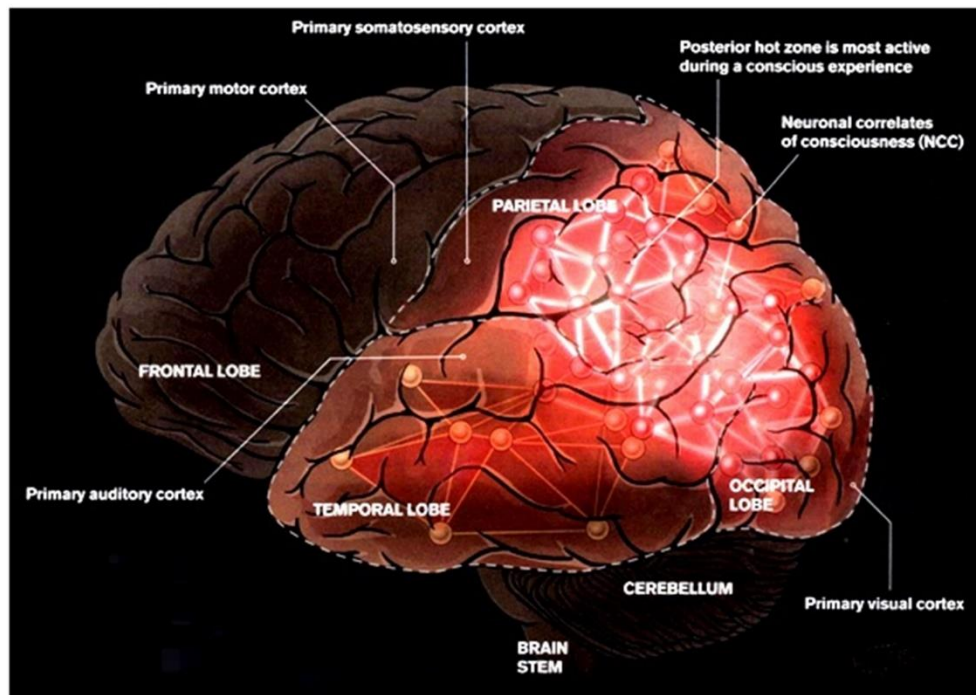
Figure 2.3. Posterior areas (highlighted) that constitute the hot zone for consciousness. Image reproduced from Koch (2018).

This hot zone hypothesis of phenomenal consciousness is an important focus of scientific investigation in this thesis. Specifically, the research in chapter 5, and to a lesser degree in chapter 6, investigates changes in effective connectivity and the power of the posterior hot zone and long-range frontoparietal connections in predicting changes in states of consciousness induced by anaesthesia and brain injury.

2.6.3 Altered states of consciousness: Disorders of consciousness

How would one go about studying the underlying neural processes that are necessary for consciousness? A reasonable starting point is to investigate the brain dynamics in transitions of consciousness: how specific properties change between different states of consciousness. A common approach is to compare normal wakefulness with states of apparent unconsciousness, such as sleep (Massimini et al., 2005; Siclari et al., 2017; Tagliazucchi & van Someren, 2017) and anaesthesia (Boveroux et al., 2010; Monti et al., 2013; see Bonhomme et al., 2019 for a review), or to contrast different states of disorders of consciousness (Gosseries et al., 2011; Laureys et al., 1999; Vanhaudenhuyse et al., 2010; see Gosseries, Di, Laureys, & Boly, 2014 for a review). Observations from sleep, DoC patients, and anaesthesia have pushed the study of consciousness beyond the study of conscious wakefulness suggesting an increasingly complex spectrum of consciousness. For example, the level-based framework for conceptualising states of consciousness was derived from the clinical literature on DoC and further developed into a two-dimensional representation with awareness of environment and of self (i.e. content of consciousness) and arousal/wakefulness (one's vigilance) as two distinct components of conscious state (figure 2.4; Laureys, 2005). More recent suggestions have called the concept of level of consciousness into question and introduced increasingly complex, multidimensional representations of the different states of consciousness (Bayne, Hohwy, & Owen, 2016).
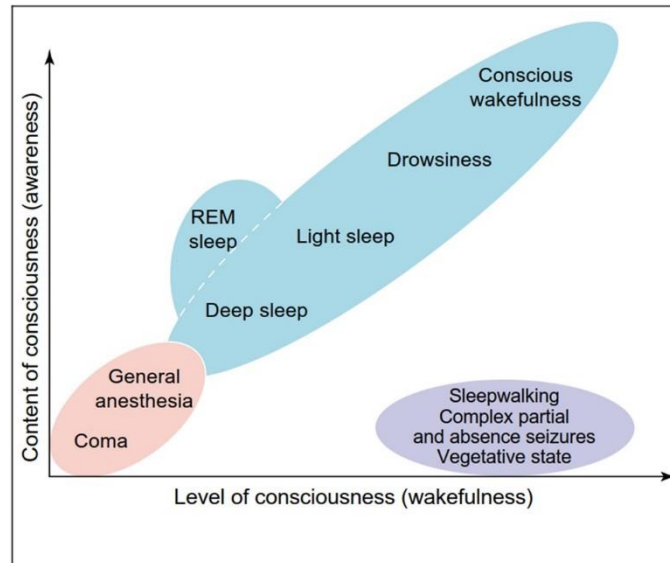
Figure 2.4. Simplified illustration of the two major components of consciousness: the level of awareness (i.e. content of consciousness) and the level of arousal/wakefulness. Reproduced from Laureys (2005).

Nevertheless, DoC – including states such as coma, unresponsive wakefulness syndrome (UWS), minimally conscious state (MCS), and locked-in syndrome (LIS) induced by either non-traumatic causes (e.g. haemorrhage, ischaemia, illness) or by traumatic brain injury (Edlow, Claassen, Schiff, & Greer, 2021) – can be characterised by alterations in arousal and/or awareness. Coma is defined as the complete absence of arousal and awareness (Teasdale & Jennett, 1974), while the UWS is characterised by preserved arousal in the absence of behavioural signs of awareness (Laureys et al., 2010). In contrast, in MCS, patients show fluctuating and incomplete awareness with preserved arousal (Giacino et al., 2002). MCS state has been further divided into MCS- and MCS+,

with the latter condition characterised by command following, intelligible verbalisation or gestural (or verbal yes/no responses) to spoken or written questions (Bruno, Vanhaudenhuyse, Thibaut, Moonen, & Laureys, 2011). LIS patients on the other hand, are characterised by undisturbed cognitive functions with total immobility except for vertical eye movements and blinking (Bauer, Gerstenbrand, & Rumpl, 1979). Recently, the concept of cognitive motor dissociation (CMD) – also known as 'covert consciousness' – has been added to the diagnostic scheme of patients with DoC (Schiff, 2015). This state is characterised by volitional brain activity typically detected by task-based functional MRI or EEG in patients in coma, UWS, or MCS.

In the absence of robust neurophysiologic markers, DoC are assessed with standardised behavioural assessments in clinical examination (Giacino et al., 2009). Currently the most sensitive behavioural scale to disentangle the DoC groups from each other is the behaviour-based Coma Recovery Scale-Revised (Giacino, Kalmar, & Whyte, 2004; Seel et al., 2010). Previous research has suggested, however, that the brain states in DoC can also be distinguished with neuroimaging, including PET (Aubinet et al., 2020; Laureys et al., 1999; Stender et al., 2014, 2016), fMRI (Crone et al., 2014; Vanhaudenhuyse et al., 2010), and EEG (Chennu et al., 2017; King et al., 2013; Sitt et al., 2014).

Early PET-imaging studies showed significant alterations in DoC patients, especially in the prefrontal, premotor, and parietotemporal association areas, in the precuneus, and in the thalamus (Laureys et al., 1999; Laureys, Faymonville, & Lamy, 2000; Laureys, Owen, & Schiff, 2004). More recent structural and

functional neuroimaging studies have consistently highlighted the role of forebrain mesocircuit (Fridman et al., 2014; Schiff, 2010) and the frontoparietal network (Vanhaudenhuyse et al., 2010; Wu et al., 2015), specifically the DMN (Annen et al., 2018; Boly et al., 2009; Fernández-Espejo et al., 2012; Guldenmund et al., 2016; Soddu et al., 2012), in the restoration of cerebral activity during recovery from DoC. Network properties have also shown to be altered in several regions associated generally with consciousness; the degree of highly connected regions ('rich-club' of hubs) is decreased in DoC patients (particularly, in medial parietal and frontal regions, and in the thalamus; Crone et al., 2014). In addition to the differences in "overt" consciousness observed in patients with differing levels of DoC, a number of studies have reported "covert" voluntary brain activity in some seemingly unconscious patients (Bodart et al., 2017; Claassen et al., 2019; Chennu et al., 2017; Cruse et al., 2011; Lechinger et al., 2013; Monti et al., 2010; Owen et al., 2006; Owen & Coleman, 2008; Schnakers et al., 2015). Such voluntary brain activity may indicate residual covert consciousness in some patients, which may evade even expert bedside behavioural examination (Edlow, 2018).

Importantly, differences in brain PET metabolism between the two hemispheres have been indicated for MCS- and MCS+ patients (Bruno et al. 2012) and recently for potentially "covertly aware" UWS patients (Thibaut et al., 2021). In addition, hemispheric differences in EEG brain connectivity have been reported between UWS and MCS patients (Lehembre et al. 2012). However, the level with which other DoC states may or may not be distinguished from hemispheric-specific EEG connectivity is unknown, especially with "covertly

aware" patients. Furthermore, whether effective connectivity modulations can identify such covert awareness in patients has not been studied to date. In chapters 4 and 6, we work towards addressing these gaps using both functional and effective connectivity; we investigate the differences between DoC patients and the extent to which they may or may not be predictive of the individual DoC states.

2.6.4 General anaesthesia

General anaesthesia – as defined by controlled and reversible induction of loss of consciousness, analgesia, and muscle relaxation – can induce altered states of consciousness (Mashour, 2004). As opposed to the previously proposed notions that general anaesthesia simply shuts the brain off, it is now clear that several brain functions may be retained by patients until very high concentrations of anaesthetic agents (Sleigh, Warnaby, & Tracey, 2018). At a small dose, anaesthetics suppress thinking, focused attention, and working memory. With increasing dosage, they cause voluntary responsiveness and consciousness to fade, ultimately producing a fully unconscious state. From a research perspective, anaesthetic agents thus provide a powerful tool to study consciousness. The ability to induce reversible alterations in the state of consciousness has led to evidence of several anaesthetic-induced changes in brain connectivity, network topology, and spatio-temporal dynamics. In turn, this knowledge has contributed towards understanding the underlying mechanisms of consciousness itself

(Bonhomme et al., 2019; Brown, Lydic, & Schiff, 2010; Guldenmund et al., 2016).

How exactly general anaesthesia abolishes consciousness is still not comprehensively explained, and different anaesthetic agents are known to have different effects on brain function. One of the most widely studied and used anaesthetic agents is the γ-amino-butyric acid (GABA) neurotransmission promoting agent propofol (Bonhomme et al., 2019; Jevtovic-Todorovic, 2016). Propofol has been demonstrated, for example, to have region-specific and dose-dependent effects on brain activity with reductions in the thalamus, the cuneus and precuneus, the posterior cingulate cortex, and the angular gyri (Alkire et al., 1995; Fiset et al., 1999). More recent studies have suggested an increase of slow-wave activity (Murphy et al., 2011), interference with activity in frontoparietal associative (Mhuircheartaigh et al., 2010) and higher-order networks such as the DMN (Stamatakis, Adapa, Absalom, & Menon, 2010), the salience network (Guldenmund et al., 2013), and the central executive network (Boveroux et al., 2010). In contrast, primary sensory networks such as the visual and auditory networks seem to stay relatively intact from the effects of anaesthetics (MacDonald, Naci, MacDonald, & Owen, 2015).

The increased slow-wave activity (SWA) under anaesthesia has been suggested as an important manifestation of perception loss, and as such, suggested as a possibly proxy for anaesthetic-induced unconsciousness (Mhuircheartaigh, Warnaby, Rogers, Jbabdi, & Tracey, 2013; Warnaby, Sleigh, Hight, Jbabdi, & Tracey, 2017). Interestingly, previous research has also

suggested a link between the level of the pre-anaesthesia brain connectivity and individual responses to propofol. Specifically, network properties at alpha-band (Chennu, O'Connor, Adapa, Menon, & Bekinschtein, 2016), inter-frontoparietal network connectivity (Deng, Cusack, & Naci, 2019), and dynamic frontoparietal alpha-connectivity (Zhang et al., 2020) have all been associated with susceptibility to propofol.

In chapter 3, we explore this research question further, and provide novel evidence for SWA – particularly for the saturation of the SWA activity (Mhuircheartaigh et al., 2013) – as a proxy for unconsciousness by investigating the link between pre-anaesthesia brain connectivity and the saturation stage. We explore the possible link between baseline connectivity and behavioural measures of unconsciousness. In chapter 5, we further investigate the changes in effective connectivity following propofol-induced LOC. In doing so we contribute to the debate around frontal vs. posterior loci of NCC by providing novel computational evidence of intra- and inter-resting state network (RSN) changes in key three key RSNs; the default mode network, the salience network (SAL), and the central executive network (CEN). We use a novel methodological combination of DCM for resting EEG cross-spectral densities (Friston et al., 2012; Moran et al., 2009) and parametric empirical Bayes (Friston et al., 2016; see further details in section 2.5). This allows us to better estimate the model parameters, and consequently, evaluate the role of different subgroups of specific intra- and inter-RSN connections in propofol-induced loss of consciousness.

2.6.5 Unresponsiveness and loss of consciousness

It should be noted that with comparisons of, for example, anaesthetic-induced LOC or DoC states, we rely only on subjective reports of the patients or behavioural measures to establish presence or absence of conscious processing. In other words, as we cannot directly measure their level of consciousness, the subjective reporting and/or behavioural responsiveness (or lack thereof) is used as a proxy for their phenomenological experience of consciousness. This proxy, however, has its problems; the presence of responsiveness does not necessarily imply consciousness (or at least not awareness). For example, patients suffering from blindsight are reported to respond to stimuli despite denying any awareness of it (Cowey, 2010; Sahraie, Hibbard, Trevethan, Ritchie, & Weiskrantz, 2010; Sanders, Warrington, Marshall, & Wieskrantz, 1974, but see also Phillips, 2021). Similarly, split-brain patients reportedly respond to stimuli presented in their right visual field verbally and correctly with their right hand. Conversely, when a stimulus is presented in the left visual field, and thus, processed by the right hemisphere, the patients verbally state that they saw nothing while still identifying the object correctly with their left hand only (Gazzaniga, Bogen, & Sperry, 1962; Sperry, 1984; Wolman, 2012, but see also Pinto et al., 2017).

The case for behavioural unresponsiveness not necessarily implying unconsciousness is even stronger (Sanders, Tononi, Laureys, & Sleigh, 2012). For example, every night in sleep we experience dreams with varying content while we are – to an extent – unresponsive to, and disconnected from, the external world (Nir & Tononi, 2010). More crucially to this thesis, although apparently unconscious due to general anaesthesia, dreams or dream-like

56

experiences are often reported by the patients (Leslie, 2017; Leslie et al., 2009; Noreika et al., 2011; Radek et al., 2018) and patients may report waking up during the surgery unbeknownst to the clinical staff (Goddard & Smith, 2013). Moreover, using the isolated forearm technique, a significant portion of patients have been demonstrated to be able to communicate during general anaesthesia (Sanders et al., 2012).

With DoC patients, a number of studies have reported "covert" voluntary brain activity in some seemingly unconscious patients (Bodart et al., 2017; Claassen et al., 2019; Chennu et al., 2017; Cruse et al., 2011; Lechinger et al., 2013; Monti et al., 2010; Owen et al., 2006; Owen & Coleman, 2008; Schnakers et al., 2015; Vanhaudenhuyse et al., 2018). Hence, it seems it may be possible for consciousness to occur in these clinical settings without any overt responsiveness detectable by bedside behavioural evaluation (Edlow, 2018).

Nevertheless, responsiveness and lack thereof undeniably provides a useful window into at least a part of the dimension extending from wakeful consciousness to unconsciousness. Indeed, as we are only privy to our own subjective experience of consciousness, outside the context of philosophical thought experiments, assessments of conscious experiences are commonly inferred from behavioural activity and responsiveness. This is especially true in the clinical setting, where behavioural responsiveness is the first and most common assay of consciousness.

Moreover, responsiveness objectively measures a neural system's ability to fully engage in the process of coordinating and producing the necessary

behaviours; at some level between unresponsiveness and the induction of flat EEG (indicative of cessation of brain activity, i.e. brain death; Wijdicks, 2001), consciousness vanishes. Subjectively, LOC does not seem to be a binary event, but rather a gradual process with intermediary stages including, for example, narrowed attention, decrease of memory, impaired cognition, and lower self-estimated level of consciousness (Esaki & Mashour, 2009; Vaitl et al., 2005). Assessments of behavioural responsiveness approximate these intermediary stages by probing the particular stage where responsiveness is lost. Finally, behavioural markers of consciousness in general have been shown to be useful in practice; for example, the Coma Recovery Scale has been shown to correlate with the prognosis of DoC patients (Bruno et al., 2012; Giacino, Fins, Laureys, & Schiff, 2014). Therefore, responsiveness provides a valuable marker for estimations of consciousness.

In this thesis, loss of behavioural responsiveness (LOBR) is used as a general marker for unconsciousness. In an attempt to control the dissociation between LOBR and LOC, in case of anaesthesia, the participants were asked if they recalled any dreams or other experiences, which none of them did. With the DoC patients, we distinguished potentially covertly aware patients based on their PET-metabolism, and separated them from unresponsive patients during model training phase. We then introduced them as an unseen dataset of (possibly) conscious patients in the test phase. We therefore follow the typical convention in the literature and refer to the behaviourally unresponsive state during anaesthesia as LOC, while using the behaviour-based diagnostic labels for the corresponding DoC groups.

2.7 Conclusions

In this chapter, we provided an overview on the methods used in this thesis, starting with EEG analysis and ending with functional and effective connectivity. We then introduced two relevant theories of consciousness before focusing on the altered states of consciousness analysed in this thesis – anaesthetic-induced sedation and disorders of consciousness. This chapter hence provides the conceptual framework for the research presented in this thesis.

3. Functional connectivity in propofol-anaesthesia

This chapter explores the correlation between pre-anaesthesia resting brain connectivity and individual measurements of propofol-concentrations in blood plasma. We employ behavioural measures of responsiveness and a previously proposed electroencephalography (EEG) marker for anaesthetic-induced loss of consciousness (LOC), namely the slow-wave activity saturation (SWAS). Specifically, we combine resting state EEG recorded from healthy volunteers under propofol-induced sedation with individual measurements of propofol-concentrations in blood plasma, alongside behavioural measures of responsiveness. We assess the participants' spectral power and functional connectivity during a pre-anaesthetic baseline period and correlate those with individual variation in propofol-concentrations, time needed to reach SWAS, level of SWA-power, and time needed to reach loss of behavioural responsiveness (LOBR). In doing so, we hypothesise and provide evidence for SWAS as a distinct and individualised index for propofol-induced loss of consciousness, that is linked with pre-anaesthesia baseline brain connectivity.

3.1 Introduction

Understanding the complex interactions of neuronal activity necessary to explain consciousness and its loss is a grand challenge for modern neuroscience.

In recent years, theoretical advances have proposed a pivotal role for the balance between integrated and differentiated neuronal activity in alterations of the conscious state (Oizumi, Albantakis, & Tononi, 2014; Tononi, 2004, 2012; Tononi, Boly, Massimini, & Koch, 2016). Nevertheless, accurately tracking the underlying changes in brain dynamics have remained a key research challenge.

General anaesthetics are powerful and commonly used tools in this context; the ability to induce reversible states of unconsciousness with anaesthetic agents has led to evidence of several changes in functional and effective brain connectivity, network topology, and spatio-temporal dynamics contributing towards understanding the underlying mechanisms of consciousness (Bonhomme et al., 2019; Brown, Lydic, & Schiff, 2010). However, despite the wide use and a set of effects common to many anaesthetic agents, a comprehensive picture of the key factors and changes influencing the brain during general anaesthesia is not yet known. Moreover, tracking brain activity to accurately assess the depth of anaesthesia in an individual has not yet been fully achieved; for example, surface EEG, a relatively easy and inexpensive method long known to index changes induced by anaesthetic agents in the brain dynamics (Gibbs, Gibbs, & Lennox, 1937), is still not universally used in clinical practice.

The lack of robust EEG markers capable of accurately tracking the loss and recover of consciousness is partly due to individual variability in susceptibility to anaesthetic agents and dosage (Palanca, Mashour, & Avidan, 2009). Indeed, many factors have been shown to have an effect on the patient's response to anaesthesia: for example, with propofol, demographic characteristics such as

weight, age, and sex have been associated with differences in the response to the drug (Gambús & Trocõniz, 2015; Schnider et al., 1999).

Furthermore, the large inter-individual differences in susceptibility to propofol remain even when such demographic factors are taking into consideration (Kaskinoro et al., 2011). For example, patients with higher preoperative anxiety scores require higher doses of propofol to reach the same level of sedation as patients with lower preoperative anxiety (Hong, Jee, & Luthardt, 2005; Kil et al., 2012). Likewise, the required dose to maintain sedation decreases with a decline of preoperative cognitive state in elderly patients (Laalou et al., 2010). These results suggest that the functional state of the brain may be pivotal factor in individual propofol susceptibility.

Indeed, a growing body of evidence suggest suppression of functional brain connectivity as a common mechanism for various anaesthetic agents in inducing unconsciousness (Lee et al., 2015; Schröter et al., 2012; Warnaby et al., 2016). Specifically, break-down of thalamo-cortical connections and disrupted frontoparietal networks (Boveroux et al., 2010; Malekmohammadi, Price, Hudson, DiCesare, & Pouratian, 2019; Schrouff et al., 2011), disruptions in frontal areas (Guldenmund et al., 2016), and diminished top-down frontoparietal connectivity (Dehaene & Changeux, 2011; Changeux, 2012; Lee et al., 2009, 2015) have been suggested as candidate mechanisms for explaining LOBR in anaesthetic state. Importantly, previous studies have reported an association between pre-anaesthesia functional connectivity and individual differences in responses to propofol. Specifically, research has indicated a link between

network properties at alpha-band and susceptibility to propofol (Chennu, O'Connor, Adapa, Menon, & Bekinschtein, 2016), with pre-anaesthesia inter-frontoparietal network connectivity and reaction time under sedation (Deng et al., 2019), and more recently, between dynamic frontoparietal alpha-connectivity and time until loss of consciousness (as measured with bispectral index, BIS; Zhang et al., 2020).

In addition to BIS, slow-wave activity (SWA; 0.5 – 1.5 Hz) – a characteristic waveform seen in both anaesthesia and sleep (Murphy et al., 2011) – has been suggested as an important manifestation of perception loss indicating individual's depth of anaesthesia (Mhuircheartaigh, Warnaby, Rogers, Jbabdi, & Tracey, 2013; Warnaby, Sleigh, Hight, Jbabdi, & Tracey, 2017). In their pivotal study, Mhuircheartaigh et al. (2013) found, that after perception loss, each individual's slow-wave activity reached a saturation point, after which it remained constant despite increasing anaesthetic concentrations. Their simultaneous functional magnetic resonance imaging indicated that typical thalamo-cortical responses to nociceptive and auditory inputs were abolished, leading them to hypothesise SWA saturation (SWAS) as a potential individualised indicator for depth of anaesthesia, and thus, an index for anaesthetic-induced LOBR.

Ergo, an association has been suggested between baseline-functional connectivity and the state of consciousness on the one hand, and between consciousness and SWA-power – specifically the SWAS – on the other (Chennu et al., 2016; Deng et al., 2019; Mhuircheartaigh et al., 2013; Warnaby et al., 2017; Zhang et al., 2020). However, to what extent baseline connectivity is

63

predictive of SWAS, if any, remains an open question. To address this gap in knowledge, we investigated the potential association between EEG power and functional connectivity at pre-anaesthesia baseline and SWAS. Specifically, we combined the measurement of resting state EEG from healthy volunteers sedated with propofol with measurements of propofol concentrations and behavioural measures (e.g. LOBR). Employing functional EEG tools, we assessed the participants' spectral power and functional connectivity at the level of electrodes during a pre-anaesthetic baseline period and linked them with individual variation in propofol concentrations, time needed to reach the saturation, level of SWA-power, and time of LOBR. Based on previous research (Chennu et al., 2016; Deng et al., 2019; Zhang et al., 2020), we hypothesised, that if indeed SWAS is a reliable index for LOC, baseline power and connectivity should predict the SWA-power at saturation and time and the amount of propofol needed for reaching saturation point.

3.2   Methods

3.2.1   Data acquisition and experimental design

The data used in the present work were acquired from a previous propofol-anaesthesia study (Mhuircheartaigh et al., 2013). This study involved two separate EEG acquisitions. Both were approved by the Local Research Ethics Committee and written consent was obtained from all volunteers. The first acquisition recorded EEG data (in a laboratory setting) from 16 healthy volunteers (8 males, 8 females with American Society of Anesthesiologists

(ASA) physical status grade I or II; age 28.6 ± 7.0 (SD) years; range 19 to 43 years), while the second session acquired simultaneous fMRI-EEG data from a subset of 12 of the 16 volunteers. In the following work, only the EEG data from the first acquisition ($N = 16$) was used.

During the study, participants experienced a 10-minute resting period with eyes closed (prior to drug administration), followed by an ultraslow induction of propofol to loss of consciousness. A target-controlled intravenous propofol infusion was used with step increases of 0.2 μg/ml to achieve a maximum effect site concentration (ESC) of 4 μg/ml over 48 minutes. During this induction period, a paradigm of nociceptive laser stimuli, 1 kHz tones, and auditory word discrimination tasks were presented. The word tasks contained a list of 200 single-syllable words which were derived from MRC Psycholinguistics Database (Machine Usable Dictionary v2.0) with a familiarity of 488 ± 99 (SD) and concreteness of 438 ±120 (SD). To determine the loss and recovery of behavioural responsiveness used here (LOBR and ROBR, respectively), motor responses (button presses) to the auditory word discrimination task were used. The details of the stimulation paradigm are not described here as those data were not analysed in the present study, but further information can be found from Mhuircheartaigh et al. (2013). After reaching the peak-dose, stimulation was turned off and participants remained fully anesthetised for 10 minutes (peak-unresponsiveness; P-UNR). This was followed by the final phase, during which the propofol infusion was turned off, and subjects were allowed to emerge from unconsciousness naturally. Alongside, the same sensory stimulation and behavioural response paradigm was resumed.

3.2.2 EEG data collection and preprocessing

The EEG data were recorded using a 32-channel EEG cap (BrainCap MR, Easycap GmbH, Gilching, Germany) with an MR-compatible amplifier system (BrainAmp MRplus, Brain Products GmbH, Gilching, Germany) using FCz as a reference electrode. EEG was measured in microvolts (μV), sampled at 5 kHz, and all electrode impedances were kept below 5 kΩ throughout the experiments. The re-referencing to common average, downsampling to 1 kHz, and identification of eye-blink artefacts was carried out by the authors of the original experiments with an automated algorithm with BrainVision Analyzer version 2.0 (Brain Products GmbH) that parsed the VEOG channel and with custom written MATLAB code (The Mathworks Inc.). Artefacts caused by eye blinks were confirmed by both visual inspection and using an independent component analysis (ICA), and removed from the EEG channels. Lastly, the artefact-free data were confirmed by a visual inspection. In the present study, for preprocessing of the baseline- and P-UNR and for data analysis, a custom pipeline was implemented using MATLAB scripts (v.2017a) that utilised the MOHAWK-pipeline (Chennu, 2018. *MOHAWK* v.1.2. Available from: https://github.com/srivaschennu/MOHAWK). The pipeline uses functions from both EEGLAB (Delorme & Makeig, 2004) and FieldTrip (Oostenveld et al., 2011) to preprocess the data, and to calculate, analyse, and visualise EEG-power and scalp-level brain connectivity. Specifically, the connectivity between EEG channels was indexed by the debiased weighted phase lag index (dwPLI; (Vinck

et al., 2011). For detailed description of the pipeline, see Chennu et al. (2016) and Chennu et al. (2017).

Data were first imported to MATLAB, downsampled to 250 Hz, and filtered between 0.5 – 30 Hz. The raw recording lengths for the baseline condition varied slightly (mean = 9.29 min., SD = 0.52), while less for the P-UNR condition (mean = 9.48 min., SD = 0.36). Next, both conditions, baseline and P-UNR, were segmented into 10-second epochs, and baseline-corrected relative to the mean voltage over the entire epoch. The epochs were checked for excessively noisy electrodes and segments by calculating their normalised variance (with thresholds of 500μV and 250μV for channels and segments, respectively), and then manually rejected or retained based on visual inspection. In the baseline condition, a mean of 1 epoch was identified as noisy (from 6 out of 16 participants) and removed. The mean number of retained epochs was 56 (range 49-63). In the P-UNR condition, a mean of 3 epochs were identified as noisy (from 7 out of 16 participants) and removed. The mean number of retained epochs was 57 (range 52-60). Visual inspection of the data during the unresponsiveness condition revealed that 4 participants demonstrated brief periods of minor burst suppression. As these periods were short and did not meet the clinical criteria for burst suppression, these segments were retained for subsequent data analysis (Fisch & Spehlmann, 1991).

Three noisy EEG channels from one participant and 1 channel from another participant were identified as noisy in the baseline condition. In the P-UNR condition, a mean of 3 channels (from 7 out of 16 participants) were identified as

noisy. These channels were interpolated from neighbouring channels using spherical spline interpolation, and each channel was then re-referenced to the average of all channels.

3.2.3 Spectral power and connectivity analysis

Spectral power values between 0.5-30 Hz and a resolution of 0.1 Hz were estimated from the cleaned EEG-datasets, using a multitaper method with five Slepian tapers. At each channel, the absolute power magnitude within the canonical frequency bands – delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (15-30 Hz) – were calculated for each participant. Further, in the P-UNR condition the slow-wave activity power (SWA; 0.5-1.5 Hz) was calculated. For all bands, the absolute power values were also converted to relative power contributions to the total power within the 0.5-30 Hz range.

Cross-spectral densities between every pair of channels (at frequency bins of 0.06 Hz) were used to calculate the debiased weighted phase lag index (dwPLI); a measure of functional connectivity that has been shown to be robust against volume conduction, uncorrelated noise and inter-subject variations in sample size (Vinck et al., 2011). The underlying idea is – as is the case with PLI (Stam et al., 2007) – to consider only the phase angle distributions predominantly on either the positive or negative side of the imaginary axis of the complex plane. Hence, dwPLI disregards the activity detected at 0° or 180° phase differences between the electrode pairs (i.e. signals), therefore effectively removing the largest contributions of volume-conducted activity (ibid.; Vinck et al., 2011). With PLI,

the extent to which phase leads or lags are non-equiprobable is measured (irrespective of the magnitudes of the phase angle differences). With wPLI, estimation error due to noise is further alleviated by weighting the observed contributions of phase angle differences (leads and lags) by the magnitude of the imaginary component of the cross-spectrum. The phase leads and lags close to $0°$ or $180°$ are considered to contribute more heavily to noise, rather than to true interaction between brain sources (see section 2.4 for more detailed description).

dwPLI has previously been utilised and tested specifically in the context of anaesthetic-driven unconsciousness (e.g. Chennu et al., 2016; Kim et al., 2016), as well as in the context of disorders of consciousness (Chennu et al., 2014; 2017). Here, following established methodology, dwPLI values across all channels within each band were used to represent the connectivity between channel pairs. Consequently, for each subject and for each frequency band, we obtained symmetric 32x32 dwPLI connectivity matrices. For further statistical analyses, we analysed connectivity values thus estimated across all channel pairs, and between specific, hypothesis-driven subsets of electrodes in frontal, posterior, and frontoposterior regions (see figure 3.1).
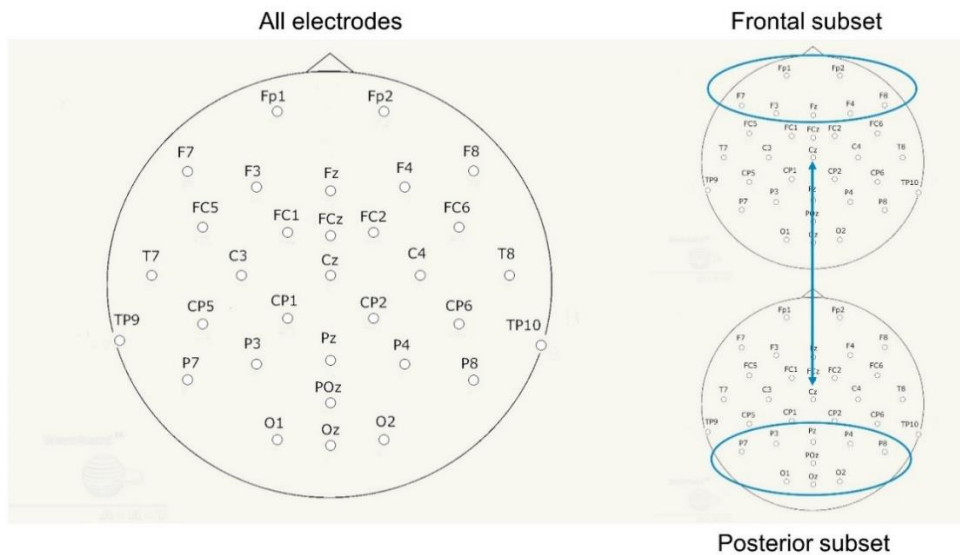
Figure 3.1. The electrode layout for all electrodes and for frontal and posterior subsets. The frontoposterior subset was formed accounting for the connectivity between frontal and posterior subsets.

3.2.4 Covariates and statistical analysis

In addition to the EEG datasets, we obtained a MATLAB dataset from the authors of the original study (Mhuircheartaigh et al., 2013), containing the covariates shown in table 3.1. Furthermore, we obtained the ages of the participants, and the State-Trait Anxiety Inventory scores (STAI) for each participant prior to the experiment were provided.

Table 3.1. Covariates for the induction and peak-unresponsiveness condition (P-UNR). The data for the induction condition were obtained from the original authors, while the covariates for the P-UNR condition were calculated during the present

work. The multivariate regression model with the P-UNR condition covariates were calculated to check the results obtained were in the same direction as with the induction condition. SWA – slow-wave activity, TSWAS – onset of slow-wave activity saturation, LOBR – loss of behavioural responsiveness.

Covariates	Induction-condition	P-UNR
SWA (power; dB)	Yes	Yes
Max. SWA	Yes	Yes
Min. SWA	No	Yes
Mean SWA-power	Yes	Yes
SWA at TSWAS	Yes	No
SWA at LOBR	Yes	No
Effect-site conc. at TSWAS	Yes	No
Effect-site conc. at LOBR	Yes	No

We used the SWA-power measures, and the LOBR and TSWAS time-points with the ESC at those two time-points as our dependent measures in a common multivariate linear regression model, with dwPLI measures calculated from the baseline as our independent variables (Matlab; mvregress-function). The multivariate regression model expresses a *d*-dimensional continuous response vector as a linear combination of predictor terms plus a vector of error terms with a multivariate normal distribution. Our response vector for the dependent variables contained eight terms:

- Maximum SWA-power

- SWA-power at TSWAS

- Mean SWA-power

- SWA-power at LOBR time-point

- LOBR time-point (s.)

- TSWAS time-point (s.)

- ESC at TSWAS time-point

- ESC at LOBR time-point

That is, with $n = 16$ participants, the dimension of the response matrix $\mathbf{Y}_i = (y_{i1}, y_{i2}, \ldots, y_{id})$ denote the response vector for observation $i = 1, 2, \ldots, 16$. The second dimension of the response matrix, $d = 8$, corresponded to the 8 dependent variables listed above.

The independent variables in the multivariate regression model, i.e., predictors, were the following:

- Average baseline dwPLI (all channels; delta-, theta-, alpha-, and beta-band)

- Avg. baseline dwPLI (frontal channels; delta-, theta-, alpha-, and beta-band)

- Avg. baseline dwPLI (posterior channels; delta-, theta-, alpha-, and beta-band)

- Avg. baseline dwPLI (frontoposterior channels; delta-, theta-, alpha-, and beta-band)

As we allowed different slopes and intercepts for the terms, each predictor $(X_1, ..., X_4)$ was an n-element cell array of d-by-K design matrices. With each model, there were $K = 40$ regression coefficients to estimate: eight intercept terms and 8 x 4 slope terms. Thus, we fitted the multivariate regression model

$$y_{ij} = \alpha_j + \beta_j x_{ij} + \epsilon_{ij},$$

where $i = 1, ..., n$ and $j = 1, ..., d$, with between-response concurrent correlation

$$COV(\epsilon_{ij}, \epsilon_{ij}) = \sigma_{jj},$$

and where the d-dimensional error term follows a multivariate normal distribution,

$$\varepsilon_{ij} \sim MVN_j(0, \Sigma).$$

The function fits multivariate regression model with a diagonal error variance-covariance matrix using a normal maximum likelihood function ('mvn' as a default for non-missing responses). With the subsets, we only changed the

electrode subset from which the baseline dwPLI values were averaged (frontal, posterior, frontoposterior).

For estimating the statistical significance of the fitted model parameters, a non-parametric randomisation approach was implemented. Specifically, after estimating the regression coefficients, we randomised the dependent variables (Y_i) by shuffling the rows (corresponding to participant dimension) 100,000 times and fitting the model again after each such shuffle. The regression coefficients obtained after each shuffle iteration formed the null hypothesis distribution. From this, non-parametric p-values were calculated as a proportion of coefficients after shuffling that were more extreme (more positive or negative than the original coefficients, if the original model regression coefficients were positive or negative, respectively) than the original coefficients.

3.3 Results

3.3.1 Confirmatory hypotheses

First, we formulated a number of confirmatory hypotheses based on commonly found effects of propofol-anaesthesia:

- we expected an increase in the slow-wave delta-power when comparing baseline and P-UNR conditions

- we expected a shift from posterior alpha-power towards anterior electrodes when comparing baseline and P-UNR conditions

The results from the confirmatory hypotheses are shown in figures 3.2A and 3.2B, respectively. Further, we compared the smoothed mean SWA-measures calculated for the P-UNR condition to the original figure (Mhuircheartaigh et al., 2013; figure 3) to confirm there were no major differences. The average SWA-power percentage (SD) normalised to each individual's maximum of 0.5 – 30 Hz was 75.3% (6%).

Similarly, we formulated two confirmatory hypotheses based on the dwPLI-measures. We expected:

- an increase in (peak) alpha-connectivity when comparing baseline and P-UNR conditions

- a shift of posterior alpha-connectivity towards anterior electrodes when comparing baseline and P-UNR conditions

These visualisations are shown in figures 3.3A and 3.3B, respectively. The functional connectivity topoplots for delta-, theta-, and beta-band are shown in figure 3.4A, 3.4B, and 3.4C.
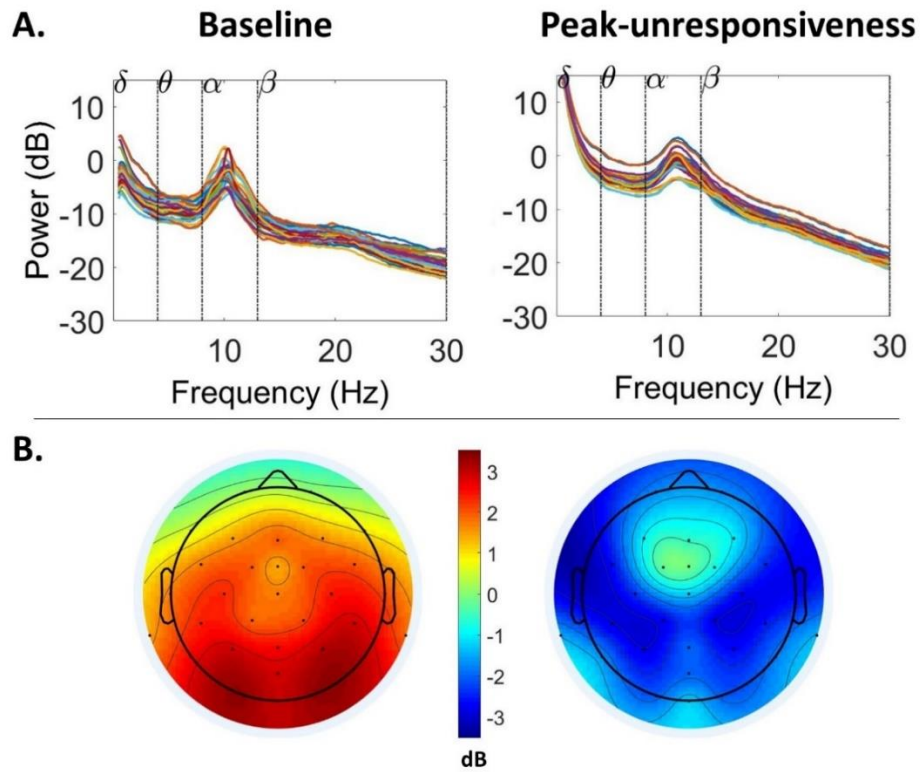
Figure 3.2A. Spectral power for the baseline and peak-unresponsiveness periods averaged over time and over participants. The separate lines reflect the electrodes. An increase in delta-power under anaesthesia was confirmed. **B.** Power topoplot for alpha-band (dB) indicating a shift from posterior electrodes towards anterior electrodes.
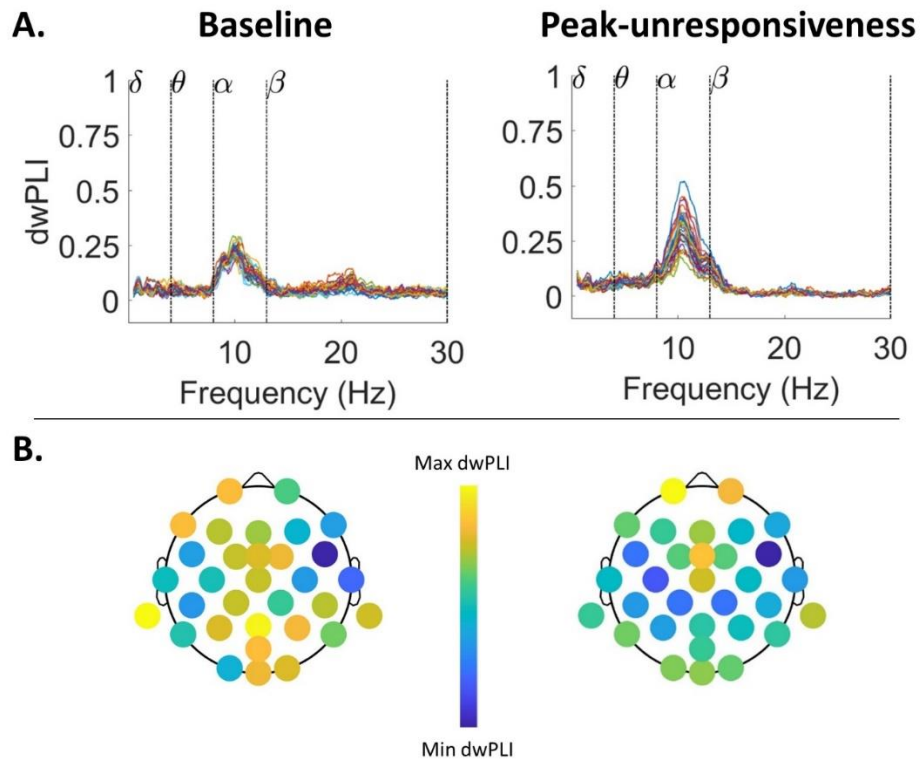
Figure 3.3A. Functional connectivity spectra for baseline and peak-unresponsiveness conditions in the alpha-band. Different lines represent different electrodes averaged over the participants. An increase in maximum alpha-connectivity (with increased across the electrodes) is clearly visible. **B.** Simple connectivity topoplots for alpha-connectivity. Each circle represents an electrode with the colour representing the average connectivity between all electrode pairs. The dwPLI varies between 0 and 0.522 (maximum range 0 – 1). The connectivity in the posterior electrodes diminishes clearly.
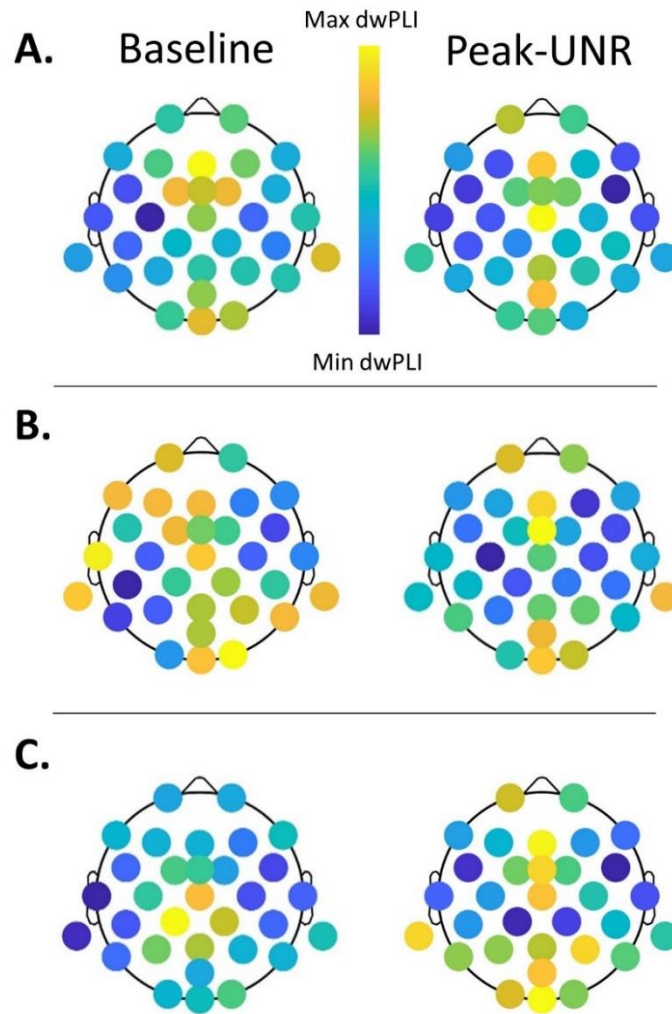
Figure 3.4. Functional connectivity topoplots for delta- (**A**), theta- (**B**), and beta-
connectivity (**C**). Each circle represents an electrode with the colour representing the
average connectivity between all electrode pairs. The connectivity varied between
0.07 – 0.25 (maximum range 0 – 1).

3.3.2 Multivariate regression

Next, we moved to multivariate regression models of the response variables

obtained from the induction condition calculating first predictions based on

absolute- and relative power. None of the predictors reached statistical significance with regards to relative power. With absolute power, there were two exceptions; absolute beta-power predicted both the LOBR time-point and the ESC at the LOBR time-point ($\beta 1 = 33131.2$, $p = 0.029$; $\beta 2 = 52.7$, $p = 0.028$, respectively).

Next, predictions for induction period based on functional connectivity (dwPLI) calculated from the baseline were generated for all electrodes, for frontal electrodes (Fp1, Fp2, F3, F4, F7, F8, and Fz), for posterior electrodes (P3, P4, P7, P8, O1, O2, Pz, Oz, and POz), and for connectivity between the frontoposterior electrodes. Table 3.2 shows the slope terms (β's) with their respective significance levels for all predictors that were found to be statistically significant. The best predictors were based on baseline connectivity averaged from the posterior electrodes: delta-, alpha-, and beta-connectivity predicted the SWA-measures, while posterior theta-connectivity predicted the behavioural measures (time) and the ESC at LOBR and TSWAS reliably. Frontal and frontoposterior connectivity produced less consistent predictions to the same direction. Further, frontoparietal theta-connectivity predicted the SWA-based measures. Predictions based on baseline connectivity averaged from all electrodes predicted only the maximum SWA-power at the beta-band.

Table 3.2. Statistically significant slope terms for baseline functional connectivity (dwPLI) based on all electrodes, frontal electrodes, posterior electrodes, and

frontoposterior electrodes for the induction condition. Here, * indicates statistical

significance at the alpha-level of < 0.05, ** at < 0.01, and *** at < 0.001.

All electrodes

Outcome	Delta	Theta	Alpha	Beta
Max SWA	-	-	-	-40.9*
SWA at TSWAS	-	-	-	-
Mean SWA	-	-	-	-
SWA at LOBR	-	-	-	-
LOBR (time)	-	-	-	-
TSWAS (time)	-	-	-	-
ESC at TSWAS	-	-	-	-
ESC at LOBR	-	-	-	-

Frontal electrodes

Outcome	Delta	Theta	Alpha	Beta
Max SWA	-	-	-	-
SWA at TSWAS	-	-31.9*	-	-
Mean SWA	26.3*	-	-	-
SWA at LOBR	-	-33.9*	-	-
LOBR (time)	-	-	-	-
TSWAS (time)	-	3987.5*	-	-
ESC at TSWAS	-	5.2*	-	-
ESC at LOBR	-	-	-	-

Posterior electrodes

Outcome	Delta	Theta	Alpha	Beta
Max SWA	55.0*	-	14.89*	-52.8***
SWA at TSWAS	45.7*	-	19.1***	-34.6**
Mean SWA	48.5*	-	16.7**	-36.1**
SWA at LOBR	50.2*	-	15.6**	-33.3*
LOBR (time)	-4272.4*	3094.9*	-	-
TSWAS (time)	-	3765.6*	-	-
ESC at TSWAS	-	5.4*	-	-
ESC at LOBR	-6.8*	5.0*	-	-

Frontoposterior electrodes

Outcome	Delta	Theta	Alpha	Beta
Max SWA	41.0 *	-	-	-36.2**
SWA at TSWAS	-	-37.6*	12.5*	-
Mean SWA	-	-30.9*	9.9*	-
SWA at LOBR	-	-39.0*	-	-
LOBR (time)	-	-	-	-
TSWAS (time)	-	-	-	-
ESC at TSWAS	-	-	-	-
ESC at LOBR	-	-	-	-

$* \ p < 0.05, \ ** \ p < 0.01, \ *** \ p < 0.001$

Lastly, based on the results from the regression models with the induction condition, we hypothesised that multivariate regression models with baseline connectivity in the canonical EEG bands should produce predictions to the same direction as with the induction period when predicting SWA in the P-UNR condition. These results are shown in table 3.3. As with the induction condition, the posterior baseline connectivity produced the most consistent results with the SWA-measures. Note, that in P-UNR condition, we only tested SWA-based measures.

Table 3.3. Statistically significant slope terms for baseline functional connectivity (dwPLI) based on all electrodes, frontal electrodes, posterior electrodes, and frontoposterior electrodes for the peak-unresponsiveness condition. The normalised SWA-power refers to subject-wise normalisation to the maximum of the broadband $(0.5 - 30$ Hz). Here, * indicates statistical significance at the alpha-level of < 0.05, ** at < 0.01, and *** at < 0.001.

All electrodes

Outcome	Delta	Theta	Alpha	Beta
Max. SWA abs.	-	-25.7*	-	-
Min. SWA abs.	-	-	-	-
Max. SWA norm.	-	-	-	-
Min. SWA norm.	-	-	1.2*	-
Mean SWA norm.	-	-	1.6*	-

Frontal electrodes

Outcome	Delta	Theta	Alpha	Beta
Max. SWA abs.	-	-20.1*	-	-
Min. SWA abs.	-	-38.6*	-	-
Max. SWA norm.	-	-	-	-6.5*
Min. SWA norm.	-	-	1.1*	-
Mean SWA norm.	-	-	-	-

Posterior electrodes

Outcome	Delta	Theta	Alpha	Beta
Max. SWA abs.	27.7*	-23.8**	-	-22.8**
Min. SWA abs.	-	-33.5*	-	-32.2*
Max. SWA norm.	15.2*	-	6.1**	-8.6*
Min. SWA norm.	4.9*	-	2.1**	-3.0**
Mean SWA norm.	7.9*	-	3.2***	-4.6**

Frontoposterior electrodes

Outcome	Delta	Theta	Alpha	Beta
Max. SWA abs.	-	-23.0*	-	-
Min. SWA abs.	-	-	-	-
Max. SWA norm.	-	-	4.9**	-5.8*
Min. SWA norm.	-	-	1.6**	-

Mean SWA norm.	-	-	2.5***	-

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.4 Discussion

Previous studies have suggested an association between baseline-functional connectivity and the state of consciousness on the one hand (Chennu et al., 2016; Deng et al., 2019; Zhang et al., 2020), and between consciousness and slow-wave activity power (SWA) – specifically the SWA saturation period (SWAS) during the peak-unresponsiveness period (P-UNR) – on the other (Mhuircheartaigh et al., 2013; Warnaby et al., 2017). Here, we aimed to connect these two aspects of the research literature. We investigated to what extent baseline EEG power and functional connectivity (FC) is predictive of SWAS in anaesthetic-induced loss of consciousness (LOC). We first measured EEG power and FC (dwPLI; Vinck et al., 2011) within the canonical frequency bands – delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (15-30 Hz) – from all electrodes, frontal and posterior electrode subsets, and between frontal and posterior subsets during pre-anaesthetic baseline and the P-UNR periods. Next, we tested four confirmatory hypotheses based on power and connectivity to check that we observe commonly found effects of propofol-anaesthesia on EEG power and connectivity. As expected, we observed an increase in slow-wave delta-power, a shift of posterior alpha-power towards anterior electrodes, an increase in the peak alpha-connectivity, and a shift of alpha-connectivity towards anterior electrodes in the anaesthetised conditions when compared to baseline. The rationale was to confirm that there were no substantial differences in the results due to the

84

methods used for the data-analysis before continuing with the multivariate regression.

Next, using a linear multivariate regression, we tested whether the maximum SWA-power, mean SWA-power, SWA-power at the beginning of the saturation period (TSWAS), SWA-power at the time of loss of responsiveness (LOBR), the time needed to reach TSWAS and LOBR, and the propofol concentration levels (effect site concentration: ESC) both at LOBR and TSWAS could be predicted based on baseline relative or absolute power. Here, the dependent measures were measured from the induction period. With relative power, none of the predictions reached statistical significance. With absolute beta-power, the LOBR time-point and the propofol concentration at LOBR reached statistical significance only when measured from all electrodes.

We then generated predictions for the induction period based on FC (dwPLI) using the same dependent variables and the electrode subsets. The best predictors were based on the posterior electrode subset: delta-, alpha-, and beta-connectivity at baseline predicted the SWA-measures at LOBR. Posterior theta-connectivity predicted the behavioural measures (time) and the ESC at LOBR and TSWAS. Both delta- and alpha-connectivity were positively correlated with all SWA-measures, while beta-connectivity had a negative correlation with the SWA-power. In other words, the lower delta- and alpha-connectivity at baseline, the lower the SWA-power, while the opposite was true for beta-connectivity. The set of results with alpha-connectivity complement the results of Chennu et al. (2016)

and Zhang et al. (2020) in suggesting that the alpha-band may play an important role in the observed individual differences in susceptibility to propofol.

Theta-connectivity was positively associated with the two time-points and with the propofol concentrations measures. The frontal- and the frontoposterior subsets produced predictions to the same direction as the posterior subset, but fewer statistically significant predictions and smaller β-values associated with them. The only notable exception was frontal- and frontoposterior theta-connectivity significantly predicting mean SWA-power and SWA-power at the two time-points (LOBR and TSWAS). Connectivity calculated from all electrodes did not produce any other significant predictions except maximum SWA-power at the beta-band.

Lastly, we calculated the SWA-power in the P-UNR condition to build a new multivariate regression model based on baseline connectivity. The rationale was to confirm the direction of the observed associations in the induction period – as calculated from the SWA data produced by the original research group – with the measures from the saturation period. Here, as the saturation had already been reached, we predicted also the minimum SWA-power in addition to the maximum and the mean with both, absolute SWA and SWA normalised to each individual's maximum of the broadband power. The best predictions were again based on posterior electrodes, with delta- and alpha- positively associated and beta-connectivity negatively associated with the normalised SWA-power. Furthermore, theta-connectivity was negatively linked with maximum or minimum absolute SWA-power, or both, in all tested electrode subsets.

86

Increasing SWA-power has been previously linked with prefrontal cortical grey matter and proposed to reflect an increasing number of cortical neurons oscillating synchronously with deepening sedation (Mhuircheartaigh et al., 2013; Warnaby et al., 2017). Moreover, previous studies have found a negative correlation between slow-wave activity and age (Purdon et al., 2015; Ringli & Huber, 2011), the latter of which is known to negatively correlate with grey matter volume (Liu et al., 2003). Accordingly, SWAS has been suggested to have a sound neurobiological basis and to potentially represent a phenotype of an underlying trait (Warnaby et al., 2017). Following this line of thought, one may speculate that the observed positive association between baseline alpha-connectivity and SWA-power at SWAS reflects lower level of required neuronal synchrony oscillating at slow frequencies necessary for reaching the saturation state. In other words, the observed SWA-power levels in participants with less robust baseline alpha-connectivity, i.e. in participants requiring less propofol to reach fully unresponsive states, may reflect an increased sensitivity to reach the saturation state in terms of neuronal synchronisation.

On the other hand, in a previous study by Chennu et al. (2016), the observed differences in the baseline alpha-network robustness were abolished at recovery, suggesting that the differences were depending on the latent alpha-state rather than any individual trait. However, our results support the notion that SWAS, as an index of unconsciousness, is distinct from direct propofol concentration levels, and therefore may indeed reflect a phenotype of a trait. To further investigate whether the SWAS power and its connection with baseline connectivity is state- or trait-based, future studies could analyse the potential link between connectivity

and measures obtained during recovery. For example, does the baseline alpha-connectivity also predict SWA-power at the end of the saturation period, at recovery of behavioural responsiveness (ROBR), or the concentration of propofol and time needed to reach ROBR?

It is worth noting that concerning alpha-connectivity, our results also differ from those of Zhang et al. (2020) who observed the strongest correlation between individual susceptibility to propofol and baseline alpha-connectivity from frontoparietal electrodes; our results support better predictive power for connectivity between the set of posterior electrodes only. However, we did also observe significant predictions based on frontoposterior alpha-connectivity, albeit with smaller slope terms and with fewer significant associations. It is also worth mentioning that Zhang et al. (2020) investigated connectivity between the electrodes in frontal and other cortical areas, and not – for example – between parietal electrodes only. Therefore, the specific contribution of the parietal electrode subset in the results of Zhang et al. (2020) remains unknown.

Interestingly, posterior baseline beta-connectivity was found to be a more robust measure than alpha-connectivity in predicting SWA-power; all of our SWA-measures were negatively associated with the baseline beta-connectivity with larger absolute slope terms than was observed with alpha-connectivity. Previous studies have reported propofol-induced changes in PLI-based graph measures in the delta, theta, alpha, and beta bands (Lee, Mashour, Noh, Kim, & Lee, 2013), functional disconnections in sensorimotor cortices specifically at the beta-band (Malekmohammadi et al., 2018), and reported significant increases in

bidirectional Granger causality strength mainly in the beta and gamma bands (Barrett et al., 2012). However, to the best of our knowledge, the present study is the first one to draw a connection between pre-anaesthesia beta-connectivity and anaesthesia-induced loss of consciousness. Moreover, the fact that previous studies have observed predictive power based only on baseline connectivity in the alpha-band further highlights the differentiation between SWAS on the one hand, and propofol concentration or bispectral index on the other, as individualised indices of the depth of anaesthesia.

Our results also suggested that posterior baseline theta- and delta-connectivity, and to a lesser degree frontal theta, can predict behavioural measures associated with anaesthetic-induced LOC. Baseline theta-connectivity significantly predicted LOBR and TSWAS (in seconds), whereas delta-connectivity predicted only LOBR. Furthermore, theta-connectivity predicted propofol concentration level at TSWAS and both bands at LOBR. Specifically, the higher the baseline delta-connectivity the participants had, the earlier LOBR occurred with lower propofol concentration level. In contrast, higher theta-connectivity at pre-anaesthesia baseline indicated later LOBR and TSWAS, with higher concentrations at both time-points.

The fact that higher baseline delta was associated with both, earlier LOBR and higher SWA-power, is perhaps not surprising given that LOBR has been associated with increases in slow-wave oscillations in anaesthesia and in NREM-sleep (Murphy et al., 2011) and with an increase in parietal and frontoparietal delta-connectivity (Lee et al., 2017). Given a higher baseline level, reaching

89

states characterised by increased slow-wave activity and connectivity more easily makes intuitive sense. Theta-waves, on the other hand, have been primarily associated with memory consolidation in REM-sleep (Adamantidis, Gutierrez Herrera, & Gent, 2019; Karakaş, 2020). Interestingly, theta-waves observed during REM-sleep have been located to originate from hippocampus, whereas cortical theta-waves have been primarily observed to occur during transitions from sleep to wake and in quiet wakefulness (Cantero et al., 2003). Hence, under the assumption that the observed baseline theta-connectivity at pre-anaesthesia resting state was generated by underlying cortical theta oscillations, it may be that behavioural measures of recovery of consciousness could be reliably predicted based on connectivity in the theta-band. Future studies should indeed consider measures from the recovery state and their potential association with baseline connectivity more rigorously.

In summary, we found pre-anaesthetic baseline alpha- and beta-connectivity to reliably predict slow-wave activity power during propofol induction and in slow-wave saturation period: the lower the baseline-alpha, the lower the observed SWA-power and the opposite for connectivity in the beta-band – the higher the baseline-beta, the lower the observed SWA-power. Furthermore, we negatively associated baseline delta-connectivity with the time needed to reach and the corresponding propofol level at LOBR. This was in contrast with positive association between baseline theta-connectivity and the time needed to reach and their corresponding concentration levels at LOBR and at the start of the saturation period. These results, if replicated and verified, support the notion that slow-wave activity saturation is a distinct and individualised index for loss of

consciousness that can be potentially used in tracking and assessing the depth of

anaesthetic-induced unconsciousness.

4. Functional connectivity in disorders of consciousness: Hemispheric connectivity is correlated with behaviour and brain metabolism

This chapter explores the level of granularity with which functional connectivity can potentially distinguish and predict states of disorders of consciousness (DoC). In particular, we are interested in the possible hemispheric differences in EEG brain connectivity. We hypothesise a specific role for left-hemisphere in distinguishing the DoC states, especially unresponsive wakefulness syndrome from minimally conscious states. Linking bedside EEG data with behaviour-based diagnostic labels, we provide evidence for left-hemisphere-specific differences in connectivity. In doing so, we lay the groundwork for further investigation of brain connectivity in DoC at the level of neural sources (chapter 6).

4.1 Introduction

Recent years have seen a rapid increase in consciousness research aiming to quantify the modulations in brain-activity – and connectivity – in altered states of consciousness, including sleep, anaesthetic-induced sedation, and prolonged DoC (for reviews see e.g. Bonhomme et al., 2019; Sanz, Thibaut, Edlow,

Laureys, & Gosseries, 2021; Tagliazucchi & van Someren, 2017). This latter case includes disorders such as the unresponsive wakefulness syndrome (UWS; Laureys et al., 2010), minimally conscious state (MCS; Giacino et al., 2002), and locked-in syndrome (LIS; Bauer, Gerstenbrand, & Rumpl, 1979). UWS is defined by preserved arousal in the absence of behavioural signs of awareness. This is in contrast to MCS, in which patients show fluctuating and incomplete awareness with preserved arousal. The MCS state has been further divided into MCS- and MCS+, with the latter condition characterised by command following, intelligible verbalisation or gestural (or verbal yes/no responses) to spoken or written questions (Bruno et al., 2011). LIS patients on the other hand, are characterised by undisturbed cognitive functions with total immobility except for vertical eye movements and blinking.

Currently the most sensitive behavioural scale to disentangle the DoC groups from each other – especially the MCS from UWS – is the behaviour-based Coma Recovery Scale-Revised (Giacino, Kalmar, & Whyte, 2004; Seel et al., 2010). Previous studies have suggested that the brain states in DoC can also be distinguished with neuroimaging, such as fMRI (Crone et al., 2014; Guldenmund, Vanhaudenhuyse, Boly, Laureys, & Soddu, 2012; Vanhaudenhuyse et al., 2010; 2011) and EEG (Cruse et al., 2011; King et al., 2013; Schnakers et al., 2008). In the case of EEG, the results from studies tracking DoC states have suggested a pivotal role for alpha- and theta-band activity and connectivity, with especially the latter showing promise in prognostic power and in automatic classification of the states (Chennu et al., 2014, 2017; Engemann et al., 2018; Sitt et al., 2014).

Currently the most sensitive neuroimaging method to detect residual consciousness related brain activity is arguably positron emission topography (Laureys, Owen, & Schiff, 2004; Stender et al., 2014; 2015). Bruno et al. (2012) found a key difference distinguishing MCS+ patients from MCS- patients; the MCS+ patients showed evidence of a relatively more preserved left-hemispheric brain metabolism. More specifically, these differences were observed in the left hemisphere "language hotspots": in Broca's and Wernicke's areas. These PET-results are consistent with the observed behavioural differences between the two patient groups; MCS+ patients typically show higher comprehension of speech than do MCS- patients, which suggests differences specifically in the comprehension-related Wernicke's area in the posterior superior temporal gyrus. Supportive results for a potential hemispheric difference in the underlying metabolism in DoC patients were suggested more recently by Thibaut et al. (2021). Their results demonstrated a gradual increase in a metabolic index measured from the best-preserved hemisphere from UWS to seemingly unaware "covertly aware" UWS PET+[2] to MCS patients.

Interestingly, results from a study by Lehembre et al. (2012) suggested also possible differences in hemispheric EEG-connectivity; EEG data were collected from 10 electrodes and power and connectivity comparisons between UWS and

---

[2] A number of studies have reported presence of "covert" voluntary brain activity in some seemingly unresponsive patients, with active and resting state paradigms (Bodart et al., 2017; Claassen et al., 2019; Chennu et al., 2017; Cruse et al., 2011; Lechinger et al., 2013; Monti et al., 2010; Owen et al., 2006; Owen & Coleman, 2008; Schnakers et al., 2015).

MCS patients were performed. Differences were found specifically in the left-hemisphere alpha- and theta-band connectivity – as measured between three electrodes – when comparing UWS to MCS patients.

Here, we aim to replicate and extend these results by distinguishing DoC states with finer granularity – specifically UWS from MCS- from MCS+ – using hemispheric EEG-connectivity. Specifically, based on previous literature (Bruno et al. 2012; Lehembre et al. 2012; Thibaut et al., 2021), we hypothesise that functional connectivity measured from the left-hemisphere electrodes distinguishes MCS- and MCS+ patients from UWS. To that end, we first calculate the debiased weighted phase-lag index (dwPLI; Vinck et al., 2011) from resting state high-density EEG recordings followed by a pre-defined set of summary metrics capturing the network properties – microscale clustering coefficient, macroscale characteristic path length (Watts & Strogatz, 1998), mesoscale modularity (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008), participation coefficient (Guimera & Amaral, 2005), and modular span (see Chennu et al., 2017 for details). We then subject dwPLI and participation coefficient – motivated by previous studies (Chennu et al., 2014, 2017) – to multiple linear regression to investigate whether the DoC group can be predicted from the hemispheric connectivity values. Moreover, we explore the potential differences in hemispheric connectivity between MCS- and MCS+ patients; we test the hypothesis that left-hemispheric connectivity is significantly different between the two MCS patient groups.

4.2 Methods

4.2.1 Participants and dataset

The original dataset was obtained from collaborators at the University Hospital of Liege, Belgium and consisted of high-density EEG recordings from 104 patients diagnosed as either UWS, MCS-, MCS+, emerging from MCS (eMCS), or LIS, alongside with 26 healthy controls. Patients with locked-in syndrome were included as a clinically relevant group for comparison. During the analyses we obtained and preprocessed additional data from 50 more patients using the same brain connectivity analysis pipeline as in the original dataset (Chennu, 2018. *MOHAWK* v.1.2. from: https://github.com/srivaschennu/MOHAWK). After preprocessing, data from nine participants (out of the 50) had to be rejected due to excessive noise (e.g. muscle spasm) or due to failed recordings. Thus, the final dataset consisting of 145 patients and 26 healthy controls was used in a follow-up analysis directly testing for hemispheric differences between MCS- and MCS+ patients.

The data collection was approved by the Ethics Committee of the University Hospital of Liège. Group sizes and mean ages are reported in table 4.1. The EEG recordings were collected during a resting state with a 256-channel high-density EEG cap and Net Amps amplifier. The original length of the recordings varied, however only the first 10 minutes of the recordings from each of the patients were used in the pipeline; this aims to maximise the viability of the pipeline for clinical application as an easy-to-use bedside assessment tool.

96

Table 4.1. Sample sizes, mean ages in years and number of females for the DoC groups. The values for the replenished dataset are shown in parentheses. UWS – unresponsive wakefulness syndrome, MCS- – minimally conscious negative, MCS+ – minimally conscious positive, eMCS – emerging from minimally conscious, LIS – locked-in syndrome.

Group	N	Mean age (years)	Females
UWS	23 (35)	40 (39)	12 (19)
MCS-	17 (22)	36 (-)	9 (12)
MCS+	49 (63)	38 (-)	30 (38)
eMCS	11 (18)	40 (-)	8 (10)
LIS	4 (7)	44 (-)	2 (4)
Control	26 (-)	44 (-)	13 (-)

4.2.2 Analysis

We present a brief description of the data analysis pipeline, please see Chennu et al. (2017) and Chennu (2018; *MOHAWK* v.1.2. Available from: https://github.com/srivaschennu/MOHAWK) for further details. The pipeline uses the functionality provided on EEGLAB (Delorme & Makeig, 2004), FieldTrip (Oostenveld et al., 2011), and the Brain Connectivity Toolbox (Rubinov & Sporns, 2010), and was here implemented in MATLAB2017a (The Mathworks Inc., 2017).

First, the data were imported and EEG-channels on the face, neck and near the eyes were removed to minimise the influence of muscular and ocular artefacts. We retained 173 channels out of 256 for further processing. The data were then downsampled to 250 Hz and filtered within a range of 0.5 – 45 Hz,

encompassing the canonical delta (0 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 13 Hz), beta (13 – 30 Hz), and gamma (30 – 45 Hz) bands. The recordings were then epoched into 10-second long segments and the time points within each epoch were baseline-corrected relative to the mean voltage over the entire epoch. Next, the artefacts were rejected using a quasi-automated procedure that flags abnormally noisy channels and epochs by calculating their variance, which then are subjected for user's visual inspection; on average, approximately 11% of channels (mean = 20, S.D. = 17) were rejected. Next, the infomax independent component analysis algorithm (Bell & Sejnowski, 1995) was used to identify and remove components of activity from potentially non-neural origins, followed by interpolation of the removed (noisy) channels using a spherical spline interpolation. The data were then re-referenced to the common average and the first 60 epochs (first 10 minutes) were retained for further analysis.

Next, spectral and cross-spectral estimates at bins of 0.1 Hz between 0.5 and 45 Hz were calculated using the FieldTrip toolbox (Oostenveld et al., 2011). Power was estimated using a multitaper method with 5 Slepian tapers. At each channel, the power in each band was transformed to normalised percentage contribution by first dividing the sum of the magnitude of power with the total power over all bands, and then multiplying this ratio by 100. Alongside, the cross-spectrum between every pair of electrodes was used to calculate the dwPLI values, which were transformed into symmetric connectivity matrices. The connectivity matrices were then proportionally thresholded, retaining between 90% - 10% of the largest dwPLI values. After applying the thresholds, the matrices were binarised by setting the non-zero values to 1. Due to considerable

electromyographic artefacts observed in high frequencies, we
restricted the analysis to the delta-, theta-, and alpha-bands.

Finally, these binarised matrices were modelled as a network with channels as nodes and the zero-values as edges. Using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010), we calculated a pre-defined set of summary metrics capturing the network properties, i.e. microscale clustering coefficient, macroscale characteristic path length, mesoscale modularity, participation coefficient, and modular span.

4.2.3 Hemispheric hdEEG analysis

Using the above pipeline, data from 104 participants alongside with 26 controls were analysed to calculate connectivity values at the level of electrodes for each hemisphere separately. That is, the electrodes were mapped into a 2D model of the head and all the electrodes on the right and left hemispheres were grouped into hemispheric subsets, disregarding the electrodes on the midline (figure 4.1A). The dwPLI connectivity values were then calculated within (and between) these areas. Furthermore, for visualisation purposes, a number of smaller subsets were defined by including electrodes only on the frontal, central, temporal, parietal, and occipital areas, and the mean dwPLI between the subsets was calculated. These electrode subsets are shown in figure 4.1B.

The main research question was to compare the predictive power of connectivity within the two hemispheres, as measured by dwPLI. Moreover, based on previous research (Chennu et al., 2014), we investigated the predictive

power of participation coefficient – a network metrics measuring the distribution of a node's edges among the communities of a graph. A node with its edges entirely restricted to its community has a participation coefficient of zero, while a node with its edges evenly distributed among all communities has a maximal participation coefficient approaching to one ("connector hubs"). To this end, the mean values of these metrics in each hemisphere were calculated and subjected to multiple linear regression. Hence, our independent variables were the average hemispheric connectivity values followed by the network metrics in each of the canonical delta-, theta-, and alpha-bands. The dependent variable was the DoC diagnosis as per the CRS-R.
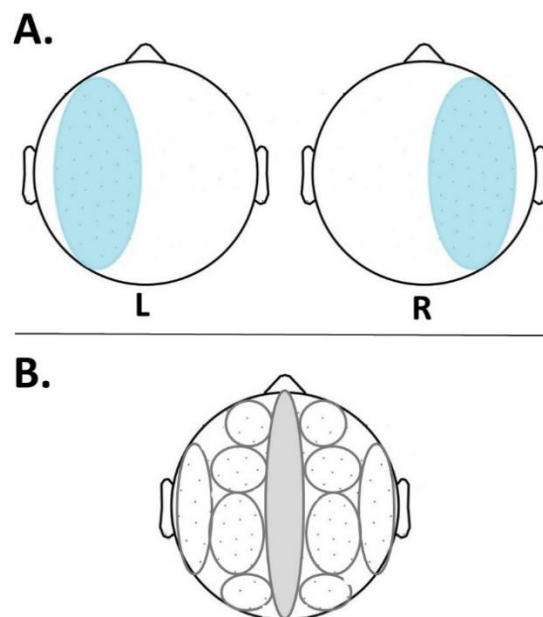
Figure 4.1A. Electrodes on each hemisphere were grouped into subsets, disregarding electrodes on the midline. **B.** For connectivity visualisations, the

electrodes on each hemisphere were divided into 5 smaller subsets, disregarding electrodes on the midline (grey area).

4.3 Results

Our goal was to investigate the differences in hemispheric functional connectivity and its predictive power on DoC states. We divided the electrodes into left and right hemisphere subsets based on the physical locations of the electrodes and calculated the dwPLI connectivity between each of the electrodes within each hemisphere. Furthermore, for visualising the trend in connectivity strength from UWS to MCS- to MCS+ to eMCS to LIS to healthy controls, each hemisphere was further divided into five subsets of electrodes, and mean connectivity between the subsets was calculated. The between-subset connectivity in the canonical delta-, theta-, and alpha-bands in each hemisphere for each DoC states is shown in figure 4.2. The mean connectivity values varied between 0.078 – 0.370 and 0.074 – 0.369 for left and right hemispheres, except for right occipital to posterior alpha-connectivity (0.546).

Visually, UWS and MCS patients show increased connectivity in the delta-band (figure 4.2A) in both hemispheres in comparison to healthy controls. In the theta-band, UWS patients and healthy controls show similar levels of connectivity (figure 4.2B). MCS patients, on the other hand, differ from UWS and healthy controls showing higher level of connectivity in the theta-band. No visually apparent interactions between the hemispheric connectivity and DoC state are visible in the theta-band for the UWS and MCS patients. In alpha-

101

(figure 4.2C) and delta-connectivity (figure 4.2A), there is a larger difference in the left hemisphere between MCS- and MCS+ in comparison to the difference in the right hemisphere. Moreover, when comparing UWS to MCS-, a larger difference is observed in the right hemisphere alpha-connectivity in comparison to the left. In addition, there is a larger difference in the left than right delta- and alpha-connectivity between LIS and healthy controls.
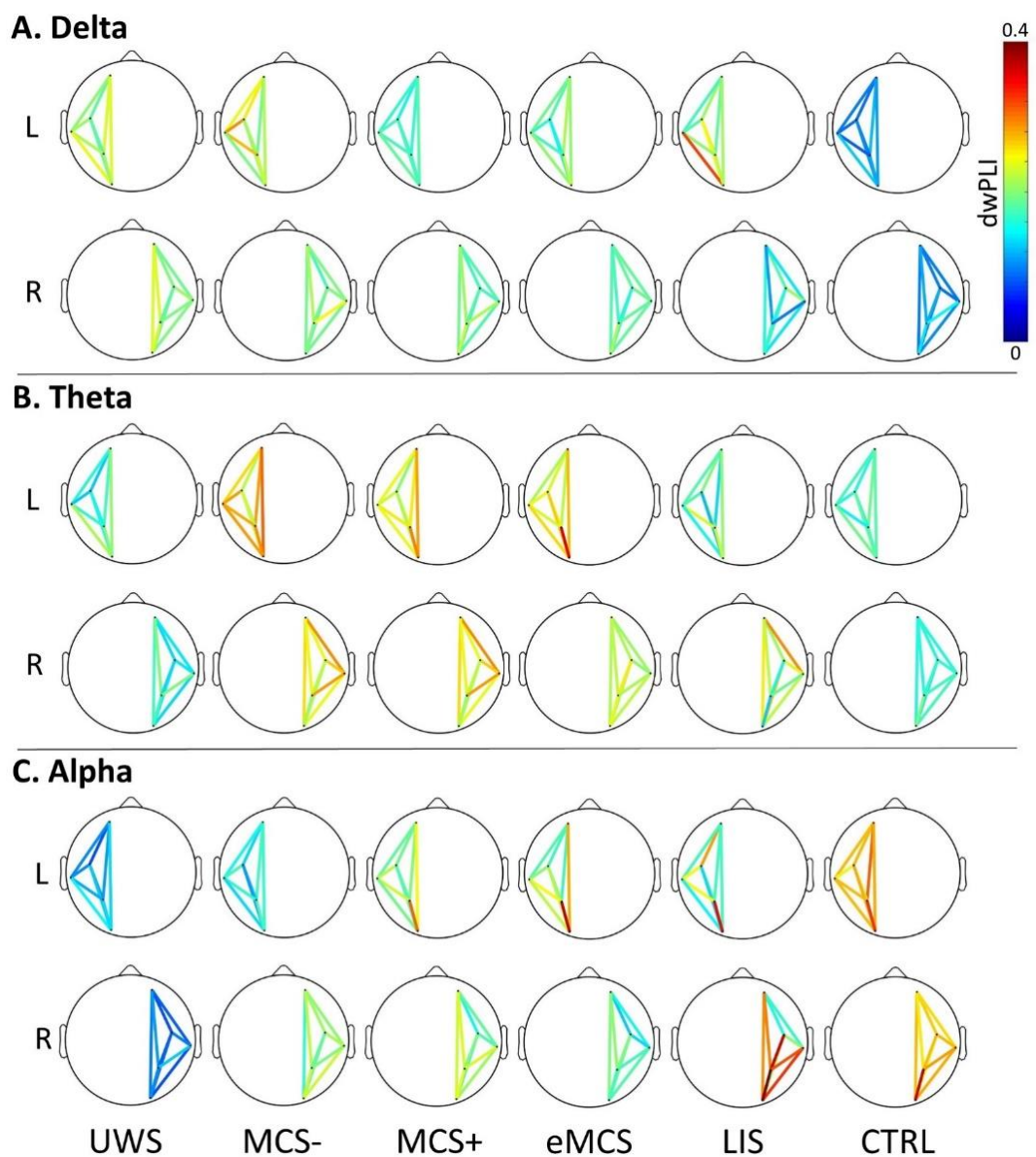
Figure 4.2. Progression of the change in mean connectivity in the delta- (A),

theta- (B), and alpha-bands (C) from unresponsive wakefulness syndrome (UWS)

to minimally conscious negative (MSC-) to minimally conscious positive

(MSC+) to emergence from MCS (eMCS) to fully conscious locked-in syndrome

(LIS) and healthy controls (CTRL) for each hemisphere. Connectivity is

calculated for five subsets in each hemisphere and means between the subsets are

visualised. The UWS and MSC patients show increased connectivity in the delta-

band in comparison to fully conscious healthy controls (A). In contrast, MCS

patients show increased levels of theta-connectivity in comparison to UWS and

healthy controls (B). In alpha- (2C) and delta-connectivity (2A), there is a larger

difference in the left hemisphere between MCS- and MCS+ in comparison to the

difference in the right hemisphere. Similarly, between LIS and healthy controls,

larger differences are observed in the left hemisphere in comparison to right in

the delta- and alpha-band (A and C, respectively). Moreover, when comparing

UWS to MCS-, a larger difference is observed in the right hemisphere alpha-

connectivity in comparison to the left.

Next, in an exploratory analysis, we built regression models of functional

connectivity (dwPLI), one for each of the delta-, theta-, and alpha-bands. In each

model, left and right hemispheric connectivity were predictors and the DoC state

was the predicted variable. In the delta- and theta-bands, neither hemispheric

predictor reached statistical significance (left hemisphere ps = 0.306 (delta) &

0.796 (theta), and right hemisphere ps = 0.423 (delta) & 0.993 (theta),

respectively). In the alpha-band, the results reached statistical significance and

indicated roughly 15% of the variance explained by the predictors (left vs. right

hemisphere dwPLI; R^2 = 0.146, $F(2,127)$ = 10.87, p = < 0.001). The results

further indicated that only left hemisphere connectivity significantly predicted

the DoC group (β = 0.26, p = .027), while right hemisphere connectivity did not

(β = 0.16, p = 0.171). Figure 4.3 shows mean hemispheric connectivity values in

the alpha-band for each data point in the patient groups. Note, that the p-values

reported here are not corrected for the multiple comparisons of the three

frequency bands.

We furthermore built a regression model combining all frequency bands to

assess the ability of four independent variables (three canonical EEG bands and

hemispheric electrode-subsets) to predict DoC states. The four predictors

explained approximately 18% of the variance in DoC states (R^2 = 0.177, $F(4,255)$

= 13.76, p = < 0.001). The effects of two predictors, delta- and alpha-band

connectivity, were statistically significant: β = -1.90, p = 0.013 (delta) & β =

4.42, p < 0.001 (alpha). Connectivity in the theta-band and hemisphere did not

reach statistical significance (ps > 0.110).
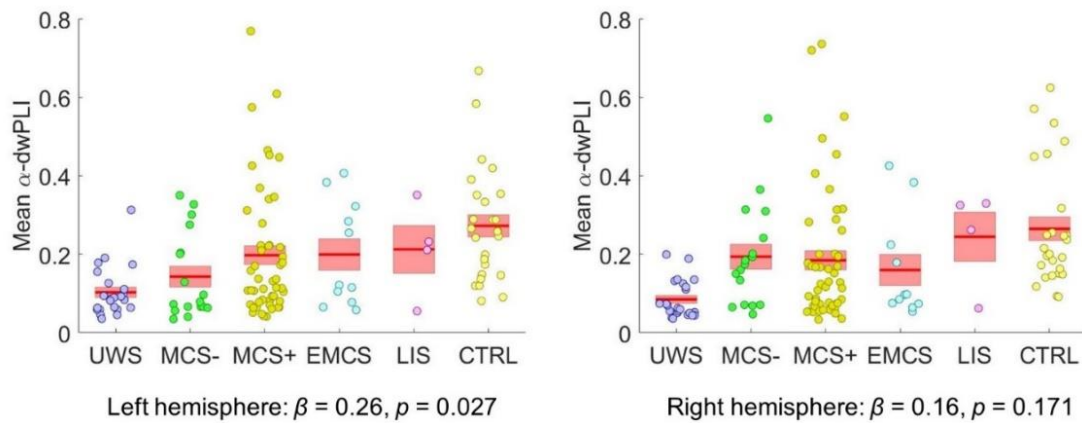
Figure 4.3. Mean connectivity in alpha-band for left and right hemispheres for each DoC state and healthy controls. Hemispheric connectivity significantly predicts the patient group; the figure shows a more robust left-hemispheric linear relationship between connectivity and behavioural diagnosis. Connectivity in the delta- and theta-bands did not reach statistical significance (ps > 0.306). Note that the p-values are uncorrected. UWS – unresponsive wakefulness syndrome, MCS- – minimally conscious negative, MCS+ – minimally conscious positive, eMCS – emerging from MSC, LIS – locked-in syndrome, CTRL – healthy controls.

We next investigated the presence and the predictive power of brain network centrality within each hemisphere (as captured by participation coefficient). In a regression analysis conducted by band, we tested if the standard deviation (SD) of alpha-band participation coefficients across electrodes within each hemisphere significantly predicted DoC group. The results indicated that the two predictors (left vs. right hemisphere) explained roughly 25% of the variance ($R^2 = 0.246$, $F(2,127) = 20.75$, $p = < 0.001$). Furthermore, we found that right hemisphere participation coefficient significantly predicted the DoC group ($\beta =$

0.40, $p < 0.001$), whereas left hemisphere participation coefficient did not reach statistical significance ($\beta = 0.15$, $p = 0.122$). With the canonical delta- and theta-band, neither predictor reached statistical significance (left hemisphere ps = 0.860 (delta) & 0.332 (theta), and right hemisphere ps = 0.352 & 0.054, respectively). Figure 4.4 illustrates the alpha-band participation coefficients in each hemisphere for each subject group.

As with connectivity, we also built a regression model to assess the ability of four independent variables (three canonical EEG bands and hemispheric electrode-subsets) to predict DoC states. The overall model was non-significant: $R^2 = 0.03$, $F(4,255) = 1.96$, $p = 0.101$. Of the four predictors, only the alpha-band participation coefficient significantly predicted the DoC state ($\beta = -1.18$, $p = 0.033$; all other ps > 0.76).
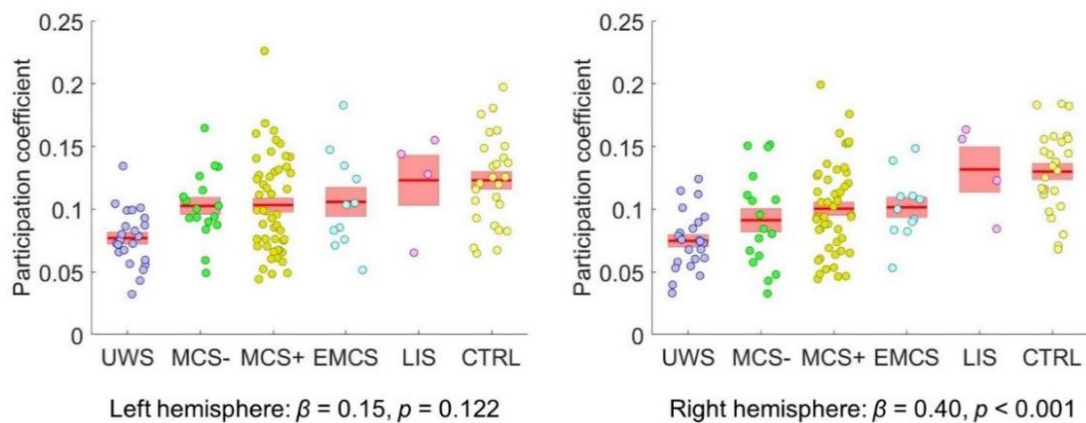
Figure 4.4. SD of participation coefficient in the alpha-band of DoC groups for both hemispheres. Hemispheric participation coefficients significantly predicted the patient group; the figure shows a more robust right-hemisphere linear

106

relationship between SD of participation coefficient and behavioural diagnosis. Participation coefficients in delta- and theta-bands did not reach statistical significance (ps > 0.054). Note that the p-values are uncorrected. UWS – unresponsive wakefulness syndrome, MCS- – minimally conscious negative, MCS+ – minimally conscious positive, eMCS – emerging from MSC, LIS – locked-in syndrome, CTRL – healthy controls.

One of the aims in the present study was to investigate if the previous report of preserved left-hemisphere PET metabolism in MCS+ patients (in comparison to MCS- patients) could also be observed in the hdEEG signal. To this end, we directly compared the left- and right-hemisphere connectivity in delta-, theta-, and alpha-bands in a larger sample of 85 patients (MCS-: N = 22; MCS+: N = 63).

To replicate the PET finding with hdEEG, we tested for an interaction between connectivity in each hemisphere and patient group (MCS- vs. MCS+). The mean alpha-connectivity values for each of the groups and for both hemispheres are depicted in figure 4.5. The results did not reach statistical significance (in an ANOVA; with alpha-connectivity $p = 0.067$; with theta $p = 0.557$; with delta $p = 0.163$).

Finally, due to previously observed correlation of right hemispheric (alpha-band) participation coefficients, we tested for an interaction between participation coefficients in each hemisphere and patient group. The results did

not reach statistical significance (in an ANOVA: with alpha-band participation coefficients $p = 0.084$; with theta $p = 0.072$; with delta $p = 0.819$).
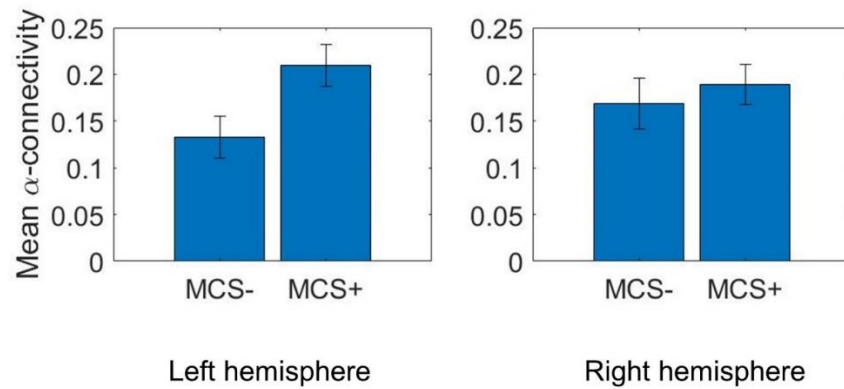
Figure 4.5. Mean alpha-connectivity for both hemispheres in minimally conscious negative (MCS-) and minimally conscious positive (MCS+) patients. The error bars represent the standard error of the mean.

4.4 Discussion

In this study, we investigated the differences in hemispheric functional connectivity and its predictive power on states of DoC. We divided the electrodes into left and right hemisphere subsets based on the physical locations of the electrodes and calculated functional connectivity between each of the electrodes within each hemisphere. Based on connectivity visualisations, UWS and MCS patients showed increased connectivity in the delta-band in both hemispheres when compared to healthy controls. In the theta-band, all MCS sub-groups (MCS-, MCS+, and eMCS) differed from healthy controls and UWS patients

showing relatively increased theta-connectivity in both hemispheres. No apparent interactions between the hemispheric connectivity and DoC state are visible in the theta-band for the UWS and MCS patients. In alpha- and delta-connectivity, there was visually a larger difference in the left hemisphere between MCS- and MCS+ in comparison to the difference in the right hemisphere. Moreover, when comparing UWS to MCS-, a larger difference was observed in the right hemisphere alpha-connectivity in comparison to the left. Additionally, there was also a larger difference in the left than right delta- and alpha-connectivity between locked-in syndrome patients and healthy controls.

The connectivity visualisations were followed by multiple linear regression models of the DoC states first based on hemispheric functional connectivity (dwPLI) in the three canonical bands (delta, theta, and alpha), followed by regression models based on intermodular centrality within the hemispheres (as captured by the standard deviation of participation coefficient). In the models based on dwPLI, only predictions in alpha-band reached statistical significance. Specifically, left-hemispheric alpha-band connectivity predicted the DoC state of the patients, while the right-hemispheric connectivity did not. The opposite result was observed with the network hubs; right-hemispheric participation coefficient in the alpha-band statistically significantly predicted the DoC states, while left-hemispheric participation coefficient did not. With the canonical delta- and theta-band, the models did not reach statistical significance.

Last, we directly compared the hemispheric connectivity in the three bands between MCS- and MCS+ patients. Although we observed a trend towards

reduced alpha-connectivity in the left hemisphere in MCS- patients when compared to MCS+ patients, these results did not reach statistical significance. This may have been because of the small and unequal sample size. Future studies could try to establish statistical significance in left-hemisphere alpha-connectivity between MCS- and MCS+ patients with larger samples.

The results from the connectivity-based regression model suggest that left hemisphere alpha-connectivity, at the level of electrodes, significantly correlates with DoC states, unlike connectivity in the right hemisphere. These results are consistent with the previous findings linking left hemisphere more strongly with behavioural DoC diagnosis (Bruno et al. 2012; Lehembre et al. 2012; see also Thibaut et al., 2021). Specifically, the observed key role for left-hemisphere alpha-band connectivity in distinguishing the DoC states may reflect the previously observed underlying differences in PET-metabolism (Bruno et al. 2012). These results suggest that hemispheric spectral connectivity provides information about brain dynamics that are not captured by simpler estimations of spectral power. It is also worth noting that the two sub-categories of MCS (MCS- and MCS+), are distinguished behaviourally based on the absence or presence, respectively, of evidence of residual language functions (Bruno et al., 2011), typically considered to be predominantly located in the left hemisphere. It is possible that the observed larger differences in the left than right hemisphere alpha-connectivity between the DoC states, particularly between MCS- and MCS+, reflect more preserved language functions.

In contrast to the results of the present study, Lehembre et al. (2012) observed not only hemispheric differences in left and right hemisphere delta- and alpha-power spectra, but their results also suggested that specifically left hemisphere alpha-band connectivity distinguished UWS from MCS patients. Here, based on the connectivity visualisations (figure 4.2), a larger difference in alpha-connectivity is observed on the right hemisphere, than left, between UWS and MCS- patients. However, when comparing UWS to MCS+ patients, this difference seems to disappear; this discrepancy in the results between this study and those of Lehembre et al. (2012) may thus reflect the further distinction of MCS patients into MCS- and MCS+ sub-groups. It should be also noted that here, we used high-density EEG recordings, whereas the connectivity in Lehembre et al. (2012) was calculated based on recordings from 10 electrodes (only three in each hemisphere). Such differences in the methodologies between the two studies might partly explain the differences in the results.

Observing a difference in left hemisphere connectivity between UWS and MCS+, in addition to the trend toward larger left hemisphere difference between MCS- and MCS+, is further consistent with the possibility of residual language functions driving the difference in functional connectivity. However, based on this study alone, this remains a speculative hypothesis at best. Future studies should further investigate the possible affiliation between more preserved left-hemispheric PET metabolism, residual language functions, and hemispheric functional connectivity.

Moreover, we observed significant prediction based on right-hemispheric, but not left, participation coefficient in the alpha-band. This observed dissociation between predictions based on functional connectivity on the one hand, and on connectivity network on the other, highlights not only that the network properties of nodes can vary independently of the average connectivity of that node (Chennu et al., 2014; 2017), but that hemispheric differences may provide complementary estimates of underlying properties of the brain's dynamical networks. Such information, if verified and replicated, could potentially be used to aid and guide diagnosis of DoC patients.

It should be noted, that in regression models encompassing all three frequency bands alongside with hemispheric electrode subsets, only the frequency bands significantly predicted the DoC states. Indeed, the reported p-values for regression models based on single frequency bands are not corrected for multiple corrections. Although previous studies (e.g. Chennu et al., 2017; Lehembre et al., 2012) provide a justification for hypothesising the predictive power for alpha-band connectivity specifically, the fact that the hemispheric effect in the alpha-band disappears when correcting for the three comparisons (frequency bands) alongside with failure to detect any difference when comparing MSC- and MCS+ patients directly, speak to a lower power in our study. Future studies should try to establish statistical significance with larger sample sizes.

In summary, we investigated the hemispheric differences in connectivity and network properties between DoC states at the level of electrodes. We observed

that left-hemispheric alpha-connectivity on the one hand, and right-hemispheric participation coefficient (centrality or "hubness") on the other, significantly predicts the DoC state. These results – in combination with previous findings from PET-imaging (Bruno et al. 2012; Thibaut et al., 2021) and functional EEG connectivity (Lehembre et al. 2012) – suggest that hemispheric connectivity may provide information over and beyond an analysis of global connectivity. However, the present analysis was performed at the level of electrodes, and consequently, we cannot directly infer the underlying brain dynamics at the source level. To address this gap, we will investigate the potential connectivity differences between the hemispheres in DoC patients at the level of neuronal sources using effective connectivity methods (chapter 6).

5. Modelling the effects of propofol-anaesthesia on the effective connectivity

This chapter explores the changes in effective connectivity following propofol-induced loss of consciousness (LOC). In particular, we assess the differences in intra- and inter-network effective connectivity in three key resting state networks using dynamic causal modelling to explain cross-spectral densities from EEG data. In doing so we contribute to the frontal vs. posterior debate of neural correlates of consciousness by providing novel computational evidence supporting a selective breakdown of posterior parietal and medial feedforward frontoparietal connectivity within the default mode network (DMN) and of parietal inter-network connectivity linking the DMN and the central executive network. Going further, we establish the generalised predictive validity of our models using a novel DCM-based cross-validation, by predicting unseen data from the post-anaesthetic recovery state

5.1 Introduction

Several cortical network-level mechanisms have been proposed to explain human consciousness and its loss, of which two, in particular, have received an increasing amount of interest and evidence. On the one hand, empirical studies

have suggested that the loss of consciousness (LOC)[3] is associated with disruptions of within- and between-network connectivity in cortical areas associated with large-scale frontoparietal networks (Bor & Seth, 2012; Laureys & Schiff, 2012). On the other, temporo-parieto-occipital areas – colloquially named as 'the posterior hot zone' – has been shown to be important in mediating changes in consciousness during sleep (Siclari et al., 2017; Lee et al., 2019), and in patients with brain damage (Vanhaudenhuyse et al., 2010; Wu et al., 2015).

In this context, general anaesthetics are a powerful tool to investigate alterations in brain connectivity during changes in the state of consciousness (see Bonhomme et al., 2019 for a recent review). Indeed, several previous studies have utilised anaesthetic drugs in investigating brain dynamics in both functional and effective/directed connectivity studies and suggested multiple explanatory mechanisms of the LOC. Note that here, effective connectivity is defined following Friston (2011) and Razi & Friston (2016) as a causal influence (in a control theory sense) of one neural population over another. In contrast, functional connectivity indicates statistical correlation between neurophysiological signals. Some of these previous studies have suggested a breakdown of thalamo-cortical connections and disrupted frontoparietal networks

---

[3] We acknowledge that anaesthetic-induced loss of consciousness (LOC) may actually be anaesthetic-induced loss of behavioural responsiveness (LOBR), as e.g. volitional mental imagery or dreaming may take place during the anaesthetic state. The participants were, however, asked afterwards if they had any recall of dreams etc., which they did not report. Thus, here, we follow the typical convention in anaesthesia-literature and refer to this state as LOC.

(Boveroux et al., 2010; Schrouff et al., 2011). Others have found disruptions in frontal areas (Guldenmund, et al., 2016), diminished frontoparietal feedback connectivity (Ku, Lee, Noh, Jun, & Mashour, 2011; Lee et al., 2009; Lee, Ku et al., 2015) , and increased frontoparietal connectivity (Barrett et al., 2012). To bring computational evidence to bear upon this discussion, we adopt one of the most commonly used methods for understanding effective connectivity, dynamic causal modeling (DCM; Friston, Harrison & Penny, 2003), to assess cortical network-level mechanisms involved in the LOC, and evaluate the evidence for the posterior hot zone. Our analysis of effective connectivity complements existing research into functional connectivity during general anaesthesia, including that presented in chapter 3. In comparison to functional connectivity, effective connectivity allows stronger inferential claims to be made about how brain regions influence one another.

There are relatively few studies assessing resting state effective connectivity with DCM during anaesthetic-induced unconsciousness, but a recent fMRI study identified impaired subcortico-cortical connectivity between globus pallidus and posterior cingulate (PCC) nodes, but no cortico-cortical modulations (Crone, Lutkenhoff, Bio, Laureys, & Monti, 2017). Boly et al. (2012) found a decrease in feedback connectivity from frontal (dorsal anterior cingulate; dACC) to parietal (PCC) nodes. Both of these studies, however, evaluated relatively simple models in terms of cortical sources (excluding subcortical nodes), consisting of only two such nodes – an anterior and a posterior node. Consequently, they do not allow us to compare the role of the posterior hot zone to other potential cortical mechanisms underpinning consciousness.

116

Here, we address this gap by modelling changes in key resting state networks (RSN) - the default mode network, the salience network (SAL), and the central executive network (CEN), due to unconsciousness induced by propofol, a common clinical anaesthetic. We employ a novel methodological combination of DCM for resting EEG cross-spectral densities (CSD; Friston et al., 2012; Moran et al., 2009) and parametric empirical Bayes (PEB; Friston et al., 2016), to better estimate model parameters (and their distributions) and prune redundant connections. Within this framework, we invert - for the first time - a single large-scale model of EEG, consisting of 14 RSN nodes, in addition to the individual RSNs themselves (figure 1). This allows us to evaluate the role of different subgroups of intra- and inter-RSN connections in the modulation of consciousness. Further, we apply robust leave-one-subject-out-cross-validation (LOSOCV) on DCM model parameters, to evaluate hypotheses about whether specific sets of connections within and between frontal and parietal nodes are not only able to explain changes between states of consciousness, but also to predict the state of consciousness from unseen EEG data. Using this combination of computational modelling, cross-validation and hypothesis testing, we indicate the importance of the posterior hot zone in explaining the loss of consciousness, while also highlighting the distinct role of frontoparietal connectivity in underpinning conscious responsiveness. Consequently, we demonstrate a dissociation between the mechanisms most prominently associated with explaining the contrast between conscious awareness and unconsciousness, and those maintaining consciousness.

5.2 Methods

5.2.1 Data acquisition and preprocessing

The data used in the present work were acquired from a previous propofol-anaesthesia study, which describes the experimental design and data collection procedure in detail (Murphy et al., 2011). The study was approved by the Ethics Committee of the Faculty of Medicine of the University of Liège, and written consent was obtained from all the participants. None of the participants suffered from mental illness, drug addiction, asthma, motion sickness, nor had a history of mental illness or suffered from any previous problems with anaesthesia. The data consisted of 15 minutes of spontaneous, eyes-closed high-density EEG recordings (256 channels, EGI) from 10 participants (mean age 22 ± 2 years, 4 males) in four different states of consciousness: behavioural responsiveness, sedation (Ramsay scale score 3, slower responses to command), loss of consciousness with clinical unconsciousness (Ramsay scale score 5-6, no response to command), and recovery of consciousness (Ramsay, Savege, Simpson, & Goodwin, 1974). Note that for the recovery state, the data consisted of 9 datasets. Participants were considered to be fully awake if the response to verbal command ('squeeze my hand') was clear and strong (Ramsay 2), and in LOC, if there was no response (Ramsay 5-6). The Ramsay scale verbal commands were repeated twice at each level of consciousness. Propofol was infused through an intravenous catheter placed into a vein of the right hand or forearm, and the propofol plasma and effect-site concentrations were estimated with 3.87 ± 1.39 mcg/mL average arterial blood concentration of propofol for LOC. Here, we only modelled data from the maximally different anaesthetic

states, behavioural responsiveness and LOC, and used recovery as a test of DCM model generalisation. These data can be made available after signing a formal data-sharing agreement with the University of Liège.

Data from channels from the neck, cheeks, and forehead were discarded as they contributed most of the movement-related noise, leaving 173 channels on the scalp for the analysis. These 173 electrodes were co-registered to a template MRI mesh in MNI coordinates, and the volume conduction model of the head was based on the Boundary Element Method (BEM). The raw EEG signals were filtered from 0.5 – 45 Hz with additional line noise removal at 50 Hz using a notch filter. The recordings were then downsampled to 250 Hz, and abnormally noisy channels and epochs were identified by calculating their normalised variance, and then manually rejected or retained by visual inspection. Last, the data were then re-referenced using the average reference.

5.2.2 Dynamic causal modeling

For the DCM modelling of the high-density EEG data, the first 60 artefact-free 10-second epochs in wakeful behavioural responsiveness and LOC were combined into one dataset with two anaesthetic states making up a total of 120 epochs per participant. The preprocessed data was imported in to SPM12 (Wellcome Trust Centre for Human Neuroimaging; www.fil.ion.ucl.ac.uk/spm/software/spm12).

To analyse effective connectivity within the brain's resting state networks, DCM for EEG cross-spectral densities (CSD) was applied (Friston et al., 2012;

Moran et al., 2009). Briefly, with this method, the observed cross-spectral densities in the EEG data are explained by a generative model that combines a biologically plausible neural mass model with an electrophysiological forward model mapping the underlying neural states to the observed data. Each node in the proposed DCM models – that is, each electromagnetic source – consists of three neural subpopulations, each loosely associated with a specific cortical layer; pyramidal cells, inhibitory interneurons and spiny stellate cells (ERP model; Moran, Pinotsis & Friston, 2013). DCM does not simply estimate the activity at a particular source at a particular point in time – instead, the idea is to model the source activity over time, in terms of interacting inhibitory and excitatory populations of neurons.

The subpopulations within each node are connected to each other via *intrinsic* connections, while nodes are connected to each other via *extrinsic* connections. Three types of extrinsic connections are defined, each differing in terms of their origin and target layers/subpopulation: forward connections targeting spiny stellate cells in the granular layer, backward connections targeting pyramidal cells and inhibitory interneurons in both supra- and infragranular layers, and lateral connections targeting all subpopulations. This laminar specificity in the extrinsic cortical connections partly defines the hierarchical organisation in the brain. Generally speaking, the backward connections are thought to have more inhibitory and largely modulatory effect in the nodes they target (top-down connections), while forward connections are viewed as having a strong driving effect (bottom-up; Salin & Bullier, 1995; Sherman & Guillery, 1998).

The dynamics of hidden states in each node are described by second-order differential equations, which depend on both, the parametrised intrinsic and extrinsic connection strengths. This enables the computation of the linear mapping from the endogenous neuronal fluctuations to the EEG sensor spectral densities, and consequently, enables the modelling of differences in the spectra due to changes in the underlying neurophysiologically meaningful parameters describing, for example, the intrinsic and extrinsic connectivity of coupled neuronal populations (i.e. sources) and their physiology. Here, for straightforward interpretability, we have focused on the changes in extrinsic connections as a result of changes in the state of consciousness. It should be noted that we did not fix any of the other parameters typically estimated by DCM using the ERP-model, rather, we estimated all our models using the default DCM setting (for further information about EEG DCM, see for example Friston et al., 2012; Kiebel, Garrido, Moran, & Friston, 2008; Moran, Kiebel et al., 2007; Moran et al., 2009). Nevertheless, from here on, we focus on the extrinsic connectivity parameters and their modulations referring to them as 'parameters'.

5.2.3 Model specification

Fitting a DCM model requires the specification of the anatomical locations of the nodes/sources a priori. Here, we modelled three canonical RSNs associated with consciousness (see for example Boly et al., 2008; Heine et al., 2012), namely the Default Mode Network (DMN), the Salience Network (SAL), and the Central Executive Network (CEN). In addition, we modelled a fourth large-scale

network (LAR) combining all the nodes and connections in the three RSNs above, with additional inter-RSN connections motivated by structural connectivity (details below). The node locations of the three RSNs modelled here were taken from Razi et al. (2017) and are shown in figures 5.1A, 5.1B, 5.1C, and 5.1D with their respective schematic representations. The node locations in figure 5.1 and the effective connectivity modulations in figures 5.4A, 5.5A, 5.6A, and 5.7A were visualized with the BrainNet Viewer (Xia, Wang, & He, 2013, http://www.nitrc.org/projects/bnv/). The MNI coordinates are listed in table 5.1. Coincidentally, these same data have been previously source localised to the same locations as some of the key nodes in the RSNs modelled here (Murphy et al., 2011). We treated each node as a patch on the cortical surface for constructing the forward model ('IMG' option in SPM12; Daunizeau, Kiebel, & Friston, 2009).

Nodes in the three RSNs were connected via forward, backward, and lateral connections as described in David et al. (2006, 2005). Thus, each node (in each RSN-model) were modelled as a point source with the neuronal activity being controlled by operations following the Jansen-Rit model (Jansen & Rit, 1995). Note that all our models were fully connected. In addition to preserving the connections within the nodes of the original three RSNs, in the LAR, we additionally hypothesised potential connections between the RSNs. Previous structural connectivity studies have identified a highly interconnected network of RSN hubs that seem to play a crucial role in integrating information in the brain, often termed the 'rich-club' (van den Heuvel & Sporns, 2011). Specifically, van den Heuvel and colleagues localised a number of these key-hubs to regions

comprising of the precuneus, superior lateral parietal cortices, and superior frontal cortex, thus, to some extent overlapping with some of the key-nodes in our RSN models. Therefore, as a structurally-informed way to investigate the potential anaesthesia-induced modulations of effective connectivity between the three RSNs, we specified – in addition to the already-specified connections in our RSNs – bi-directional connections between PCC/precuneus and left/right superior parietal nodes (connecting DMN and CEN), and between PCC/precuneus and anterior cingulate cortex (connecting DMN and SAL).

These three different types of connections in each model were specified in what is referred in the DCM literature as the 'A-matrix'. In addition, to explicitly parameterise the effect of the session – i.e. the effect of the anaesthetic – on the connections, we allowed every connection to change (specified in the 'B-matrix').

Table 5.1. All the nodes and their corresponding MNI coordinates for the three resting state networks (adapted from Razi et al., 2017). The large model incorporated all these nodes as a single model.

Network	Coordinates (in mm)		
	x	**y**	**z**
Default Mode Network			
Left lateral parietal	-46	-66	30
Right lateral parietal	49	-63	33
Posterior cingulate/precuneus	0	-52	7
Medial prefrontal	-1	54	27
Salience Network			
Left lateral parietal	-62	-45	30
Right lateral parietal	62	-45	30

Dorsal anterior cingulate	0	21	36
Left anterior PFC	-35	45	30
Right anterior PFC	32	45	30
Central Executive Network			
Left superior parietal	-50	-51	45
Right superior parietal	50	-51	45
Dorsal medial PFC	0	24	46
Left anterior PFC	-44	45	0
Right anterior PFC	44	45	0
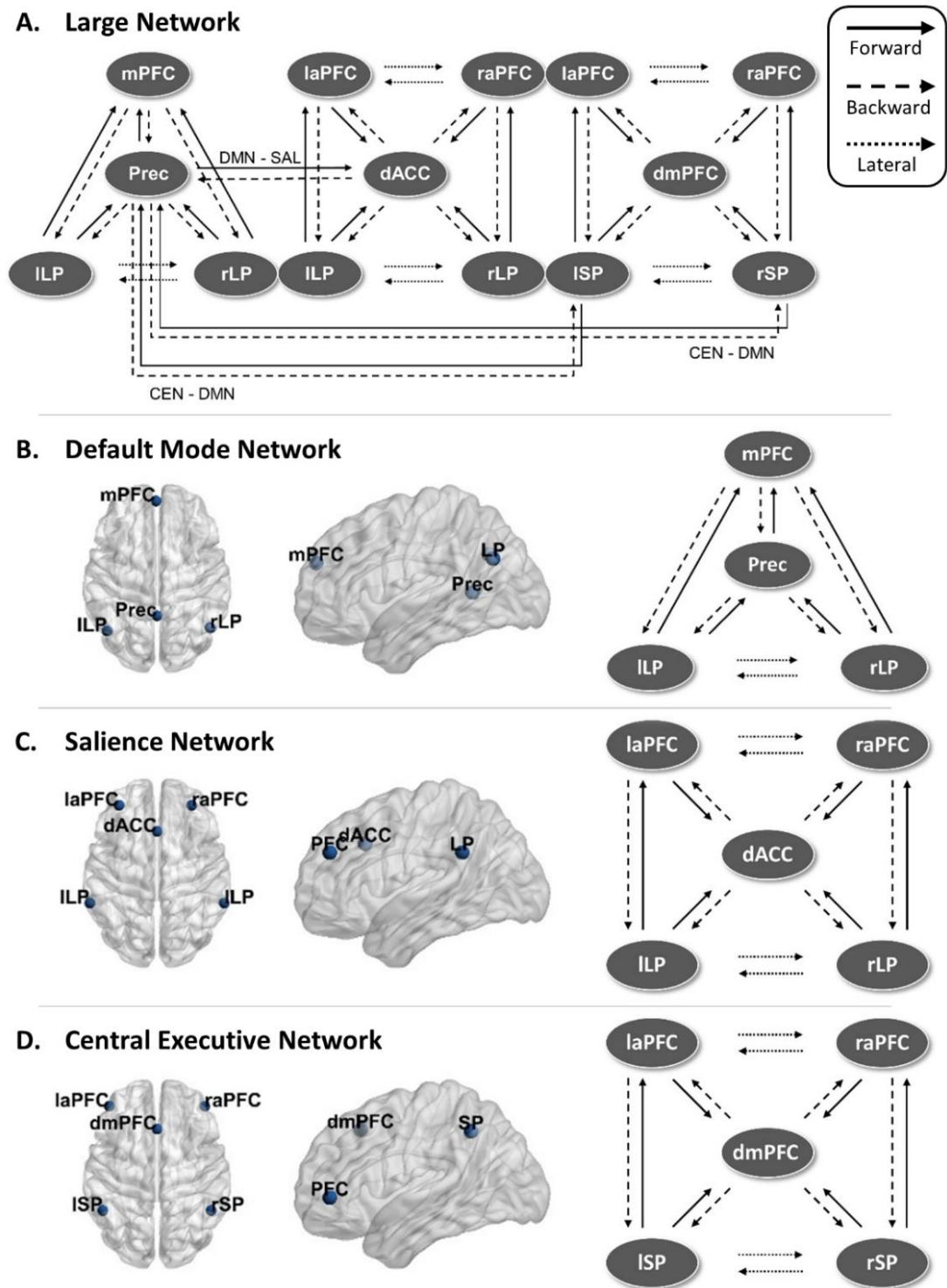
Figure 5.1. Full model schematics and node locations. mPFC – medial prefrontal

cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal

cortex, rLP – right lateral parietal cortex, laPFC – left anterior prefrontal cortex,

raPFC – right anterior prefrontal cortex, dACC – dorsal anterior cingulate cortex, dmPFC – dorsomedial prefrontal cortex, lSP – left superior parietal cortex, and rSP – right superior parietal cortex. **A.** Schematic view of the large DCM model consisting of the 14 nodes and connections combining three RSNs. Inter-RSN connections were specified between PCC/precuneus and bi-lateral superior parietal nodes, and between PCC/precuneus and anterior cingulate cortex. **B-D.** Location of the nodes and the schematic representation of the full model for DMN, SAL, and CEN, respectively.

5.2.4 Model inversion

In DCM, model inversion refers to fitting the models to best explain the empirical data of each participant's dataset, and thereby inferring a full probability density over the possible values of model parameters (with the expected values and covariance). Here, we first modelled the effects of propofol in terms of changes in connectivity that explained the differences in the empirical data observed in LOC as compared to behavioural responsiveness baseline (figure 5.3A). The EEG data used contained considerable peaks at the alpha range (8-12 Hz), and the default parameter settings in DCM for CSD failed to produce satisfactory fits to these peaks when inspected visually (see van Wijk et al., 2018, *p.* 824). To address this issue, we doubled the number of maximum iterations to 256 and estimated the models with two adjustments to the hyperparameters: first, we set the shape of the neural innovations (i.e. the baseline neuronal activity) to flat (-32) instead of the default mixture of white and

pink (1/f) components (Moran et al., 2009). Second, we increased the noise precision value from 8 to 12 to bias the inversion process towards accuracy over complexity (see Friston et al., 2012 and Moran et al., 2009 for a detailed description of DCM for cross-spectral densities). In addition, for LAR the number of spatial modes was increased to 14 instead of the default of 8. The modes here refer to a reduction of the dimensionality of the data (done for computational efficiency) by projecting the data onto the principal components of the prior covariance, such that a maximum amount of information is retained (David et al., 2006; Fastenrath, Friston, & Kiebel, 2009; Kiebel, Garrido, Moran, & Friston, 2008).

These adjustments led to our full models (i.e. DMN, SAL, CEN, and LAR) converging with satisfactory fits (inspected visually) to the spectrum for 30/40 subject model instances (similar fits to what can be seen as the end result in figure 5.2). We then applied Bayesian Parameter Averaging (BPA) for each of the full models separately, averaging over the posteriors from the subject model instances that did converge and setting these averaged posteriors as new priors for the respective non-converged subject model instances. Estimating these subject model instances again with these BPA-derived priors produced satisfactory fits for all 10 remaining instances. Finally, we estimated all the full models again for all the participants with setting the posteriors from the earlier subject model estimations as updated priors, but this time with the neural innovations and noise precision set back to default settings. In doing so, all the models produced satisfactory fits with the default parameter settings for all of the participants (see figure 5.2).

To validate that the priors we used in the final inversion were suitable, we compared the group-level model evidence obtained with and without the adjusted noise levels. With all full models, the default hyperparameter settings with the updated priors generated better model evidence (difference in free energies for LAR, DMN, SAL, and CEN were +47260, +9440, +15700, and +660, respectively). To qualitatively assess the model fits, the observed and model-predicted cross-spectra were visually compared in each participant and judged sufficiently similar. To be sure about our conclusions, we also performed the PEB modelling (see below) leaving out the fitted subject model instances that produced the worst fits (1-2 per model); this had no notable influence on the interpretation of the results. The same approach was followed when inverting the full models separately for individual states of consciousness (figure 5.3B); in addition to the full models, here the BPA was also restricted to the same state of consciousness. The model-predicted and original spectral densities averaged over participants are shown in figure 5.2A, 5.2B, 5.2C, and 5.2D for LAR, DMN, SAL, and CEN, respectively.
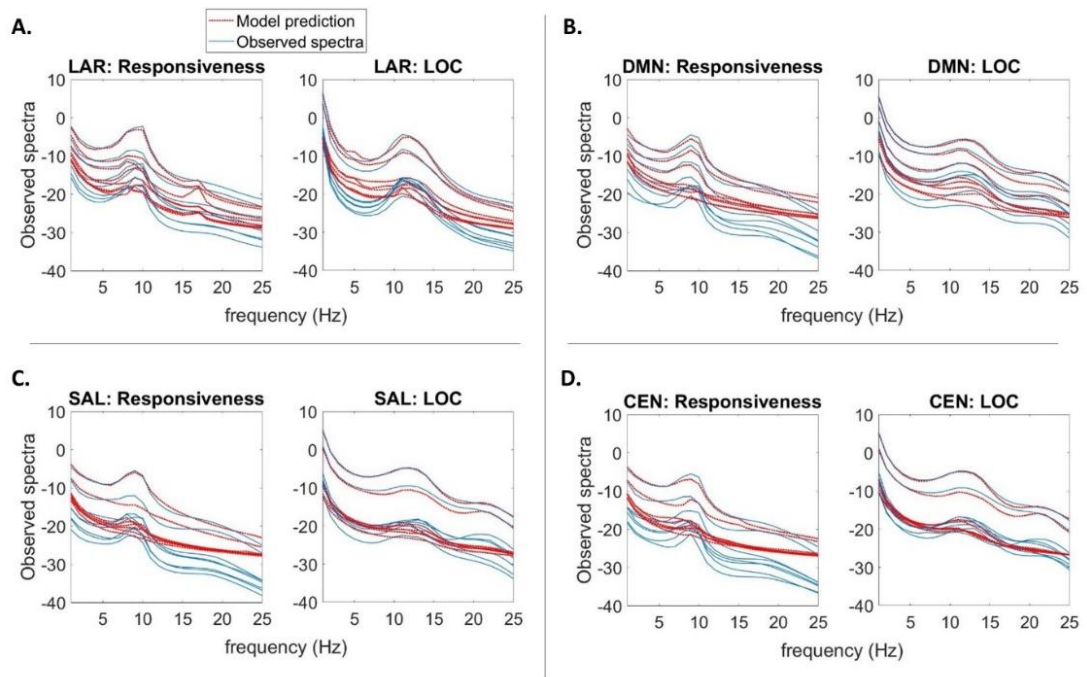
Figure 5.2. Average model fits. **A-D.** Subject-averaged power spectra of the observed EEG channel-space data, juxtaposed with that predicted by the fitted DCM models of each RSN, in normal behavioural responsiveness and loss of consciousness (LOC). Individual lines reflect the first eight spatial modes. LAR – large model, DMN – default mode network, SAL – salience network, and CEN – central executive network.

5.3 Parametric Empirical Bayes

In DCM, a variational Bayesian scheme called Variational Laplace is used to approximate the conditional or posterior density over the parameters given by the model inversion process, by maximizing a lower bound (the negative free energy) on the log-evidence (Friston et al., 2007). The Parametric Empirical Bayes (PEB) framework is a relatively recent supplement to the DCM procedure

used, for example, to infer the commonalities and differences across subjects (Friston et al., 2016). Briefly, the subject-specific parameters of interest (here, effective connectivity between nodes in a DCM model) are taken to the group-level and modelled using a General Linear Model (GLM), partitioning the between-subject variability into designed effects and unexplained random effects captured by the covariance component. The focus is on using Bayesian model reduction (BMR) – a particularly efficient form of Bayesian model selection (BMS) – to enable inversion of multiple models of a single dataset and a single hierarchical Bayesian model of multiple datasets that conveys both the estimated connection strengths and their uncertainty (posterior covariance). As such, it is argued that hypotheses about commonalities and differences across subjects can be tested with more precise parameter estimates than with traditional frequentist comparisons (Friston et al., 2016).

A particular advantage of PEB is that as part of the BMR process – when no strong a priori hypotheses about the model structure exist, as in the present study – a greedy search can be used to compare the negative free energies for the reduced models, iteratively discarding parameters that do not contribute to the free energy (originally 'post-hoc DCM analysis', Friston & Penny, 2011; Rosa, Friston & Penny, 2012). The procedure stops when discarding any parameters starts to decrease the negative free energy, returning the model that most effectively trades-off goodness of fit and model complexity in explaining the data. Last, a Bayesian Model Average (BMA) is calculated over the best 256 models weighted by their model evidence (from the final iteration of the greedy search). For each connection, a posterior probability for the connection being

present vs. absent is calculated by comparing evidence from all the models in which the parameter is switched on versus all the models in which it is switched off. Here, we applied a threshold of >.99 posterior probability, in other words, connections with over .99 posterior probability were retained.

For the DCMs that were fitted to the contrast between two states of consciousness using the procedure described in the previous section, we used PEB for second-level comparisons and Bayesian model reduction to find the most parsimonious model that explained the contrast by pruning away redundant connections. The focus was explicitly on the group-level comparison of the connectivity modulations (B-matrix). The whole sequence of steps is summarized in figure 5.3A.

A. Model fitting pipeline for PEB

Specify the full model

Estimate the full model
DCM.xU.X = [0 1]'

Specify and estimate a PEB model (B-matrix)
M.X = ones(N,1)

Search over nested models – pruning using BMR

Group effects & The best model

B. Model fitting pipeline for LOSOCV

Specify the full model

Estimate the full model *separately for the conditions*
DCM.xU.X = [ ]

LOSOCV with PEB
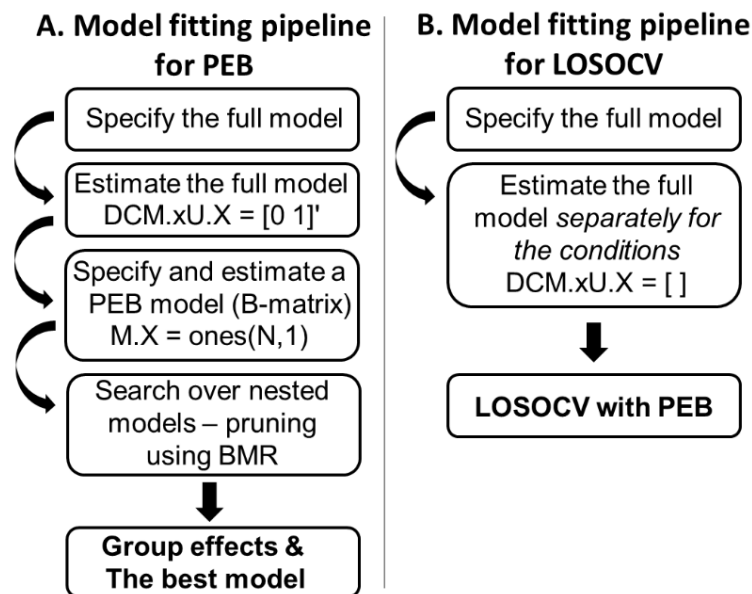
Figure 5.3. Modelling pipelines. **A.** The pipeline for inverting the DCM models in terms of changes in connectivity that explain the differences in the empirical

data observed in LOC as compared to wakeful consciousness baseline. The DCM model inversion was followed by PEB modelling with BMR to find the most parsimonious model and the modulatory effects on the group-level effective connectivity. **B.** The pipeline for inverting the DCM models separately for individual states of consciousness. This was done as a prerequisite for the LOSOCV classification with PEB modelling.

5.3.1 Leave-one-out cross-validation paradigm

As a crucial form of validation of our modelling framework, we investigated which network connections are predictive of the state of consciousness in unseen data. We adapted a standard approach in computational statistics, leave-one-subject-out cross-validation (LOSOCV; spm_dcm_loo.m). Here, we iteratively fitted a multivariate linear model (as described in detail in Friston et al., 2016) to provide the posterior predictive density over connectivity changes, which was then used to evaluate the posterior belief of the explanatory variable for the left-out participant: in the present case, the probability of the consciousness state-class membership.

To conduct LOSOCV analysis, the DCM models were now fitted to each state of consciousness separately, as shown in the procedure visualised in figure 5.3B. To cross-validate a fitted DCM model, both datasets from one participant were left-out each time *before* conducting PEB for the training dataset, and the optimised empirical priors were then used to predict the state of consciousness (behavioural responsiveness/LOC) to which the datasets from the left-out

participant belonged (see Friston et al., 2016 and Zeidman et al., 2019 for details). This procedure, repeated for each participant, generated probabilities of state affiliation, which were used to calculate the Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values with 95% point-wise confidence bounds across the cross-validation runs (see MATLAB perfcurve). In addition, the corresponding binary classification accuracy was calculated as the sum of true positives and true negatives divided by the sum of all assigned categories, i.e. (TP+TN) / (TP+TN+FP+FN), where TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

It is worthwhile to note, that – excluding the specific alterations reported in section 5.2.4 – we have estimated the DCM models using the default parameter settings recommended in the literature (Ashburner et al., 2017; Friston et al., 2003; Friston et al., 2012; Kiebel, Garrido, Moran, Chen, & Friston, 2009). This is also true for the LOSOCV procedure; no tweaking or adjusting of the hyper parameters within the model was performed. Here, we trained the model with the data from all but the left-out participant (training set), and predicted the state based on the data from the left-out participant (test set) and repeated this procedure by leaving out a different participant each time.

We first estimated LOSOCV metrics for all connections in all models. Next, LOSOCV metrics of subsets of hypothesis-driven connections were tested; the connections were divided into frontal, parietal, frontoparietal, and between-RSN subsets, based on the anatomical location of the connected nodes. The rationale was to investigate where in the brain the most consistent inter-subject-level

effects were located, in addition to the largest effect sizes identified by the PEB analysis.

Finally, we extended our validation of the DCM models by introducing a more difficult classification problem: we used the DCM parameters from responsiveness and LOC for training, and then tested them on unseen data collected during the post-drug recovery state of each subject (recovery state prediction). Again, during training, both datasets (behavioural responsiveness/LOC) from one participant were left out each time *before* conducting PEB, and the optimised empirical priors were then used to predict the state of consciousness to which the recovery-dataset from the left-out participant belonged. We hypothesised that if our modelled effects are valid, it should classify the recovery state as behavioural responsiveness rather than LOC - even though recovery is not identical to normal wakeful responsiveness, it is clearly closer to normal responsiveness than LOC. Here, we used recall - as calculated by (true positive) / (true positive + false positive) - and mean posterior probability for responsiveness to quantify classification performance. The 95% CIs were calculated over the posterior probabilities using a simple approximation for the unbiased sample standard deviation (Gurland & Tripathi, 1971).

5.4 Results

5.4.1 Dynamic causal modeling and parametric empirical Bayes

Our goal was to investigate the effective connectivity modulations caused by anaesthesia-induced loss of consciousness on three resting state networks

together and separately. We modelled time-series recorded from two states of consciousness – wakeful behavioural responsiveness and loss of consciousness (LOC) – with DCM for CSD at a single-subject level, followed by PEB at the group-level. In doing so, we estimated the change in effective connectivity with RSNs during LOC, relative to behavioural responsiveness before anaesthesia. For the DMN, we estimated 12 inter-node connections, and for both SAL and CEN 16 connections. With LAR, in addition to including all the connections in each RSN, additional connections were specified to model the modulatory effects of anaesthesia on between-RSN connections, increasing the estimated inter-node connections to fifty.

Following the inversion of the second-level PEB model, a greedy search was implemented to prune away connections that did not contribute significantly to the free energy using BMR. This procedure was performed for LAR and for all the three resting state networks separately. The most parsimonious model (A) and estimated log scaling parameters (B) for LAR, DMN, SAL, and CEN are shown in figures 5.4-5.7, respectively. Here, we applied a threshold of >.99 for the posterior probability; in other words, connections that were pruned by BMR and connections with lower than .99 posterior probability with their respective log scaling parameter are faded out (figures 5.4B-5.7B).

Of the fifty connections in the large model (figure 5.4), five were pruned away by BMR. The results indicate that typically effective connectivity decreased going from behavioural responsiveness to LOC between nodes in the DMN, with parietal connections showing consistent and large decreases.

135

Similarly, between-RSN parietal connections linking DMN and CEN also decreased. Backward connections between the dACC and PCC/precuneus, linking the DMN and SAL, increased slightly. A clear majority of connections forming the SAL and CEN networks increased.

On inverting the DMN separately (figure 5.5), we found that no connections were pruned away by BMR. In other words, all of the effective connectivity in the DMN was modulated by the loss of consciousness. In particular, forward connectivity to and from PCC/precuneus largely decreased, whereas direct parietofrontal forward connectivity from lateral parietal cortices to the medial prefrontal cortex was increased. Backward connectivity between all the sources was increased.

In contrast, seven connections out of 16 were pruned away from the full SAL model when it was inverted separately (figure 5.6). These consisted of all but one lateral connections between both, the lateral prefrontal nodes and lateral parietal nodes, and all but one backward connection originating from the dACC. The strength of change in connectivity within the SAL was lower than in DMN, and all but one of the retained connections showed an increase in strength when losing consciousness.

When inverting the CEN separately, two connections were pruned away (figure 5.7). Most of the retained connections showed a small increase in strength, with the largest effects in frontoparietal connections from the dmPFC to the left superior parietal cortex. Further, right hemisphere frontoparietal connections showed more modulatory changes than left hemisphere connections.
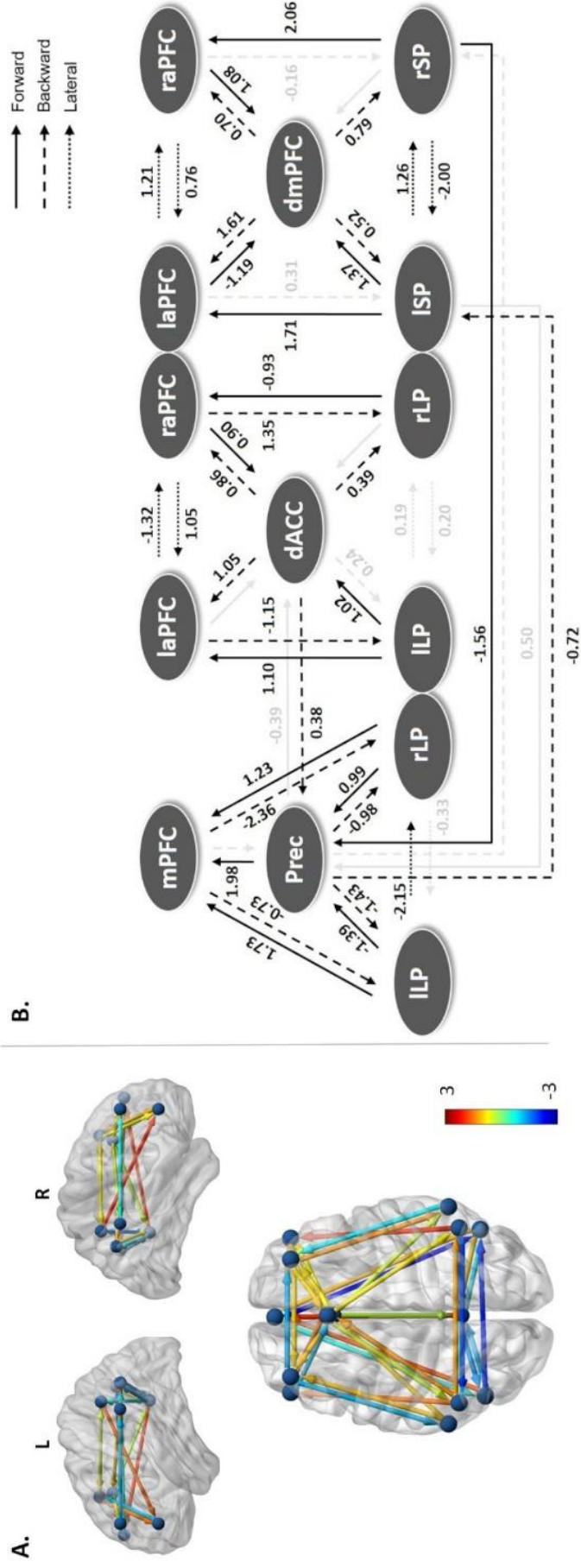
Figure 5.4. Estimated model parameters for LAR. **A.** Effective connectivity modulations on the most parsimonious LAR model. 5 connections were pruned away by BMR and a further 8 had lower than .99 posterior probability of being present. Colour shows modulation strength and direction. **B.** The log scaling parameters for the connections in the large model after BMR and BMA. Connections that were pruned by BMR and connections with lower than .99 posterior probability with their respective log scaling parameter are faded out. mPFC – medial prefrontal cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral parietal cortex, laPFC – left anterior prefrontal cortex, raPFC – right anterior prefrontal cortex, dACC – dorsal anterior cingulate cortex, dmPFC – dorsomedial prefrontal cortex, lSP – left superior parietal cortex, and rSP – right superior parietal cortex.
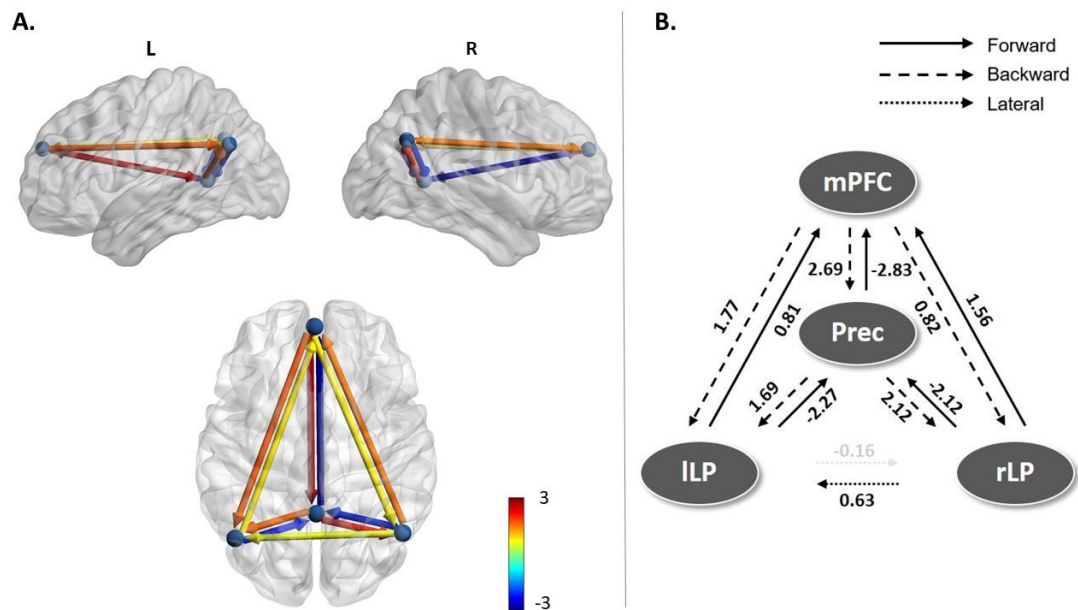
Figure 5.5. Estimated model parameters for DMN. **A.** Effective connectivity
modulations on the most parsimonious DMN model. Colour of connections show
strength and direction of modulation. None of the connections were pruned away,
and only one connection had lower than .99 posterior probability. **B.** The log
scaling parameters for the connections in DMN after BMR and BMA. The
below-threshold posterior probability connection with its corresponding log
scaling parameter is faded out. mPFC – medial prefrontal cortex, Prec – posterior
cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral
parietal cortex.

Figure 5.6. Estimated model parameters for SAL. **A.** Effective connectivity
modulations on the most parsimonious model for SAL. 7 connections were
pruned by BMR. **B.** The log scaling parameters for the connections in SAL.
Several connections were pruned away (faded out). The retained connections
were almost all positive modulations, but smaller in strength than in the DMN.
laPFC – left anterior prefrontal cortex, raPFC – right anterior prefrontal cortex,
dACC – dorsal anterior cingulate cortex, lLP – left lateral parietal cortex, rLP –
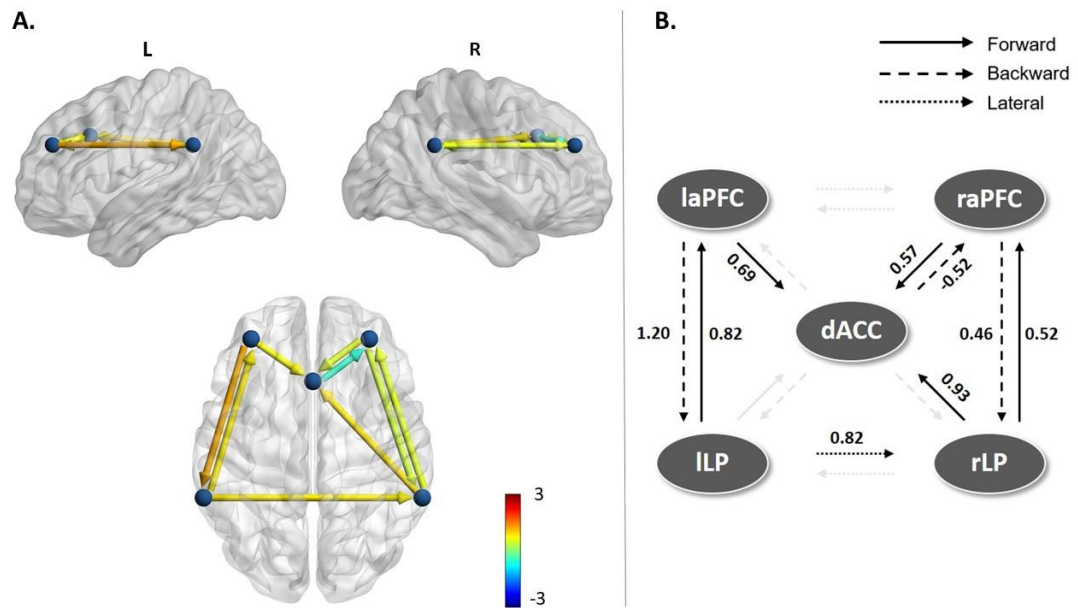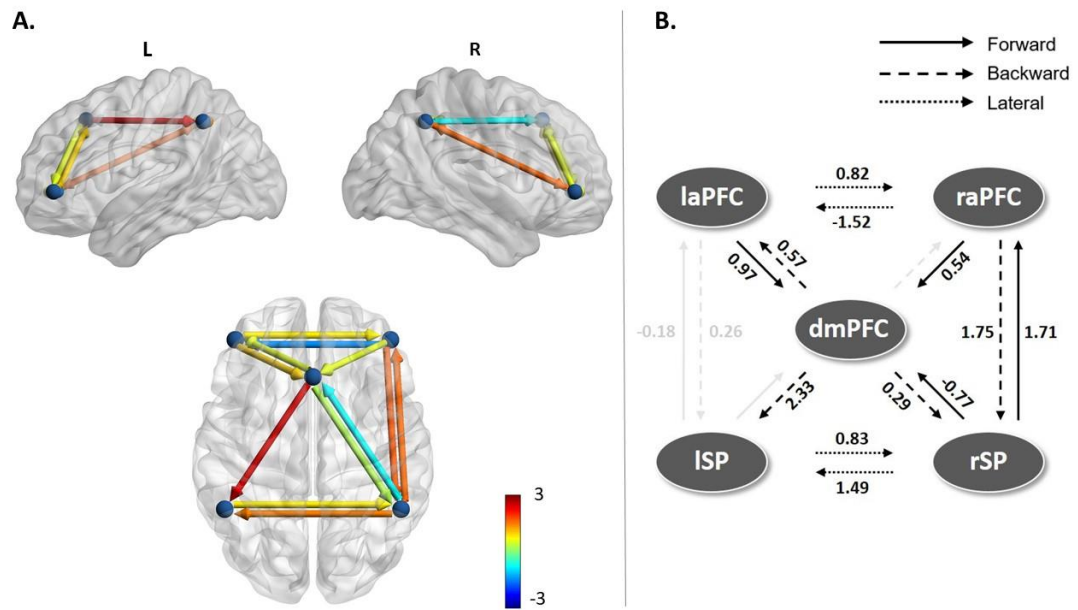right lateral parietal cortex.

Figure 5.7. Estimated model parameters for CEN. **A.** Effective connectivity modulations on the most parsimonious model for CEN. 2 connections were redundant in addition to 2 connections having lower than .99 posterior probability for being switched on. **B.** The log scaling parameters for the connections in CEN. Pruned connections and low posterior probability connections with the corresponding log scaling parameters are faded out. Effects on the remaining connections were almost all positive modulations, with strengths in-between those observed in the SAL and DMN. laPFC – left anterior prefrontal cortex, raPFC – right anterior prefrontal cortex, dmPFC – dorsomedial prefrontal cortex, lSP – left superior parietal cortex, and rSP – right superior parietal cortex.

5.4.2 Leave-one-subject-out cross-validation

To conduct LOSOCV, the DCM models were inverted again, this time for each state of consciousness in each subject separately. With the states modelled separately, PEB was conducted repeatedly (on the training set in each cross-validation run) alongside LOSOCV analysis to generate AUC values (see Methods). The AUC/ROC values for all full models are shown in figure 5.8A, and table 5.2 shows all tested AUC values with accuracy for all tested sets of connections. The results indicate that leave-one-subject-out cross-validated predictions based on the LAR and SAL models had accuracy significantly different from chance, i.e. with the lower bound of the 95% CI of the AUC above chance. However, for predictions based on the DMN and CEN, the lower bound of the 95% CI of the predictions did not exceed chance.

To understand whether specific connections within cortical brain networks were driving changes in consciousness, we evaluated the predictive power of four different hypothesis-driven subsets of connections – frontal, parietal, frontoparietal, or between-RSN – to predict the two states of consciousness in left-out subjects. As shown in figure 5.8B, frontoparietal connectivity in LAR, DMN, and SAL produced the best predictions of the state of consciousness with LOSOCV. Further, the posterior subset in the SAL performed statistically better than chance. None of the subsets in the CEN reached statistical significance.

Finally, the predictive power of these RSN connectivity subsets were tested in a more difficult classification problem: each model subset was trained on behavioural responsiveness and LOC, and then tested on the previously unseen

'recovery' state, the data that was collected after the participant regained

consciousness. In figures 5.9A and 5.9B each data point represents one

participant. Figure 5.9A shows the mean posterior probabilities of the recovery

state being correctly classified as behavioural responsiveness when using all

connections in a model as predictors. Figure 5.9B shows the same results for the

frontal, parietal, frontoparietal, and between-RSN connections as predictors.

When predicting with all connections, only classifications based on all

connections in LAR performed significantly better than chance. With the

hypothesis-driven subsets of connections, frontoparietal connectivity within the

DMN generalised best to the recovery state. Only one other subset – parietal

connections in SAL – performed significantly better than chance, and almost as

well as frontoparietal DMN connectivity (.82 vs. .79 posterior probability). All

subsets with LAR performed statistically better than chance, however, with poor

mean posterior probability values in comparison to DMN frontoparietal and SAL

parietal connections. Table 5.2 shows the mean posterior probabilities and the

corresponding recall values for all the tested connection sets and for all models.

We verified that the predictive accuracy (of the unseen recovery state) was not

driven by subject effects or bias, as evident in the individual posterior

probabilities plotted in figures 5.9C and 5.9D.
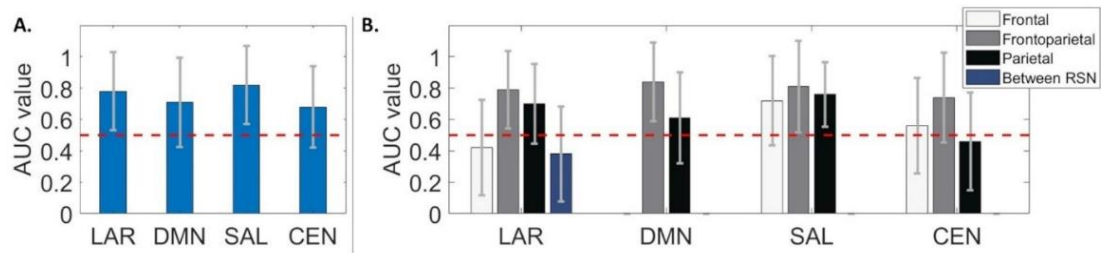
Figure 5.8. The AUC values for classifying the state of consciousness in LOSOCV paradigm. **A.** For the full models, only predictions based on LAR and SAL performed statistically better than chance (red dashed line), with classifications based on the connections in SAL reaching the overall best prediction. The error bars represent the 95% point-wise CI calculated using leave-one-out cross-validation for both A and B (MATLAB perfcurve). **B.** AUC values for hypothesis-driven connections for all models in LOSOCV paradigm. The DMN is missing frontal connections as it had only one anterior node. Best prediction performance was obtained with frontoparietal connections in LAR, DMN, and SAL. Further, predictions based on posterior SAL connections reached statistical significance.
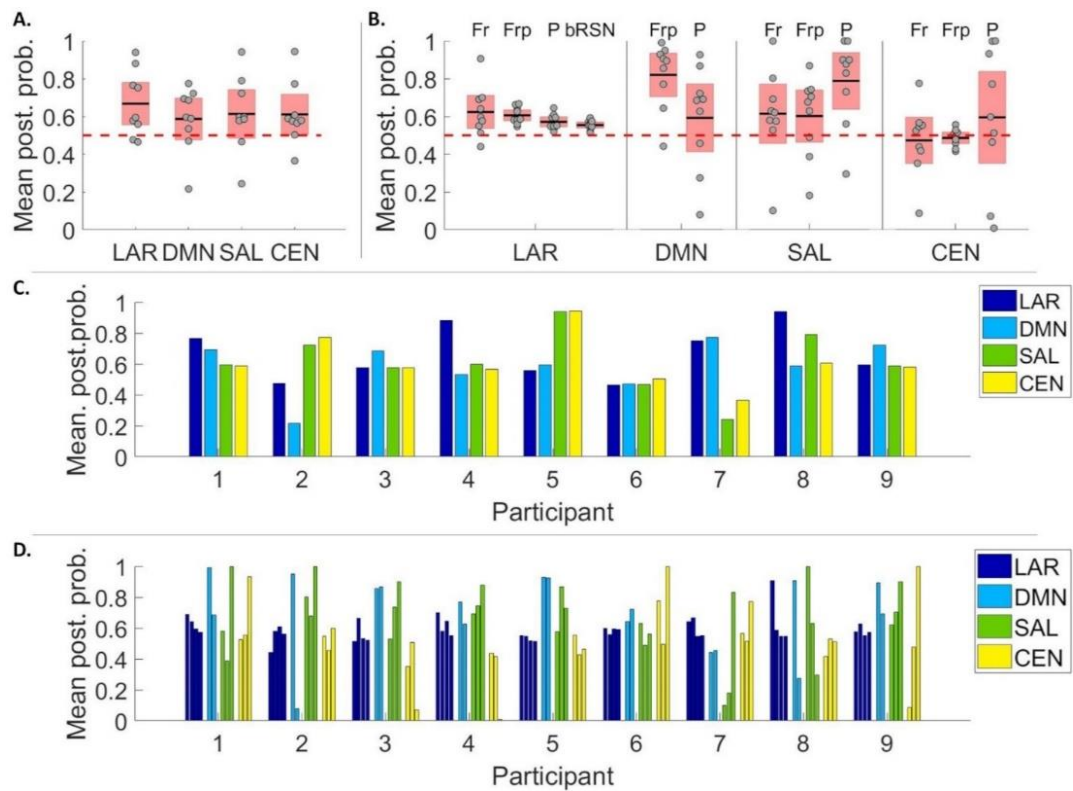
Figure 5.9. Mean posterior probabilities for prediction of recovery data. On panels A and B the individual data points represent individual participants. **A.** Predictions based on all connections in LAR performed better than chance (red dashed line). Data points representing participants are laid over a 1.96 SEM (95% confidence interval over posterior probabilities) in red with the black lines marking the mean. **B.** Mean posterior probabilities for hypothesis-driven connection subsets of all models in the recovery state: top labels refer to frontal (Fr), frontoparietal (Frp), parietal (P), and between-RSN (bRSN) connections. DMN frontoparietal connectivity had the best performance across all sets and all models. Parietal connections in SAL performed statistically better than chance but with lower posterior probability value in comparison to DMN frontoparietal connections. All subsets with LAR performed statistically better than chance,

however, with poor posterior probability values in comparison to DMN frontoparietal and SAL parietal connections. **C-D.** Posterior probabilities predicted for individual datasets, based on all connections (C) and on hypothesis-driven subsets (D). In Panel D, the individual bars depict different connection subsets: frontal, frontoparietal, parietal, and between-RSN in LAR, frontoparietal and parietal in DMN, and frontal, frontoparietal, and parietal in SAL and CEN.

Table 5.2. AUC (accuracy) values calculated with LOSOCV, and mean posterior probabilities (recall) in the recovery state, for all connections, all hypothesis-driven connection subsets (frontal, parietal, frontoparietal, and between-RSN connections), and for the large model (LAR), default mode network (DMN), salience network (SAL), and central executive network (CEN). No values are given if no such connection-subsets exist for the model. Accuracy/recall values were not calculated for connection subsets with performance close to chance (between 0.4 - 0.6). * indicates significance estimated at 95% confidence intervals in both AUC and posterior probability.

Model	Responsiveness/LOC AUC (Accuracy)		Recovery Mean PP. (Recall)	
	All connections		**All connections**	
LAR	0.78 (0.80)*		0.67 (0.78)*	
DMN	0.71 (0.70)		0.59 (--)	
SAL	0.82 (0.80)*		0.61 (0.78)	
CEN	0.68 (0.70)		0.61 (0.89	
	Frontal	**Parietal**	**Frontal**	**Parietal**
LAR	0.42 (--)	0.70 (0.65)	0.62 (0.89)*	0.57 (--)*

DMN	--	0.61 (0.65)	--	0.59 (--)
SAL	0.72 (0.65)	0.76 (0.65)*	0.61 (0.89)	0.79 (0.89)*
CEN	0.56 (--)	0.46 (--)	0.47 (--)	0.60 (--)

	Frontoparietal	BRSN	Frontoparietal	BRSN
LAR	0.79 (0.80)*	0.38 (0.55)	0.61 (1.00)*	0.55 (--)*
DMN	0.84 (0.85)*	--	0.82 (0.89)*	--
SAL	0.81 (0.75)*	--	0.60 (--)	--
CEN	0.75 (0.70)	--	0.49 (--)	--

5.5 Discussion

We computationally evaluated the evidence for the posterior hot zone theory of consciousness by modelling the relative contributions of three resting state networks (DMN, SAL, and CEN) for propofol-induced LOC. Using the recently introduced PEB framework, we characterised modulations in effective connectivity accompanying the loss of consciousness within and between these key RSNs. We found a selective breakdown of posterior parietal and medial feedforward frontoparietal connectivity within the DMN, and of parietal inter-network connectivity linking DMN and CEN. These results contribute to the current understanding of anaesthetic-induced LOC, and more generally to the discussion of whether the neural correlates of consciousness are predominantly anterior (Del Cul et al., 2009), frontoparietal (Bor & Seth, 2012; Chennu et al., 2014; Chennu, O'Connor, Adapa, Menon, & Bekinschtein, 2016; Laureys & Schiff, 2012), or posterior (Koch et al., 2016; Koch et al., 2016b; Siclari et al., 2017).

147

We used a novel DCM-based cross-validation to establish the predictive validity of our models, addressing an issue commonly present in DCM studies, including previous consciousness-related DCM studies - that the best model identified by BMS is only the best model among the models tested. Significant generalisation performance with cross-validation increases the level of confidence we can ascribe to our results. This analysis highlighted that frontoparietal effective connectivity consistently generated accurate predictions of individual states of consciousness. Furthermore, we demonstrated generalisation of this predictive power by showing that effective frontoparietal connectivity within the DMN and parietal connectivity within the SAL predicted the state of consciousness in unseen data from the post-anaesthetic recovery state.

With the large model combining all three RSNs, we observed consistent and widespread decreases in connectivity between posterior DMN nodes and between parietal connections linking DMN and CEN (figure 5.4). With the individual RSNs, we observed a selective breakdown of the DMN, specifically, decreases in feedforward connectivity to and from PCC/precuneus (figure 5.5). It is worth highlighting that most decreases in effective connectivity - both when the RSNs were modelled individually and as one large network - were between nodes located within the posterior hot zone, and related specifically to PCC/precuneus – a key structure in the hot zone (Koch et al., 2016; Siclari et al., 2017). In other words, the network-level breakdown characterising the difference between behavioural responsiveness and LOC was mostly located within the parietal hot zone.

In the SAL and CEN networks, when fitted on their own, several connections were pruned away by BMR, with small increases in the majority of preserved connections; ¼ of the connections in CEN and almost half of the connections in SAL (7 out of 16) were pruned, in contrast to the DMN in which no connections were pruned (figures 5.6 and 5.7). The same pattern was present, although to a smaller degree, when the three RSNs were estimated together (LAR): fewest of the connections pruned were in the DMN, when compared with the SAL and CEN networks. This highlights the relative importance of the DMN over the SAL and CEN in explaining differences between states of consciousness and is consistent with the previous evidence from disorders of consciousness (Crone et al., 2011; Fernández-Espejo et al., 2012; Laureys, 2005; Laureys et al., 1999), anaesthesia (Boveroux et al., 2010), and sleep (Horovitz et al., 2009).

It is important to note, however, that there are multiple possible approaches to parameter estimation in DCM, both at the individual and at the group-level. The joint estimation method we chose utilises BMR and PEB. An alternative would be a step-by-step approach, which uses individually estimated RSN posteriors as fixed priors when fitting the LAR, thereby reducing the number of free parameters. The joint estimation method hence enables us to fit comparatively larger models, but potentially with a risk of a more complex free energy landscape (Litvak et al., 2019). Due to these modelling choices, we have limited our granularity of our inference to models and cortical regions within them, instead of interpreting the posterior densities of all possible fitted model parameters. The fact that we were able to demonstrate out-of-sample

generalisation using our fitted models gave us confidence that the methodology was valid.

Keeping the above in mind, we did find that PCC/precuneus-related feedforward connectivity in the DMN is impaired during LOC. This is in contrast to two previous DCM studies of propofol-anaesthesia, which have suggested either selective impairments in frontoparietal feedback connectivity from dACC to PCC (Boly et al., 2012), or subcortico-cortical modulations from globus pallidus to PCC (Crone et al., 2017). However, there are major methodological differences between the present study and the previous two that could explain these different results. Firstly, the examined model space was different. Secondly, both previous studies used models with only two cortical nodes summarising activity of frontal and parietal regions. They did not implement a wide search over a large model space using BMR and instead focused on evaluating a small number of hypothesis-specific models. We adopted a broader approach to model formulation and evaluation. In doing so, we expand upon these previous results by suggesting a selective breakdown of PCC/precuneus-related forward connectivity within the DMN. Our results differed from Boly et al. (2012) even when the direct connections between dACC and PCC/precuneus were modelled (in LAR) – we found an increase in feedback connectivity from dACC to PCC/precuneus and a small, low probability decrease in feed-forward connectivity. Our results are, however, in line with previous studies showing increased frontoparietal connectivity with partial directed coherence (Maksimow et al., 2014) and with Granger Causality (Barrett et al., 2012; Nicolaou, Hourris, Alexandrou, & Georgiou, 2012) during anaesthesia.

It is noteworthy that impaired feedforward connectivity has been suggested to be the main modulation caused by propofol-anaesthesia in a recent DCM study with TMS-evoked potentials by Sanders et al. (2018). Their models consisted of 6 cortical sources (bilateral inferior occipital gyrus (IOG), bilateral dorsolateral PFC, and bilateral superior parietal lobule (SPL). They found predominantly impaired feedforward connectivity from right IOG to right SPL (specifically with theta/alpha-gamma coupling). Although they suggested that resting state activity was driven by feedback connectivity, while induced responses were driven by feedforward connectivity, it may be that restricting modulations to just two free parameters (connections) in the cortex simplifies the effects of propofol-induced LOC to the degree that they differ from estimations of more complex models.

Finally, the observed *increase* in effective connectivity between specific nodes (especially front-to-back) has been suggested previously to be due to the drug-specific effects of propofol rather than changes in states of consciousness (Maksimow et al., 2014). Hence, it may be that the relatively uniform increases in connectivity in the SAL and CEN, and the increased feedback connectivity in the DMN, were specific to propofol.

While the results of the LOSOCV cross-validation should be interpreted with caution given the limited number of participants in our study, the results indicated that, when using all connections, the above-chance prediction performance of conscious state was only obtained with LAR and SAL, with the latter performing the best (figure 5.8A). With smaller, hypothesis-driven subsets, we found that the frontoparietal connections provided consistently the most

151

accurate predictions in all models except the CEN (figure 5.8B). When predicting

the unseen state of recovery (figure 5.9B), frontoparietal DMN connections

performed the best, followed by parietal connections in SAL. It is worth

highlighting that the frontoparietal DMN and parietal SAL connections predict

the state correctly, even when the state actually differs from the true training

state; recovery differs from normal wakeful responsiveness not only

behaviourally, but also in terms of the residual propofol in the blood. However,

the participants are conscious and responsive, and thus, recovery is considered as

a state clearly closer to normal wakeful responsiveness than LOC.

Taken together, our prediction results highlighted an important role for

frontoparietal connections. This is perhaps not surprising, as wakeful awareness

is known to recruit the DMN (Raichle & Snyder, 2007); maintaining a state of

conscious responsiveness requires an interaction between the posterior hot zone

(the role of which is highlighted when modelling the *change* between states) and

frontal areas, mediated by the frontoparietal connections. Previous literature has

suggested dynamic changes in connectivity between brain networks during

cognitive control (Cocchi, Zalesky, Fornito, & Mattingley, 2013; Leech, Braga,

& Sharp, 2012) and anaesthetic-induced loss of consciousness (Luppi et al.

2019). The importance of frontoparietal connections in the present study when

predicting states of behavioural responsiveness – a state of higher integration

than LOC – is consistent with the notion that conscious, behavioural

responsiveness requires a brain-wide "global workspace" supported by the

frontoparietal network (Baars, 1997; Dehaene & Changeux, 2011; Dehaene et al.,

2011; Mashour, Roelfsema, Changeux, & Dehaene, 2020). Hence, it is perhaps

no surprise that the role of frontoparietal connections became prominent when we predicted individual states of consciousness rather than the contrast between them.

A number of previous studies have suggested a pivotal role of subcortical structures in transitions to unconsciousness (e.g. Baker et al., 2014; Liu et al., 2013; White & Alkire, 2003). Crone et al. (2017) reported a breakdown of connectivity between the globus pallidus and posterior cingulate cortex connectivity during LOC, followed by a reversal at recovery. It remains a possibility that the effective connectivity modulations found in the present study – especially in relation to the PCC/precuneus - are driven by subcortical structures that we did not model here, given the limitations of scalp EEG signals (Goldenholz et al., 2009). It might be worthwhile to further investigate the effects of LOC with fMRI DCMs, including large-scale models combining cortical and subcortical nodes with PEB with BMR to conduct a wider exploration of the model space.

In addition to the modelling being limited only to cortico-cortical connections, some of our results are arguably propofol-specific; for example, very different alterations have been observed between propofol and ketamine (Driesen et al., 2013; Sarasso et al., 2015). Hence, it may be that modelling the cortical effects of other anaesthetic agents would lead to very different sets of results. Further, we have modelled the effects using DCM and the standard ERP neuronal model, rather than modelling frameworks designed to capture more fine-grained properties of the EEG spectrum during anaesthesia (see for example

Bojak & Liley, 2005; Hutt & Longtin, 2010). DCM and the ERP neuronal model were chosen primarily in order to produce results that could be compared with the prior DCM work on modelling consciousness. Furthermore, we aimed to model consciousness at the network level, rather than at the level of the known molecular effects of propofol, e.g., prolongation of inhibitory post-synaptic potential time constants, that are known to take place within individual cortical and sub-cortical sources. A valuable future direction would be to investigate the predictive power of such effects and the extent to which they may drive the modulations in extrinsic connectivity. This could be done, for example, by using the LFP model or the Canonical Microcircuits model which are better suited for estimating the intrinsic connectivity and the molecular effects within the sources (Bastos et al., 2012; Moran et al., 2007). Lastly, as we tested only a pre-specified model space, the limitations imposed by this scope might have missed important mechanisms of conscious awareness not modelled here.

Notwithstanding these points, our results highlight a selective breakdown of inter- and intra-RSN effective connectivity in the parietal cortex, reinforcing the role of the posterior hot zone for human consciousness. However, modulations of frontoparietal connections were consistent enough to predict states in unseen data, demonstrating their causal role in maintaining behavioural responsiveness.

6. Effective connectivity in traumatic disorders of consciousness

This chapter continues our work exploring the brain connectivity in disorders of consciousness (DoC). We assess differences in effective connectivity between cortico-cortical regions of the default mode network (DMN) between DoC patients and healthy controls using dynamic causal modeling (DCM) for EEG. Following the methodology set in chapter 5, we first test the prospective performance of the DCM-connectivity within DMN in classifying states of consciousness in DoC based on two connectivity subsets: frontoparietal and posterior connections. We follow this by adopting a data-driven approach to the classification problem by investigating the predictive performance of single connections, and provide evidence for a key difference in left top-down frontoparietal connectivity in distinguishing unresponsive wakefulness syndrome (UWS) patients from minimally conscious patients (MCS+) and healthy controls. Further, we demonstrate that our DCM models generalise to a more difficult classification problem: in a leave-one-*state*-out cross-validation paradigm, we train the models on UWS patients with a confirmed PET negative diagnosis on the one hand, and either healthy controls or MCS+ patients on the other. The models are then tested on datasets from 'covertly aware' UWS patients with a PET positive diagnosis. With this generalisation, we provide evidence for the hypothesis that if our modelled effects are valid, and if the sustained PET metabolism reflects covert consciousness in the UWS PET+

155

patients, our model should classify these patients as healthy

controls/MCS+ rather than UWS PET-.

6.1 Introduction

The unresponsive wakefulness syndrome is defined by preserved arousal in

the absence of behavioural signs of awareness (periodic sustained eye opening

with unpurposeful movements; Laureys et al., 2010). In contrast, patients in the

minimally conscious state show fluctuating and incomplete awareness with

preserved arousal. The MCS state has been further divided into MCS- and

MCS+, with the latter condition characterised by command following, intelligible

verbalisation or gestural (or verbal yes/no responses) to spoken or written

questions (Bruno et al., 2011). These states are collectively known as disorders of

consciousness (with other states, such as coma and locked-in-syndrome).

The exclusive use of clinical consensus for diagnosing these disorders of

consciousness based on observed behaviours has been shown to result in high

rates of misdiagnosis of the true level of consciousness of the DoC patients,

especially in the case of UWS (van Erp et al., 2015; Schnakers et al., 2009).

Consequently, with the advent of modern neuroimaging techniques, there has

been increasing interest in characterising the underlying neuronal basis for the

presence or lack of awareness in DoC using structural and functional MRI

(Demertzi et al., 2015; Di Perri et al., 2016), PET (Laureys et al., 1999; Stender

et al., 2014), and EEG (Chennu et al., 2014; King et al., 2013; Sitt et al., 2014).

Structural and functional neuroimaging studies have suggested an important role of the DMN in DoC – an intrinsic brain network encompassing the posterior cingulate cortex/precuneus, bilateral parietal cortices, and the medial prefrontal cortex (Annen et al., 2018; Boly et al., 2009; Fernández-Espejo et al., 2012; Guldenmund et al., 2016; Soddu et al., 2012; Vanhaudenhuyse et al., 2010). Cerebral metabolism – as measured by PET – has been shown to differentiate UWS from MCS (Stender et al., 2014, 2016), with regional differences often in areas associated with DMN (Stender et al., 2015; Thibaut et al., 2012). This extends to minimally conscious patients; MCS+ can be differentiated from MCS- with the former group showing higher cerebral metabolism especially in left-sided cortical areas, including Broca's and Wernicke's areas, premotor, presupplementary motor, and sensorimotor cortices (Aubinet et al., 2020; Bruno et al., 2012).

Similarly, effective connectivity studies (as measured with DCM for fMRI) have suggested disruptions within the DMN specifically related to posterior cingulate cortex (PCC; Crone et al., 2015) and in subcortical networks potentially driving the disruptions in the DMN (Chen et al., 2018; Coulborn, Taylor, Naci, Owen, & Fernández-Espejo, 2021). As far as we are aware, just one study has used DCM with EEG for measuring and diagnosing cognitive functioning in DoC population. Using a mismatch negativity paradigm, Boly et al. (2011) showed that the difference between UWS and healthy controls was due to an impairment of backward connectivity from frontal to temporal cortices, emphasising the importance of top-down processing for conscious perception.

Importantly, a number of studies have reported "covert" voluntary brain activity in some seemingly unresponsive patients, with active and resting state paradigms (Bodart et al., 2017; Claassen et al., 2019; Chennu et al., 2017; Cruse et al., 2011; Lechinger et al., 2013; Monti et al., 2010; Owen et al., 2006; Owen & Coleman, 2008; Schnakers et al., 2015). However, it is unknown whether effective resting state connectivity between key nodes within the DMN, as measured with EEG, could be used to identify such covertly aware patients.

Here, as a preliminary investigation, we address this gap by using spectral DCM for EEG with parametric empirical Bayes (PEB). We investigate the difference in causal interactions between cortico-cortical regions of the DMN, between DoC patients (UWS and MCS+) and healthy controls. We then demonstrate the prospective performance of the connectivity within DMN in classifying states of consciousness in DoC. Based on previous studies (Boly et al., 2011), we hypothesise that there will be top-down/backward connectivity differences in UWS vs. healthy controls and in UWS vs. MCS+ comparisons.

Next, in a leave-one-subject-out cross-validation, we test the classification performance of models based on the full DMN network and on two connectivity subsets: the posterior connections and the frontoparietal connections. Following this, we adopt a data-driven approach to the classification problem by investigating the predictive performance of single connections. The aim here is to identify the direction and location of the largest, most consistent modulations between the subjects.

Finally, we demonstrate that our DCM models generalise to a more difficult classification problem: in a leave-one-*state*-out cross-validation paradigm, we train the models on UWS patients with a confirmed PET negative diagnosis (a complete bilateral hypometabolism of the associative frontoparietal cortex; see Methods) on the one hand and either healthy controls or MCS+ patients on the other. We then test the models on datasets from "covertly aware" UWS patients with a PET positive diagnosis (partially preserved metabolism and activity within these areas). We hypothesise that if our modelled effects are valid, and if the sustained PET metabolism reflects covert consciousness in the UWS PET+ patients, our model should classify these patients as healthy controls/MCS+ rather than UWS PET-.

6.2   Methods

6.2.1   Data acquisition

We assessed effective connectivity within the default mode network and whether modulation of this connectivity predicted states of consciousness in patients with Disorders of Consciousness (DoC). The patients included were referred to the University Hospital of Liège from clinical centres across Europe since 2008. The data collection was approved by the Ethics Committee of the University Hospital of Liège and the patients' legal guardians gave written informed consent. Data were also collected from healthy controls as a reference group, all of whom gave informed written consent before participation.

The dataset consisted of the patient data with 26 healthy controls ($N = 188$). From the dataset, we identified patients admitted due to traumatic brain injury (TBI; $N = 76$), and those diagnosed with unresponsive wakefulness syndrome (UWS; Laureys et al., 2010) or minimally conscious state plus (MCS+; Giacino et al., 2002; Bruno, Vanhaudenhuyse, Thibaut, Moonen, & Laureys, 2011). Patients admitted due to any other aetiology, e.g. anoxia or haemorrhage, and patients diagnosed with any other condition than UWS or MCS, were excluded from the further analyses. The patient groups were further divided based on their respective PET-scans – either into a PET-positive (PET+) or a PET-negative (PET-) sub-group (table 6.1). Amongst the healthy controls, we pseudo-randomly drew a cohort of 11 control subjects to adjust for the group-size discrepancies. There were no meaningful differences in the mean ages between the groups (in a Bayesian ANOVA the probability for the model including the main effect of age: p(M|data) = 0.247, Bayes factor = 0.328).

6.2.2 Behavioural and Positron Emission Tomography assessments

Patients were behaviourally assessed on the day of the PET and EEG imaging using the Coma Recovery Scale – Revised (CRS-R; Kalmar & Giacino, 2005) five to seven times a day, and the diagnosis was based on the highest score obtained.

Positron Emission Tomography Data Collection and Analysis Flurodesoxyglucose-PET (FDG-PET) scans were acquired from all the included patients using the methodology as described in Stender et al. (2014). The scans

were acquired on a Philips Gemini TF PET-CT scanner (Philips Medical Systems) approximately 30 minutes after an intravenous injection of 150 or 300 MBq of the radioactive tracer, fluor-18 fluordeoxyglucose (FDG). The brain imaging was obtained during an awake-period and while eyes open in a silent and dark room (ensured by an examiner present by administering tactile or auditory stimuli when the patients were closing their eyes). The data analysis identified relatively preserved and decreased metabolism in patients in comparison to controls and was conducted using Statistical Parametric Mapping (SPM8).

The analysis results were visually inspected by a trained neurologist to reach a PET+- or PET- -diagnosis following previous findings. A PET--diagnosis was produced by a complete bilateral hypometabolism of the associative frontoparietal cortex with no voxels with preserved metabolism, whereas PET+- diagnosis was produced by an incomplete hypometabolism and partial preservation of activity within these areas (Laureys, Owen, & Schiff, 2004; Nakayama, Okumura, Shinoda, Nakashima, & Iwama, 2006; Thibaut et al., 2012).

Table 6.1. The mean age in years (SD), the total number of patients, and the number of PET$^+$ and PET$^-$ of patients in each of the different DoC-groups. UWS – unresponsive wakefulness syndrome, MCS+ – minimally conscious positive.

Patient group	Mean age (SD) in years	N^{TOTAL}	PET+	PET-	Aetiology
UWS	30.7 (8.5)	11	5	6	TBI
MCS+	38.3 (10.3)	12	11	1	TBI

| **Controls** | 30.9 (6.7) | 11 | - | - | - |

6.2.3 EEG data acquisition and preprocessing

The EEG data collection is described in more detail in chapter 4. In short, the data consisted of high-density EEG recordings of 20-30 minutes (256-channels, EGI), acquired during the FDG uptake, just prior to the start of the PET-imaging. The patients were ensured to stay awake during the data collection.

The data were recorded at a sampling rate of either 250 Hz or 500 Hz (downsampled to 250 Hz). Data from the channels from the neck, cheeks, and forehead were discarded due to contributing most of the movement-related noise. As in chapter 4, we were left with the data from 173 channels on the scalp for further analysis. The raw signals were filtered from 0.5 – 45 Hz, with additional line noise removal at 50 Hz (notch-filter). We further restricted the DCM analysis to 1 – 30 Hz due to excessive high-frequency noise components. Via calculating the normalised variance, the excessively noisy channels and epochs were identified and either manually rejected or retained by visual inspection. Lastly, the data were re-referenced to a common average.

6.2.4 Dynamic causal modeling

A more detailed description of the dynamic causal modeling can be found in chapter 5. In short, we first imported the first 60 artefact-free 10-second epochs, in to SPM12 (Wellcome Trust Centre for Human Neuroimaging; www.fil.ion.ucl.ac.uk/spm/software/spm12). To analyse the resting effective

connectivity within the default mode network (DMN), DCM for EEG cross-spectral densities (CSD) was applied (Friston et al., 2012; Moran et al., 2009). Here, the observed cross-spectral densities in the resting-EEG are explained by a generative model that combines a biologically plausible neural mass model with an electrophysiological forward model that maps the underlying neural states to the observed data (ERP-model; Moran et al., 2013). The idea is to model the source activity over time in terms of causal relationships between interacting inhibitory and excitatory populations of neurons (see section 2.5.1 for detailed description).

Each source – or node – is connected to each other via extrinsic connections, while each subpopulation within each source is connected to each other via intrinsic connections. Here, however, we only estimated extrinsic connectvity between the nodes within the DMN to ensure identical methods and comparability of the results between chapters 5 and 6. Among the extrinsic connectivity, the top-down – or backward – connections are thought to have inhibitory and modulatory effects on the nodes they target, while forward connections are viewed as having a strong excitatory driving effect (bottom-up; Salin & Bullier, 1995; Sherman & Guillery, 1998).

Second-order differential equations describe the hidden state of neural activity within each node depending on both, the parametrised intrinsic and extrinsic connection strengths. This enables the computation of the linear mapping from the endogenous neuronal fluctuations to the EEG sensor spectral densities, and consequently, permits the modelling of differences in the spectra

due to changes in the underlying neurophysiologically meaningful parameters.

These parameters are describing, for example, the intrinsic and extrinsic

connectivity of coupled neuronal populations (i.e. sources) and their physiology.

For further information about EEG DCM, see for example Friston et al. (2012),

Kiebel et al., (2008) & Moran et al. (2009).

6.2.5 Model specification

Fitting a DCM model requires the specification of the anatomical locations of

the nodes/sources a priori. DoC patients typically have widespread structural

brain damage often accompanied by distributed white matter anomalies (Annen

et al., 2018; Fernández-Espejo et al., 2012; Tshibanda et al., 2009). Hence,

the feasibility and validity of applying DCM to damaged brains can be

questioned, especially with patients with anoxic brain damage (King,

Bekinschtein, & Dahaene, 2011). Here, we mitigated these concerns with the

following methods: firstly, we select only patients suffering from traumatic brain

injury. Traumatic aetiology has been associated with more focal injury centred

often on areas susceptible to rotational forces, acceleration and deceleration such

as the brainstem, midbrain, thalamus, hypothalamus, cerebellum, and posterior

corpus callosum (Guldenmund et al., 2016; Newcombe et al., 2010).

Second, we will use a special case of Bayesian model selection (BMS),

Bayesian model reduction (BMR) to invert multiple nested models from a single,

fully connected DMN (see Methods). A particular advantage here is that BMR

can be applied using an explanatory approach, in which no strong a priori

hypotheses about the model parameters are needed (Friston & Penny, 2011; Rosa et al., 2012; Zeidman et al., 2019). This enables more flexibility in the estimated model space. Third, the focus here is not so much to model the underlying connections within the DMN, as it is on comparing the predictive performance of effective connectivity represented by a set of DCM models.

Finally, to reduce the risk of spurious findings, we only modelled the DMN, which have been previously associated with DoC (see for example Boly et al., 2008; Crone et al., 2011; Crone et al., 2015; Heine et al., 2012; Lin et al., 2017). The node locations were the same as in chapter 5; the schematic representation and the node locations (from Razi et al., 2017) are shown in figures 6.1A and 6.1B, respectively (node locations visualized with the BrainNet Viewer (Xia, Wang, & He, 2013, http://www.nitrc.org/projects/bnv/). The MNI coordinates are listed in table 6.2.
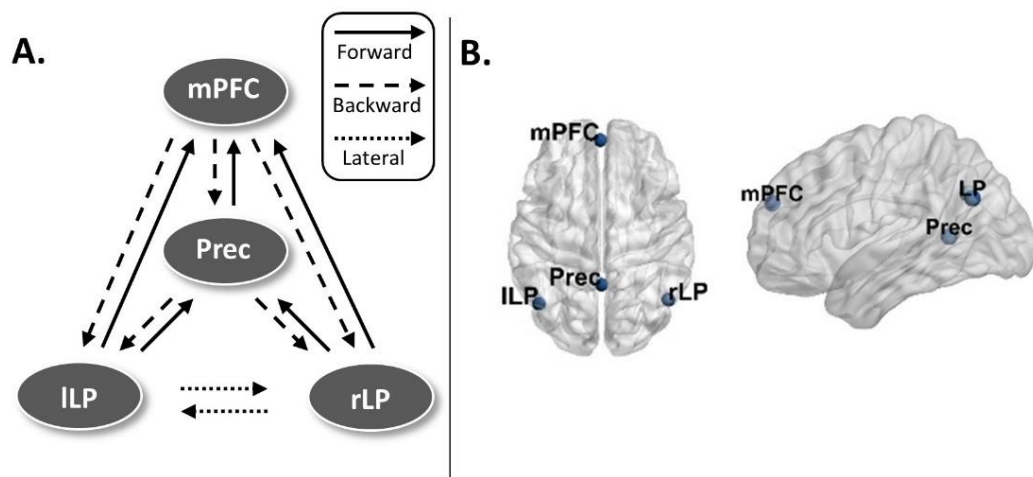
Figure 6.1A. The fully connected, schematic representation of the default mode network. **B.** The node locations for the DMN. mPFC – medial prefrontal cortex, Prec

– posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right

lateral parietal cortex.

As shown in figure 6.1A, the nodes in the DMN were connected via forward,

backward, and lateral connections as described in David et al. (2006, 2005).

Thus, each node was modelled as a point source with the neuronal activity being

controlled by operations following the Jansen-Rit model (Jansen & Rit, 1995).

These three different types of connections in each model were specified in what

is referred in the DCM literature as the 'A-matrix'.

Table 6.2. The default mode network nodes and their corresponding MNI

coordinates (adapted from Razi et al., 2017).

Network	Coordinates (in mm)		
	x	**y**	**z**
Default Mode Network			
Left lateral parietal	-46	-66	30
Right lateral parietal	49	-63	33
Posterior cingulate/precuneus	0	-52	7
Medial prefrontal	-1	54	27

6.2.6 Model Inversion

In DCM, model inversion refers to the process of fitting a model to explain

the empirical data of each participant's dataset, and thereby inferring a full

probability density over the possible values of model parameters (with the

expected values and covariance). The default parameter settings in DCM for CSD led to inaccurate fits of the model when inspected visually (see van Wijk et al., 2018, *p.* 824). To address this, similar to chapter 5, we doubled the number of maximum iterations to 256 and estimated the models with two adjustments to the hyperparameters: first, the shape of the neural innovations (i.e. baseline neural activity) were set to flat (-32) instead of the default white and pink (1/f) noise-component mixture (Moran et al., 2009). Second, we increased the noise precision value from 8 to 12 to bias the inversion process towards accuracy over complexity (see Friston et al., 2012 and Moran et al., 2009 for a detailed description of DCM for cross-spectral densities). With these adjustments, we estimated the full DMNs again, and applied the Bayesian Parameter Averaging (BPA) for each of the subject-groups separately, averaging over the posterior fits from the subjects for whom the model did converge satisfactory and setting these averaged posteriors as new priors for the respective non-converged subjects. Finally, we estimated all the full models again for all the subjects with setting the posteriors from the earlier subject model estimations as updated priors, but this time with the neural innovations and noise precision set back to default settings. To validate that the priors we used in the final inversion were suitable, we compared the group-level model evidence obtained from the BMA with and without the adjusted noise levels. With all comparisons, the default hyperparameter settings with the updated priors generated better model evidence (difference in free energies for control vs. UWS PET-, control vs. MCS+, and MCS+ vs. UWS PET- were +29364, +3096, and +39726, respectively. To qualitatively assess the model fits, the observed and model-

167

predicted cross-spectra were visually compared in each participant and judged sufficiently similar. This process led to satisfactory fits for the subjects (table 6.3). The average fits over subjects for each group are shown in figure 6.2.

Table 6.3. The number of satisfactory fits with the default hyperparameters and after adjusting the neural innovations and the noise precision for the different subject groups.

Patient group	N	Satisfactory fits	After BPA	Final
UWS	11	5	11	11
MCS+	12	9	12	12
Controls	11	9	11	11
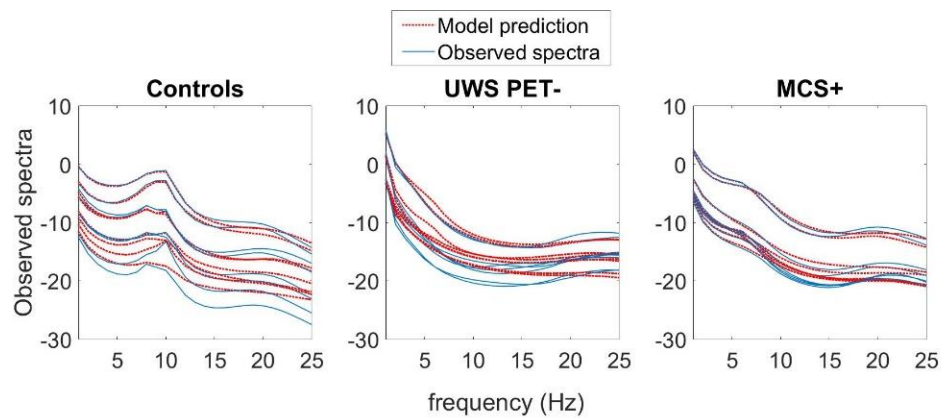
Figure 6.2. The average model fits across the participants in all subject-groups. **A-C.** Subject-averaged power spectra of the observed EEG channel-space data, juxtaposed with that predicted by the fitted DCM models of each subject group. Individual lines reflect spatial modes.

6.2.7 Parametric empirical Bayes

In DCM, the posterior density over the parameters given by the model inversion process is approximated via a variational Bayesian scheme by maximizing a lower bound (the negative free energy) on the log-evidence (Variational Laplace; Friston et al., 2007). A more recent addition, the Parametric Empirical Bayes (PEB) framework, can be utilised to infer, for example, the group-level commonalities and differences (Friston et al., 2016).

In PEB, the subject-specific parameters – here, the effective connectivity modulations between nodes in DMN – are taken to the group-level and modelled using a General Linear Model (GLM). In doing so, PEB partitions the between-subject variability into designed effects and unexplained random effects (captured by the covariance component). As a special case of Bayesian model selection (BMS), Bayesian model reduction (BMR) enables the inversion of multiple nested models from a single, fully connected ('full') model in a hierarchical manner. In doing so it enables a greedy search to compare the negative free energies for the nested models (reduced models), iteratively discarding the parameters that do not contribute to the free energy (originally 'post-hoc DCM analysis'; Friston & Penny, 2011; Rosa, Friston & Penny, 2012). Consequently, PEB conveys both the estimated group-level connection strengths and their respective uncertainty (posterior covariance component). As such, it is argued that hypotheses about commonalities and differences across subjects can be tested with more precise parameter estimates than with traditional frequentist comparisons (Friston et al., 2016).

A Bayesian Model Average (BMA) is calculated over the best 256 models weighted by their model evidence; for every connection, a posterior probability for the connection being present vs. absent is calculated by comparing evidence from all the models in which the parameter is switched on versus all the models in which it is switched off. Here, we applied a threshold of >.99 posterior probability, in other words, connections with over .99 posterior probability were retained. The overall process is shown in figure 6.3.
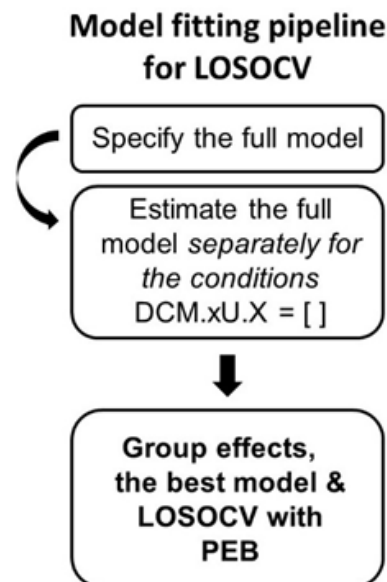
Figure 6.3. The pipeline for inverting the DCM model for different subject-groups. This was done to find the best models for each patient group, to estimate the effective connectivity modulations between the patient groups, and as a prerequisite for the LOSOCV classification with PEB modelling.

6.2.8 Leave-one-out cross-validation

Following the pipeline used in chapter 5, to validate our modelling framework, we investigated which DMN connections are predictive of the subject group by adapting a standard approach in computational statistics, leave-one-subject-out cross-validation (LOSOCV; spm_dcm_loo.m). Here, we iteratively fitted a multivariate linear model (as described in detail in Friston et al., 2016) to provide the posterior predictive density over connectivity changes, which was then used to evaluate the posterior belief of the explanatory variable for the left-out participant: in the present case, the probability of the subject group membership.

To cross-validate a fitted DCM model, one participant was left out each time *before* conducting PEB analysis on the training dataset, and the optimised empirical priors were then used to predict the subject-group to which the dataset from the left-out participant belonged (see Friston et al., 2016 for details). We repeated this procedure for each participant, and in doing so generated probabilities of state affiliation (here, posterior probabilities for subject group-membership).

It is worthwhile to note, that – excluding the specific alterations reported in section 6.2.6 – we have estimated the DCM models using the default parameter settings recommended in the literature (Ashburner et al., 2017; Friston et al., 2003; Friston et al., 2012; Kiebel, Garrido, Moran, Chen, & Friston, 2009). This is also true for the LOSOCV procedure, just as in chapter 5, no hyper parameter optimisation was done. Here, we trained the model with the data from all but the left-out participant (training set), and predicted the state based on the data from

the left-out participant (test set) and repeated this procedure leaving out a different participant each time.

6.2.9 Leave-one-subject-out cross-validation

We first estimated predictive performance in a leave-one-subject-out cross-validation paradigm in which LOSOCV metrics for all connections in the DMN and for a hypothesis-driven subsets were estimated (frontoparietal and parietal subsets; figure 6.4). Next, a data-driven approach was used in which we started the estimation from the connection associated with the largest connectivity reduction between the subject-groups and repeated the procedure for all connections. Here, we utilised a forward stepwise regression in which we started the estimation from the connections with the largest changes and continued through the parameters based on their respective modulation effect sizes. Lastly, we combined connections into data-driven subsets, starting from the connections with the best classification performance, until the classification accuracy stopped improving. The rationale was to investigate the location and direction of the most consistent inter-subject-level effects, in addition to the largest effect sizes identified by the PEB analysis.

6.2.10 Leave-one-state-out cross-validation

Finally, the validation process was generalised by introducing two more difficult classification problems: first, we trained the model on the DCM parameters from the control and the UWS PET- groups, and then tested it on

unseen data collected from the UWS PET+ patient-group. Second, we trained the

model on the data from the MCS+ and the UWS PET- groups, and again tested

on the UWS PET+ datasets. Here, the model was trained on all training

datasets. As above, the model used the optimised empirical priors to predict the

more likely patient-group the test dataset (UWS PET+) belonged. We

hypothesised that if our modelled effects are valid, and if the sustained PET-

metabolism reflects higher level of consciousness present in the

UWS PET+ patients in comparison to UWS PET- patients, in the former case the

model should classify the test datasets as controls rather than UWS PET-

. Similarly, in the latter case, given that the MCS+ patients are conscious, the test

data should be classified as MCS+ rather than UWS PET-. Here, we used

posterior probability for subject group-membership to quantify classification

performance.
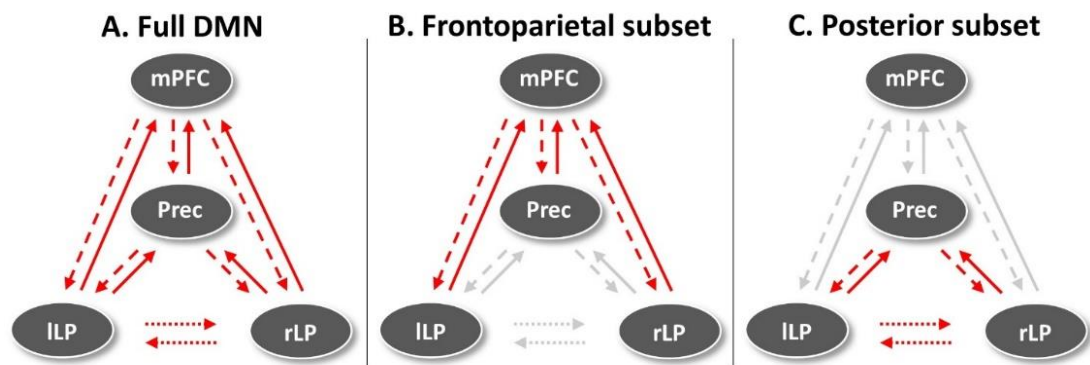
Figure 6.4. The hypothesis-driven subsets for the LOSOCV-paradigm. The red

arrows indicate the connections included in each subset. First, predictions based on

all connections were estimated (A). Next, predictions based on two connection

subsets – frontoparietal (B) and parietal subsets (C) – were estimated. Lastly, we

estimated predictions based on single connections in a data-driven approach. mPFC – medial prefrontal cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral parietal cortex.

6.3   Results

6.3.1   Dynamic causal modeling and parametric empirical Bayes

Our first goal was to investigate the effective connectivity modulations best explaining the difference between healthy controls, UWS PET-, and MCS+ patients. We modelled time-series recorded from the three groups with DCM for CSD at a single-subject level, followed by PEB at the group-level. In doing so, we estimated the change in effective connectivity in 12 inter-node connections in the DMN, contrasting 11 healthy controls both with 6 UWS PET- patients and with 12 MCS+ patients, and the MCS+ patients with UWS PET- patients.

Following the inversion of the between-groups PEB model, a greedy search was implemented to prune away connections not contributing significantly to the free energy using BMR. Figure 6.5 shows the most parsimonious models and figure 6.6 shows the estimated log scaling parameters contrasting healthy controls with UWS PET- (A), MCS+ with UWS PET- (B), and finally, healthy controls with MCS+ (C). Here, we applied a threshold of >.99 for the posterior probability; in other words, connections that were pruned by BMR and connections with lower than .99 posterior probability with their respective log scaling parameter are faded out (figures 6.6A, 6.6B, 6.6C).

174

On inverting the DMN for the control and UWS PET- groups, 3 connections were pruned away by BMR with additional 4 connections having lower than .99 posterior probability (figures 6.5A and 6.6A). All but one of the pruned connections were located within the posterior cortices between lateral parietal cortices and PCC/precuneus (with the exception of the right backward frontoparietal connection). The largest reduction in effective connectivity was located on left frontoparietal connection; the backward connection between mPFC and left lateral parietal node was largely diminished for the UWS PET- group in comparison to healthy controls.

On inverting the DMN contrasting MCS+ and UWS PET-, only three connections survived the BMR process with at least .99 posterior probability (with additional three connections surviving pruning with lower than .99 posterior probability; figures 6.5B and 6.6B). As with the control vs. UWS PET- contrast, the largest reduction was on the left backward connectivity from mPFC to lLP, with left lLP-mPFC forward connectivity increasing.

On inverting the DMN for the contrast between healthy controls and MSC+, two connections were pruned by the BMR with additional 4 connections having lower than .99 posterior probability for being present (figures 6.5C and 6.6C). The largest reductions were between the posterior nodes, to and from the lateral parietal cortices and PCC/precuneus. In addition, the left frontoparietal backward connectivity was reduced, although with smaller than .99 posterior probability and with clearly smaller effect size than with UWS PET-

175

. Other non-pruned connections were associated with small to medium
increases.
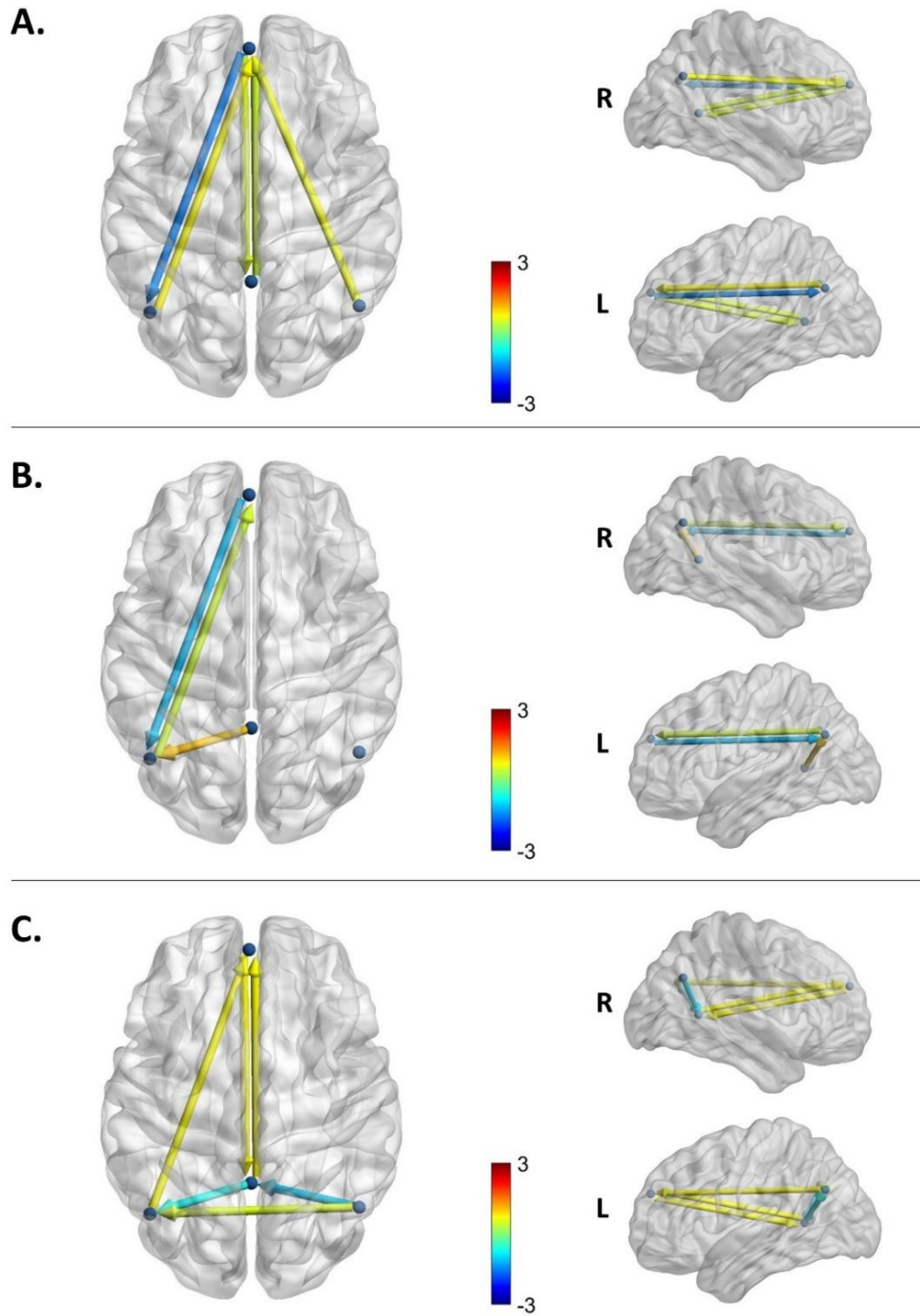
Figure 6.5. The most parsimonious DMN models after BMA and BMR for the difference between the healthy controls and the UWS PET-, MCS+ patients and UWS PET-, and healthy controls and MCS+. Colour shows modulation strength and direction. **A.** The most parsimonious model best explaining the difference between healthy controls and UWS PET- patients. Three connections were pruned with additional four having lower than .99 posterior probability of being present. All but one pruned connection was located between lateral parietal and PCC/precuneus nodes. The largest reduction in effective connectivity was in the backward connection from the medial prefrontal cortex to left lateral parietal cortex. **B.** The most parsimonious model best explaining the difference between the MCS+ and UWS PET- patients. Six connections were pruned by the BMR with additional three connections having lower than .99 posterior probability of being present. The largest reduction was on the backward connection from the medial prefrontal cortex to left lateral parietal cortex, similar to the reduction between healthy controls and UWS PET- patients. **C.** The most parsimonious model explaining the difference between healthy controls and MCS+ patients. Two connections were pruned with additional four connections having lower than .99 posterior probability of being present. The largest reductions on effective connectivity were on posterior connections between the lateral parietal cortices and PCC/precuneus.
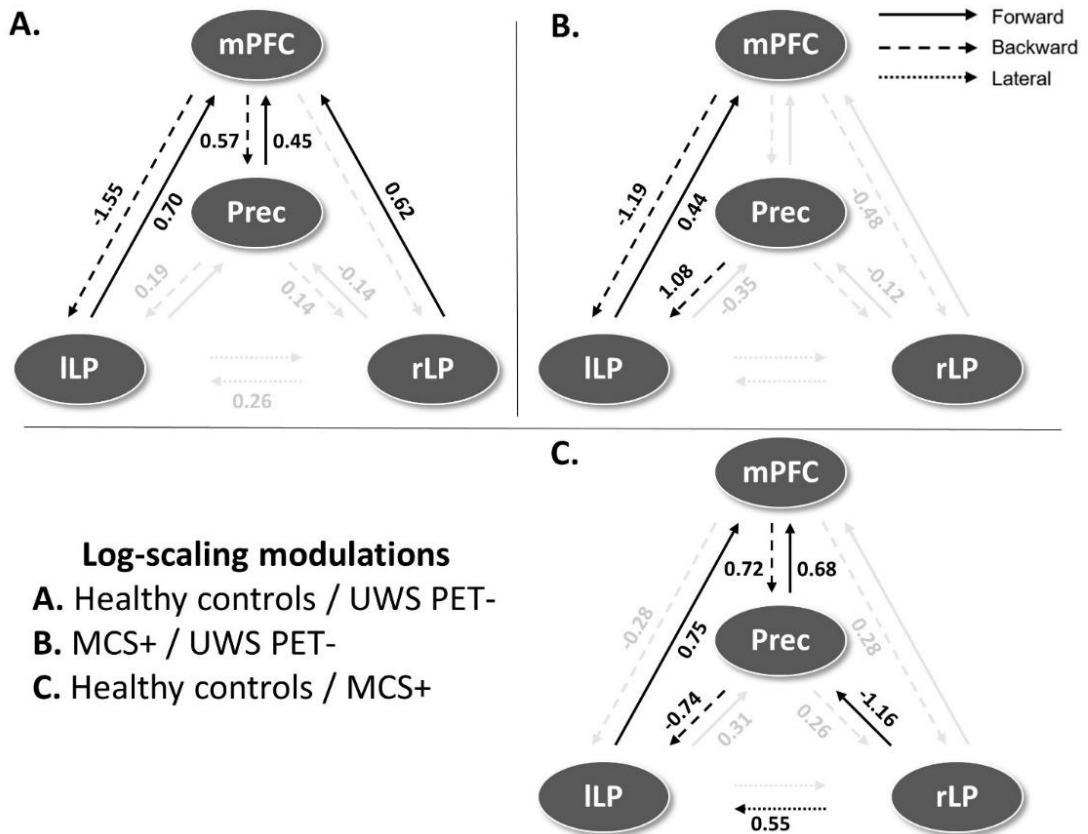
Figure 6.6. The log scaling parameters for the connections in the DMN after BMR and BMA. Positive values represent an increase and negative values a decrease in effective connectivity for the three group comparisons. Connections that were pruned by BMR and connections with lower than .99 posterior probability with their respective log scaling parameter are faded out. **A.** The modulatory effects best explaining the difference between healthy controls and UWS PET- patients. Connectivity between lateral parietal and PCC/precuneus nodes were either pruned away by BMR or had lower than .99 posterior probability with low modulatory effects. The largest reduction was found on backward lateral frontoparietal connection from medial prefrontal cortex to left lateral parietal cortex. **B.** The modulatory effects best explaining the difference between MCS+ and UWS PET-

. The modulatory effects on left bidirectional frontoparietal connections were both in the same direction as when comparing healthy controls to UWS PET-, with the largest reduction on left backward frontoparietal connection. In addition, right backward frontoparietal connectivity and posterior forward connectivity between lateral parietal nodes and PCC/precuneus reduced (smaller effect sizes), albeit with lower than .99 posterior probability of being present. **C.** The modulatory effects best explaining the difference between healthy controls and MCS+. The largest reductions were between the posterior nodes, between the lateral parietal nodes and PCC/precuneus. Bidirectional medial frontoparietal connectivity was increased in MCS+ in comparison to healthy controls. mPFC – medial prefrontal cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral parietal cortex.

6.3.2 Leave-one-subject-out cross-validation

To conduct LOSOCV, the DCM model was inverted again, this time separately for each patient group. Following the inversion process, PEB was conducted repeatedly on the training set in each cross-validation run alongside LOSOCV analysis to generate the posterior probabilities for group-membership (see Methods).

First, the UWS PET- patients were classified alongside the controls based on the full DMN model, and two hypothesis-driven connection subsets (frontoparietal- and parietal connections; figure 6.7). Similar approach was applied classifying MCS+ patients alongside UWS PET- patients, and finally, healthy controls alongside MCS+ patients. Figures 6.7 and 6.8 show violin plots

representing the individual posterior probabilities for the hypothesis-driven

classifications and data-driven approach, for all three contrasts, respectively (A:

control vs. UWS PET-, B: MCS+ vs. UWS PET-, C: control vs. MCS+). As

seen in figure 6.7, frontoparietal connections classified correctly most of the

controls and MCS+, and around half of the UWS patients in the controls vs.

UWS PET- and MCS+ vs. UWS PET- contrasts. Both, full DMN and parietal

subsets clustered most of the subjects around the chance level of 0.5.

We then moved to a data-driven approach in which we first predicted the

patient group membership based on the connections with the largest reductions in

PEB, one at a time, working through all connections. Lastly, we checked

combinations based on the connections' respective classification accuracies (see

Methods). The bi-directional left frontoparietal connectivity produced the best

predictions, especially when classifying the UWS PET- from both,

healthy controls and the MCS+ patients (figures 6.8A and 6.8B), with the best

predictions based on the backward mPFC-lLP connectivity. Forward lLP-

mPFC connectivity produced the best predictions for controls vs. MCS+ contrast,

especially for the healthy controls (6.8C). None of the tested combinations

improved classification performance.
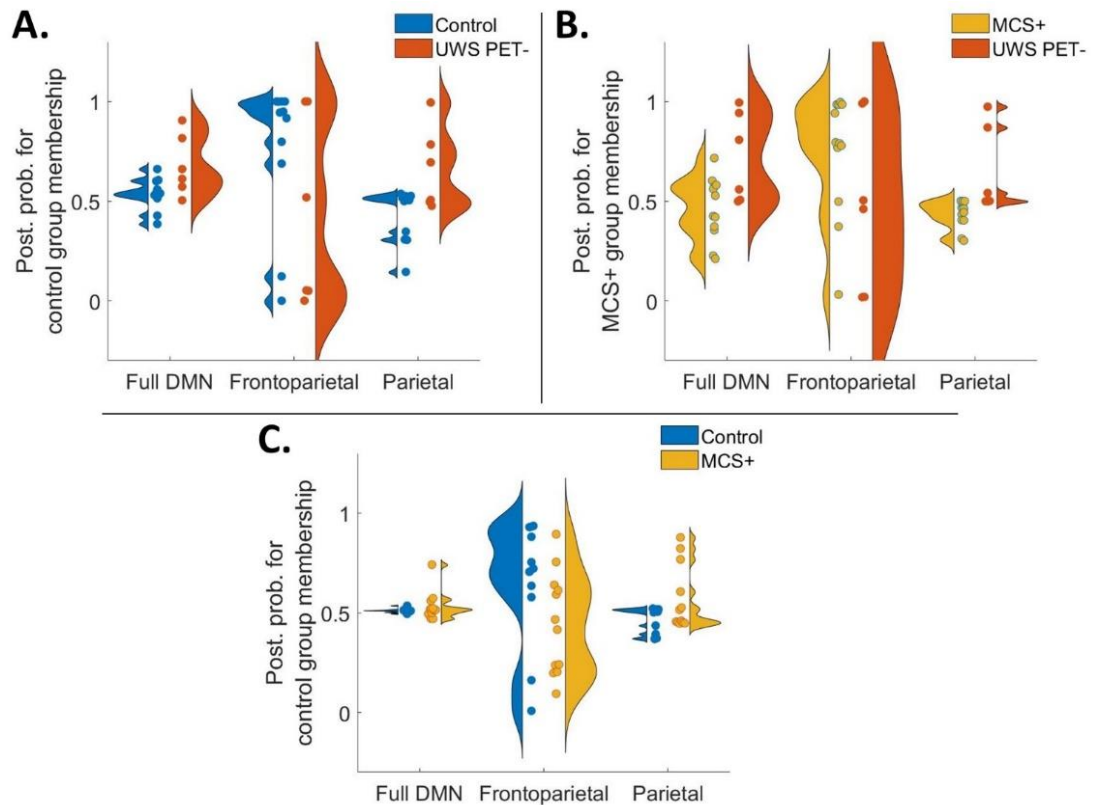
Figure 6.7. Violin plot representing diversity in posterior probabilities for control group membership (A and C) and for MCS+ group membership (B) for the hypothesis-driven subsets for all three group contrasts. Each coloured point specifies a subject. In a perfect model in panels A and B, the UWS PET- patients ($N = 6$), and in panel C, the MCS+ patients ($N = 12$) should approach to a posterior probability of zero. Overall, the results show a trend for frontoparietal connections producing the best predictions. **A.** When classifying UWS PET- patients alongside healthy controls, the frontoparietal subset produced the best results. The individual data points reveal more consistent classifications of healthy controls. On all three panels, full DMN model and parietal subset produced classifications with most posterior probabilities bordering the 0.5 chance level. **B.** As in panel A, the best predictions

when classifying UWS PET- patients alongside MCS+ were based on frontoparietal connections, specifically with MCS+ patients. **C.** Classification of MCS+ alongside healthy controls. Frontoparietal subset produced the best predictions, however with large variability on the performance across the subjects.
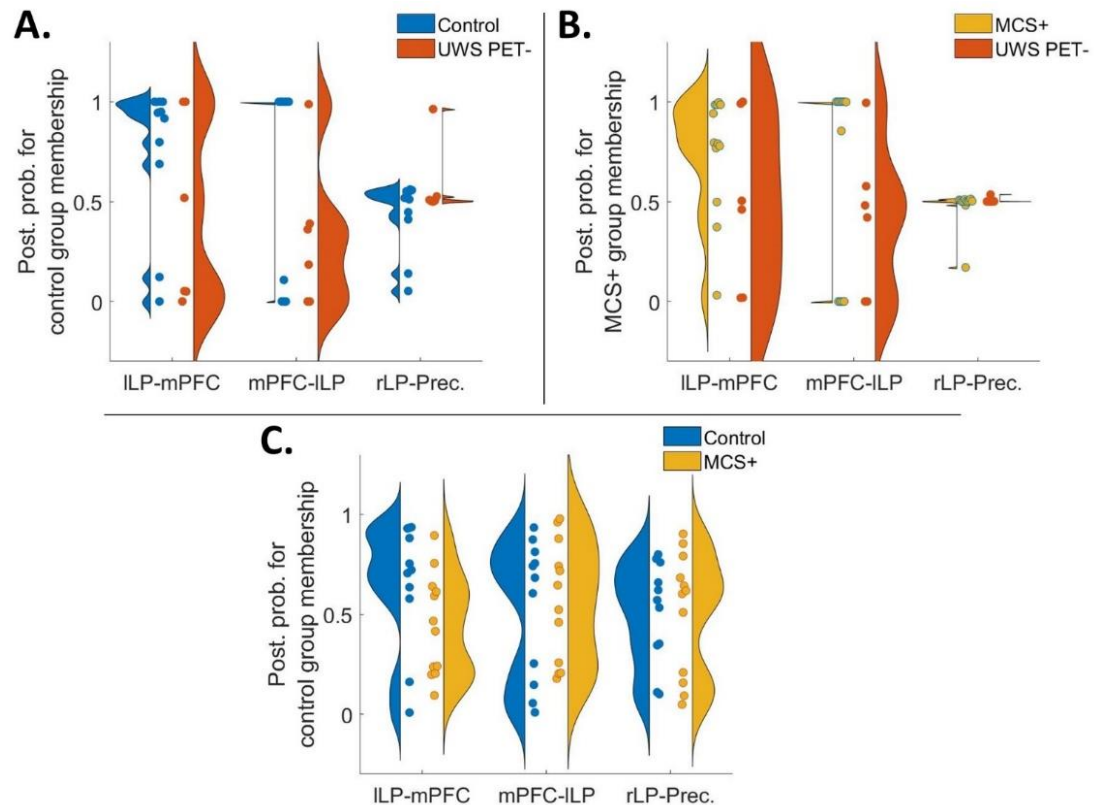
Figure 6.8. Violin plot representing diversity in posterior probabilities for control group membership (A and C) and for MCS+ group membership (B) for the data-driven connections for all three contrasts. Each coloured point specifies a subject. In a perfect model in panels A and B, the UWS PET- patients, and in panel C, the MCS+ patients should approach to a posterior probability of zero. Overall, the best predictive performance was based on the left bi-directional frontoparietal

connections when classifying UWS PET- alongside controls (A) and MCS+ (B). Largest inconsistencies and variability were on classifications of MCS+ alongside healthy controls. **A.** Left frontoparietal connectivity from mPFC to lLP produced the best predictions (mean posterior probabilities) of the group-membership when classifying UWS PET- alongside healthy controls. As with the hypothesis-driven subsets, the classifications were more accurate with healthy controls than with patients. **B.** As in panel A with healthy controls and UWS PET- patients, the classification performances based on mPFC-lLP and lLP-mPFC produced the most consistent results when contrasting UWS PET- patients alongside MCS+. **C.** Mean posterior probabilities for classification of MCS+ alongside healthy controls. The performance of the models based on the single connections did not produce consistently accurate classifications. mPFC – medial prefrontal cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral parietal cortex.

Figures 6.9 and 6.10 show confusion matrices of prediction accuracy calculated by labelling posterior probabilities greater than 0.5 as a positive classification, for the hypothesis-driven subsets and the data-driven approach, respectively. In figure 6.9, the frontoparietal subset performed most consistently in terms of both, classification accuracy and mean posterior probability, especially with healthy controls.

In figure 6.10 with the data-driven approach, the frontoparietal backward connection from mPFC to lLP performed best in terms of both, classification accuracy and mean posterior probability for healthy control vs. UWS PET- and

183

MCS+ vs. UWS PET- contrasts. Forward frontoparietal connectivity from lLP to mPFC classified healthy controls and MCS+ patients from UWS PET- with high accuracy but bordered the chance level with UWS PET-. Overall, the bi-directional left frontoparietal connections provided the best classification performances amongst all connections and combinations tested.
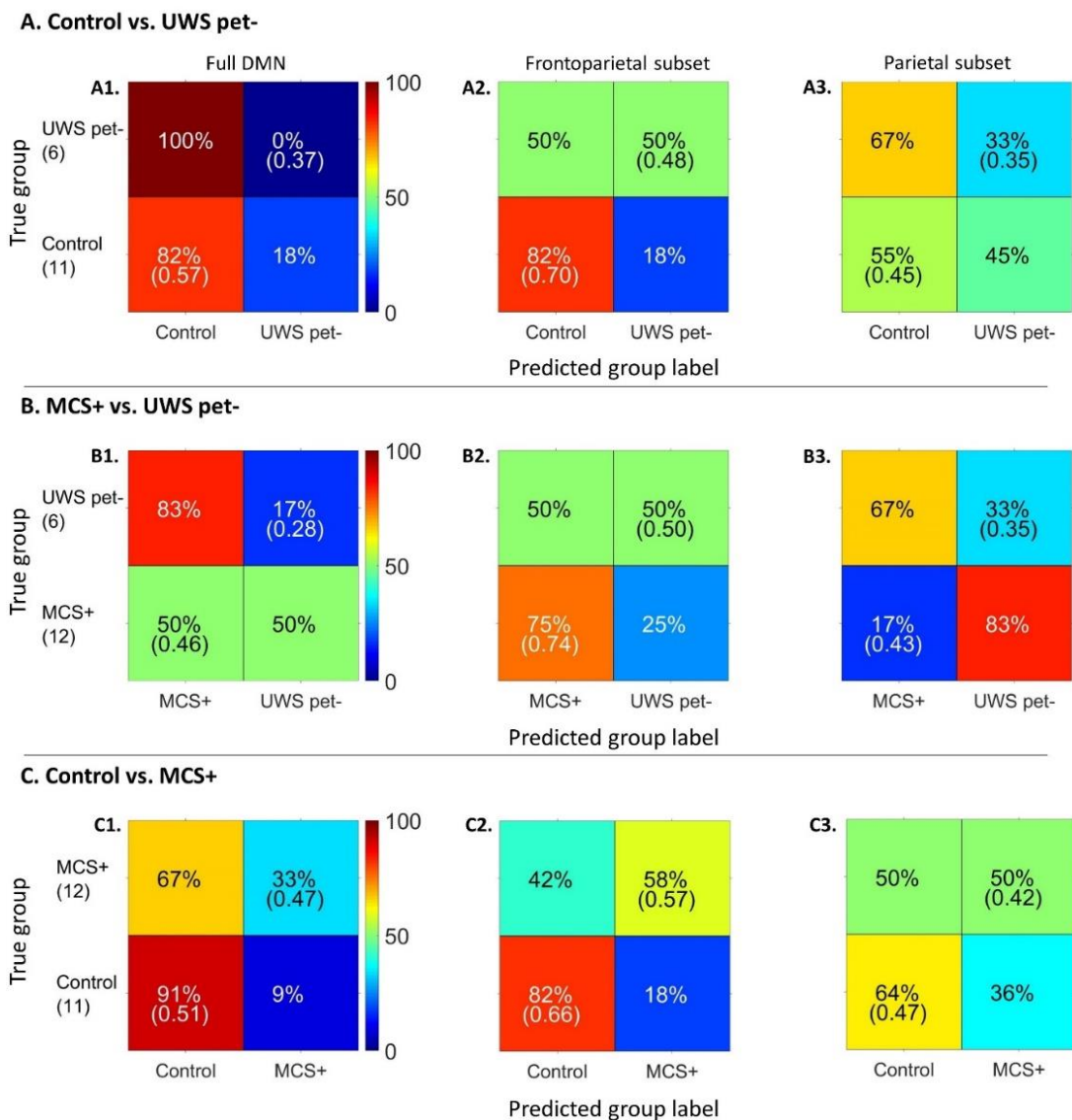
Figure 6.9. Classification accuracy percentage (mean posterior probability for correct classification) in the leave-one-subject-out cross-validation paradigm for the hypothesis-driven subsets. The number of subjects in each group is shown in parenthesis under the true group labels. The frontoparietal subset performed the best in terms of both classification accuracy and mean posterior probability, especially with healthy controls for healthy control vs. UWS PET- and MCS+ vs. UWS PET-contrasts (panels A2 and B2, respectively). Classification based on full DMN had high accuracy for healthy controls; however, the mean posterior probabilities bordered the chance level.

185
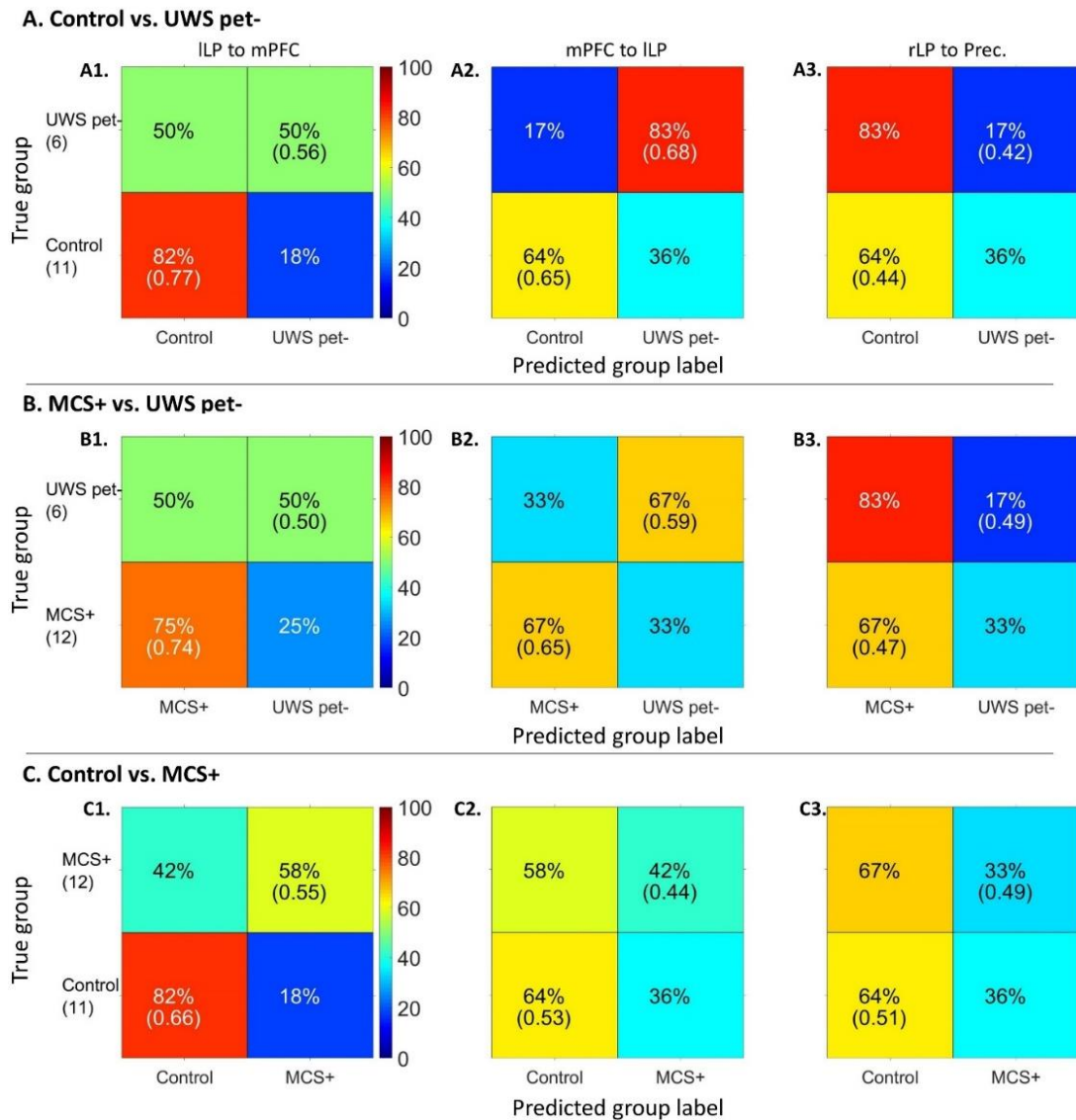
Figure 6.10. Classification accuracy percentage (mean posterior probability for correct classification) in the leave-one-subject-out cross-validation paradigm for the data-driven approach. The number of subjects in each group is shown in parenthesis under the true group labels. For the healthy controls vs. UWS PET- and MCS+ vs. UWS PET- contrasts, the frontoparietal backward connection from mPFC to lLP performed best in terms of both, classification accuracy and mean posterior probability. Forward frontoparietal connectivity from lLP to mPFC classified healthy controls and MCS+ patients from UWS PET- with high accuracy but bordered the

chance level with UWS PET-. Similarly, lLP to mPFC connectivity performed the best with the healthy controls vs. MCS+ contrast.

6.3.3   Leave-one-state-out cross-validation

Finally, the predictive power of DCM modelling was generalised in two more difficult classification problems; each model was trained first on healthy controls and UWS PET- and then tested on the previously unseen UWS PET+ group. A similar approach was used with a training set consisting of MCS+ and UWS PET- patients. The individual posterior probabilities for the five UWS PET+ patients represented in a violin plot for both, the hypothesis subsets (panels A and B; controls vs. UWS PET- and controls vs. MCS+, respectively) and for data-driven connections (panels C and D) are shown in figure 6.11. The hypothesis-driven subsets did not classify the UWS PET+ as controls or MCS+. Instead, when trained on datasets from healthy controls and UWS PET-, the frontoparietal subset classified four out of five UWS PET+ patients as UWS PET-.

With the data-driven approach, the left backward frontoparietal connectivity from mPFC to lLP produced nearly perfect predictions classifying the UWS PET+ datasets as controls (6.11C) and MCS+ (6.11D) rather than as UWS PET- group. Similar as with frontoparietal subset (6.11A), the left forward connectivity from lLP to mPFC classified four of five patients as UWS PET- patients rather than healthy controls – three of them with a high posterior probability. This dissociation was not as prominent when training the model with MCS+ and UWS PET- patients; the backward connectivity produced nearly perfect classifications

187

connectivity showed larger variability.
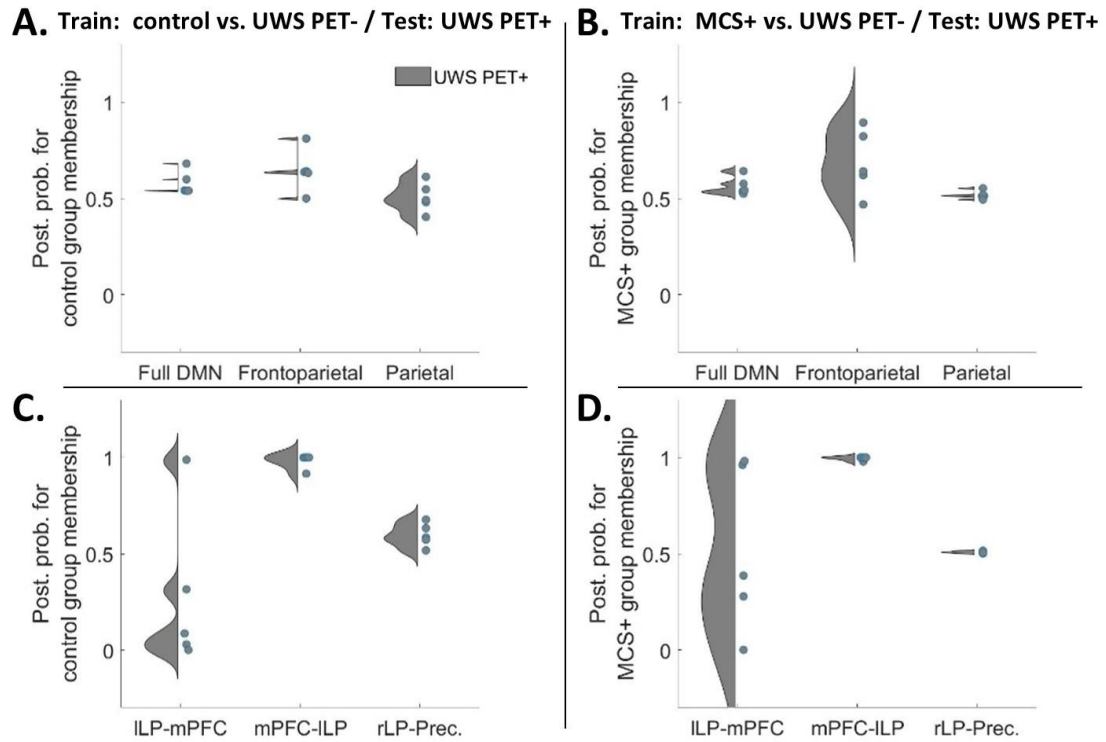
Figure 6.11. Violin plot representing diversity in posterior probabilities for control

group membership (A and C) and for MCS+ group membership (B and D) for both,

the hypothesis- and data-driven predictions. Here, the models were trained on

datasets from controls and UWS PET- (A & C) or from MCS+ and UWS PET- (B

and D) and tested on unseen data from UWS PET+ patients ($N = 5$). Each coloured

point specifies a subject. Overall, the left backward frontoparietal connectivity

produced the best group membership predictions. **A and B.** Mean posterior

probabilities for classification of UWS PET+ patients alongside healthy controls (A)

and MCS+ patients (B). Neither of the hypothesis-driven subsets, nor the full DMN,

clearly classified the unseen UWS PET+ patients as members of either train-group. **C**

and D. Left frontoparietal connectivity from mPFC to lLP produced almost perfect predictions for the UWS PET+, classifying all five patients as either controls or MCS+ rather than UWS PET- patients. Unlike the backward connectivity, predictions based on the left forward connectivity from lLP to mPFC, the model classified four of five UWS PET+ patients as UWS PET- rather than as controls (C). Similarly, when trained on MCS+ and UWS PET-, the model classified three of five patients as UWS PET- rather than MCS+. mPFC – medial prefrontal cortex, Prec – posterior cingulate cortex/precuneus, lLP – left lateral parietal cortex, rLP – right lateral parietal cortex.

6.4 Discussion

In this cross-sectional, retrospective analysis, we applied spectral DCM to high-density EEG data with parametric empirical Bayes to investigate the difference in effective connectivity dynamics between cortico-cortical regions of the default mode network in DoC patients (UWS and MCS+) and healthy controls. Overall, the modelling results indicate a key difference between healthy controls or conscious patients and completely unresponsive patients in left-hemispheric backward frontoparietal connectivity. Furthermore, with out-of-sample cross-validation, we demonstrated that this association is robust enough to not only distinguish patient groups from each other, but also generalises to an unseen data subset, collected from seemingly unresponsive patients showing evidence of hidden consciousness which became apparent when assessed with functional neuroimaging (UWS PET+). These results identify specific alterations

in the DMN after severe injury and highlight the clinical utility of EEG-based measurement of effective connectivity for identifying covert consciousness.

6.4.1 Dynamic causal modelling

The most parsimonious model explaining the difference between healthy controls and completely unresponsive patients (UWS PET-) indicated a large relative reduction in left-hemisphere backward frontoparietal connectivity in UWS PET- patients. Additionally, a small, lower-probability reduction from right lateral parietal cortex to precuneus was found. Interestingly, excluding the right parietal connection, the estimated connectivity in the posterior nodes – within the 'posterior hot zone of conscious contents' (Koch et al., 2016a; Siclari et al., 2017) – was either pruned away from the model best explaining the difference or returned only small, low-probability increases suggesting lower relative importance for the posterior hot zone in explaining the difference between healthy controls and UWS PET- patients.[4] In contrast, in a previous fMRI DCM study with DoC patients precuneus/PCC-related connectivity reduction was the key difference; specifically, the recurrent connectivity (down-regulation of the PCC itself) was found to be diminished on UWS patients in comparison to both, MCS patients and healthy controls (Crone et al., 2015).

---

[4] Connections pruned by BMR are considered not to contribute towards the model evidence. See Zeidman et al. (2019).

However, not only are the data between Crone et al. (2015) and the present

study from different modalities, and thus, direct comparisons of the results

unsound, the underlying neurobiologically motivated models used by DCM are

different for hemodynamic versus electrophysiological data leading to a different

interpretation of the modulatory effects.[5] Here, we modelled the data with the

default ERP neuronal model (David et al., 2005) primarily in order to produce

comparable results with prior DCM for EEG work modelling consciousness

(Boly et al., 2011, 2012; Ihalainen et al., 2021). Further, as in our previous work

with propofol-anaesthesia, we aimed to model disrupted consciousness at the

level of active networks rather than focusing on e.g. synaptic hypotheses or

recurrent self-connections (intrinsic connectivity), which could be better captured

with other neuronal models such as the LFP model (Moran et al., 2007) or the

Canonical Microcircuits model (Bastos et al., 2012; see also Moran et al., 2013

for a review of neuronal population models). Hence, we only estimated extrinsic

connectivity – i.e. connectivity between cortical areas. It is possible that the

observed differences in the network dynamics are driven by modulations in self-

---

[5] The interpretation of the modulatory effects in DCM for fMRI versus EEG differ in that positive and

negative values indicate excitatory and inhibitory effects in fMRI data (except for recurrent

connections, for which the connection strength is always negative and modulations reflect increases or

decreases in comparison to the prior). In DCM for EEG, positive modulations indicate an increase and

negative a decrease in connectivity relative to the prior. Generally speaking, the backward connections

are thought to have more inhibitory and largely modulatory effect in the nodes they target (top-down

connections), while forward connections are viewed as having a strong driving effect (bottom-up;

Salin & Bullier, 1995; Sherman & Guillery, 1998).

inhibiting, recurrent connectivity within the cortical sources or within and between subcortical networks driving the disruptions in the DMN (Chen et al., 2018; Coulborn, Taylor, Naci, Owen, & Fernández-Espejo, 2021). A worthwhile endeavour for future DCM for EEG studies would be to model the extent to which recurrent, within-source cortical connectivity may or may not drive the modulations in extrinsic connectivity.

As far as we are aware, just one study has used DCM for EEG in DoC populations. Boly et al. (2011) showed that in an auditory mismatch negativity paradigm, the difference between UWS, MCS, and healthy controls was an impairment of backward connectivity from frontal to temporal cortices in the UWS patients, emphasising the importance of top-down processing for conscious perception. Similarly, in the present resting-paradigm, the key difference distinguishing UWS PET- from both, MCS+ and healthy controls was decreased left-lateralised backward frontoparietal connectivity in UWS PET- patients, although from medial prefrontal cortex to lateral parietal cortex and not to superior parietal cortex. It is important to note, however, that the differences in the paradigm, methodology and in the models estimated render rigorous, direct comparisons of the results between Boly et al. (2011) and the present study infeasible. Despite the methodological differences, the results of both studies suggest a crucial role for lateral backward connectivity originating from the frontal cortex. Future studies should investigate this further by modelling the connectivity related to temporal areas, and the backward frontoposterior connectivity in other resting networks (see below) in DoC patients.

192

Like with healthy controls vs. UWS PET- comparison, the largest difference between MCS+ and UWS PET- patients was left frontoparietal backward connectivity, with UWS patients again showing reduced connectivity. Furthermore, the left forward parietofrontal connectivity and backward connectivity from precuneus to lLP both increased, reproducing the modulations between healthy controls and UWS PET-. These changes were accompanied by smaller, low-probability reductions in forward connectivity within the posterior hot zone and in right backward frontoparietal connectivity. These modelling results highlight the left frontoparietal backward connectivity as the key distinguishing difference when comparing healthy controls or conscious patients with completely unconscious patients and complement those of previous studies discriminating DoC patients from scalp-level EEG connectivity; especially frontal and parietal functional connectivity has been shown to consistently discriminate DoC patients (Chennu et al., 2014; 2017). As the direction and the spatial location of the changes in connectivity were similar with the comparisons involving UWS PET- patients, and distinguishable from those when comparing healthy controls vs. MCS+, we were motivated to further test the predictive power of the modulatory effects (see below).

In contrast, the largest connectivity reductions between healthy controls and MCS+, although relatively smaller than in previous contrasts, were associated with the precuneus node in the posterior hot zone. The left backward frontoparietal connectivity was again reduced, but by a smaller effect and with lower than .99 probability of being present in the most parsimonious model. The activity changes in the posterior hot zone of conscious content have been

193

associated with changes in consciousness not only during sleep (Lee et al., 2019; Siclari et al., 2017) and general anaesthesia (Alkire, Hudetz, & Tononi, 2008; Ihalainen et al., 2021), but also in DoC patients (Vanhaudenhuyse et al., 2010; Wu et al., 2015). Moreover, previous studies have suggested a subdivision of the frontoparietal network into two anticorrelated subnetworks; an "intrinsic" network encompassing precuneus/PCC, anterior cingulate/mesofrontal cortices, and parahippocampal areas associated with internal awareness, and into an "extrinsic" central executive network (CEN) encompassing dorsolateral prefrontal and posterior parietal areas linked with the intensity of external awareness (Boveroux et al., 2010; Vanhaudenhuyse et al., 2010). The observed decrease in the left lateral frontoparietal connectivity in the present study between UWS PET- patients and healthy controls or MCS+ patients may reflect such diminished internal awareness in the UWS PET- patients. To that end, future endeavours should investigate the modulatory effects and the possible predictive power of such modulations in other resting state networks, such as the CEN.

However, changes in the physiological state of the frontoparietal network alter not only consciousness but also several other brain functions such as vigilance and attention (Hohwy, 2009; Koch et al., 2016a). Moreover, the specific areas of the DMN have been associated with specific cognitive functions; for example, the frontal areas seem to be important for self-reference, whereas the precuneus/PCC in autobiographical memory (Whitfield-Gabrieli et al., 2011). It remains a possibility that the found modulations in the DMN reflect changes also in other cognitive functions, rather than in consciousness alone.

194

It is important to bear in mind that DoC patients typically suffer from widespread structural brain damage often accompanied by distributed white matter anomalies (Annen et al., 2018; Fernández-Espejo et al., 2012; Tshibanda et al., 2009). Hence, it is important to consider the feasibility and validity of applying DCM to DoC patients; this is particularly true for DCM for EEG which requires the specification of the anatomical locations of the nodes/sources a priori, and with patients with non-traumatic aetiology, e.g. patients with anoxic brain damage (see King et al., 2011 and Boly et al., 2011). Here, we mitigated these concerns first by limiting our modelling to DMN, a resting state network previously associated with disorders of consciousness in patients (Boly et al., 2008, 2009; Crone et al., 2011, 2015; Fernández-Espejo et al., 2012; Heine et al., 2012; Laureys et al., 1999; Laureys, 2005; Vanhaudenhuyse et al., 2010) and with changes in the conscious state e.g. due to anaesthesia (Boveroux et al., 2010; Greicius et al., 2008; Stamatakis, Adapa, Absalom, & Menon, 2010) and sleep (Horovitz et al., 2009).

Second, we selected only patients with traumatic brain injury; traumatic aetiology, as compared to non-traumatic aetiology, has been associated with more focal injury centred often on areas susceptible to rotational forces, such as the brainstem, midbrain, thalamus, hypothalamus, cerebellum, and posterior corpus callosum (Guldenmund, Soddu, et al., 2016; Newcombe et al., 2010). Third, we applied a special case of Bayesian model selection (BMS), Bayesian model reduction (BMR) to invert multiple nested models from a single, fully connected DMN (see Methods). A particular advantage here is that BMR can be applied using an explanatory approach, in which no strong a priori hypotheses

about the model parameters are needed. This enables a greedy search to compare the negative free energies of the reduced (nested) models by iteratively discarding parameters that do not contribute to the free energy. The procedure stops when discarding any parameters starts to decrease the negative free energy, returning the model that most effectively trades-off goodness of fit and model complexity in explaining the data. BMR applied in this way allows one to estimate a large model space from a single, specified full model in a relatively short period of time (Friston & Penny, 2011; Rosa et al., 2012; Zeidman et al., 2019).

Nevertheless, it is possible that not all true influences on the specific regions are captured by the specified full model. Moreover, the explanatory approach to BMR is conducted under the assumption that all reduced models are equally probable a priori, and thus, the full model should only contain parameters that are biologically plausible. Here, we cannot exclude physical damage to cortical areas and pathways crucial to the functioning of the DMN.

That said, our aim was not to make any strong claims about the "true" model; to draw stronger conclusions about the "true" underlying neuronal basis using DCM for EEG, structural MRI imaging assessing the extent of the damage in specific patients, possibly in adjunct with source localisation of the EEG signals, should be applied. Here, the aim was to demonstrate and to compare the predictive performance of effective connectivity in the clinical context. Additionally, demonstrating predictive value with significant generalisation

performance with cross-validation, the level of confidence we can ascribe to our modelling results increases.

6.4.2 Leave-one-subject/state-out cross-validation (LOSOCV)

To test whether the effective connectivity modulations were consistent enough across the patient groups and healthy controls to reliably distinguish the groups from each other, we first conducted a leave-one-subject-out cross-validation based on hypothesis- (full DMN, frontoparietal, and parietal subsets; figure 6.7) and data-driven subsets of connections (figure 6.8). Amongst the connection subsets, the frontoparietal connections performed the best, classifying a majority of controls and MCS+, and around half of the UWS patients correctly in the healthy controls vs. UWS PET- and in MCS+ vs. UWS PET- contrasts. The full DMN and parietal subsets clustered most of the subjects around the chance level in all three contrasts.

We then moved to a data-driven approach in which we first predicted the patient group membership based on the connection with the largest reduction in PEB, one at a time, working through all connections. It is important to note that searching for the best connection in this way increases the risk of overfitting the model by potentially extracting some of the residual variation – noise – as if representing the underlying model structure. However, the congruence between the present results and those discussed in chapters 4 and 5, and the fact that the best model generalised to an unseen dataset suggest that the results may reflect a genuine effect (see below).

197

With the data-driven approach, the bi-directional left frontoparietal connectivity produced the best predictions, especially when classifying the UWS PET- from both, healthy controls and the MCS+ patients. The single best performing connection was the backward frontoparietal connectivity (figure 6.8). Not surprisingly, the classifications were more accurate and consistent with healthy controls than with patients; classifications of patients suffering from brain damage, and thus, from disrupted brain functioning, were expected to vary more than those of healthy controls. Next, we combined the single connections into data-driven subsets based on the classification performance: none of the combinations improved the performance of the single connections. Last, the predictive power of DCM modelling was generalised in two more difficult classification problems; following the hypothesis- and data-driven approaches above, we trained each model on healthy controls or MCS+ on the one hand, and UWS PET- patients on the other, and then tested the models on the previously unseen, "covertly aware" UWS PET+ patients. The hypothesis-driven subsets did not classify the UWS PET+ patients as controls or MCS+. Crucially, with the data-driven approach, the left backward frontoparietal connection produced nearly perfect predictions classifying all five patients as either controls or MCS+ rather than UWS PET- patients.

While the results of the cross-validation should be interpreted with caution due to the relatively low number of subjects in our study, the results highlight the importance of the frontoparietal connectivity – particularly the left-lateralised backward connectivity – when predicting individual states of consciousness. Although direct comparison between the present results and the predictions made

in chapter 4 with DoC patients is difficult due to the methodological differences (e.g. EEG signal at the level of electrodes in chp. 4 vs. source level signal, functional vs. effective connectivity, UWS PET+ not distinguished from UWS PET- in chp. 4), we highlight that the present results align with and expand those presented in the previous chapter. Specifically, in chapter 4 we established predictive value of the left-hemisphere functional connectivity in DoC patients; the left-lateralised alpha-band connectivity – but not the right – significantly distinguished not only the UWS patients from MCS patients and healthy controls, but also implied a distinction between MCS- and MCS+ patients (as calculated with debiased weighted phase-lag index; see chapter 4). Here, using effective connectivity, we showed complementary predictions distinguishing DoC groups from controls and from each other at the level of the brain sources in the left hemisphere.

Similarly, the cross-validation results in chapter 5 highlighted the importance of frontoparietal connections in distinguishing loss of consciousness due to propofol-anaesthesia from behavioural responsiveness. As in the present chapter, the best prediction performance in anaesthetised participants was associated with frontoparietal connectivity within the DMN. The modelling results from PEB, however, indicated higher relative importance for the posterior hot zone in anaesthetised participants in comparison to the PEB results in the present chapter.

The importance of frontoparietal connections within the DMN for dissociating states of consciousness in DoC patients is perhaps not surprising given the previously established association between conscious awareness and

the DMN (e.g. Boly et al., 2008, 2009; Boveroux et al., 2010; Vanhaudenhuyse et al., 2010). More specifically, consciousness is thought to require brain-wide broadcasting of information by a "global workspace" associated with brain areas within the frontoparietal network (Baars, 1988; 1997; Dehaene et al., 2011). It is important to note, however, that the global neuronal workspace theory (GNW) is not a localisationist approach but rather posits a distributed "router" for conscious access (ibid.). The extensive and rapid bidirectional connectivity between the hubs of the GNW is thought to trigger the sudden collective and coordinated activity mediating global broadcasting (Mashour et al., 2020). Aptly, these hubs initially included the prefrontal cortex and parietal cortex (in combination with a set of specialized and modular perceptual, motor, memory, evaluative, and attentional processors) although it has later been complemented with other, potentially equally important hubs (such as the anterior and posterior cingulate and the precuneus). The observation that the changes in the long-range frontoparietal connectivity best predicts the state of consciousness is in accordance with the suggested importance of the connectivity between the hubs in the GNW. This is in contrast with the more restricted, content-specific neural correlates of consciousness often associated with the posterior hot zone (Koch et al., 2016a).

However, it is important to keep in mind that presumably, when the patient becomes "more" conscious, different content becomes more globally available for conscious processing throughout the brain, affecting and employing different cognitive systems (Hohwy, 2009). In other words, any major changes in the physiological state alter not only consciousness but other cognitive systems as

well, many of which depend on levels of arousal-promoting neuromodulators (ibid.). Therefore, it remains possible that the predictive performance of the frontoparietal effective connectivity is related not only to the state of consciousness, but also to other arousal-related cognitive processes.

In summary, our results indicate a key difference between healthy controls or conscious patients and completely unresponsive patients in left-lateralised backward frontoparietal connectivity. With out-of-sample cross-validation, we demonstrated that this association is robust enough to not only distinguish patient groups from each other, but also generalises to an unseen data subset, collected from seemingly unresponsive patients. These results contribute towards identifying specific alterations in network interaction after severe brain injury, and importantly, suggest clinical utility of EEG-based effective connectivity in identifying covertly aware patients who seem behaviourally unresponsive.

7. General discussion

Over the previous chapters, we have described research exploring a key challenge in neuroscience: how does the brain sustain and lose consciousness, and further, how are the brain networks affected during unconsciousness? In recent years, it has been assumed that prefrontal cortex is essential for conscious experience, either alone (Del Cul et al., 2009) or in conjunction with parietal areas (Bor & Seth, 2012; Laureys & Schiff, 2012). More recently, this assumption has been called into question, and temporo-parieto-occipital areas – colloquially named as the posterior hot zone – have been suggested to primarily mediate conscious experience. Under this view, frontal activation is associated primarily with cognitive processes subsequent to conscious experience, such as attention and working memory (Boly et al., 2017; Koch et al., 2016a). Here, in chapters 3 and 5, we provided novel evidence supporting the importance of the posterior areas; we contrasted unconsciousness due to propofol anaesthesia with normal wakefulness and characterised the differences with functional and effective connectivity. We found the posterior electrodes (at the scalp level) to best predict various indirect measures of consciousness. In the dynamic causal models (DCM) best explaining the difference between normal wakefulness and unconsciousness, most reductions were observed within the posterior hot zone, between posterior neuronal sources. This was true particularly for the default mode network, and for the large model combining three key resting state networks.

In parallel to theoretical considerations of how consciousness is sustained or lost and what are the full neural correlates of consciousness (NCC), an equally important approach concerns the practicality of the current proxies for consciousness – this is particularly important in the clinical context. Assessing the depth of anaesthesia or consciousness in brain-damaged patients robustly is critical as this may have far-reaching consequences in each individual case, for example, when making decisions regarding life-sustaining treatment (Buckley et al., 2004; Demertzi et al., 2013; Kuehlmeyer et al., 2014) and in diagnostic precision (Peterson, Cruse, Naci, Weijer, & Owen, 2015). Previous research has suggested that states such as disorders of consciousness (DoC) and anaesthetic-induced loss of consciousness (LOC) can be distinguished with neuroimaging (e.g. Bonhomme et al., 2019; Crone et al., 2014; Guldenmund et al., 2016; Laureys et al., 1999). Moreover, these states can be distinguished and predicted from EEG (Chennu et al., 2016; Deng et al., 2019; Engemann et al., 2018; Sitt et al., 2014; Zhang et al., 2020).

Hence, in chapter 5, we adapted a standard approach in computational statistics, leave-one-subject-out cross-validation (LOSOCV), to first, validate our effective connectivity modelling work, and second, to produce novel effective connectivity-based predictions of propofol-induced LOC. We observed that the frontoparietal effective connectivity best predicted the state of consciousness for the left-out participants. We further generalised this finding to unseen datasets consisting of recordings captured during the recovery phase. The fact that we observed the most consistent modulations with large enough effect sizes to

predict the state of consciousness in frontoparietal effective connectivity supports a role for frontal areas in maintaining wakeful conscious processing.

Further evidence supporting a crucial role for long-range frontoparietal connections in alterations of consciousness were provided in chapters 4 and 6. Here, we estimated functional and effective connectivity and their respective predictive power in patients suffering from disorders of consciousness. In chapter 4, we provided evidence for hemisphere-specific predictive power of functional connectivity; particularly left-hemisphere connectivity, but not right, predicted the DoC states. Going further, by contrasting patients assumed to be fully unaware with clinical states characterised by gradually progressing behavioural signs of consciousness, we provided evidence for a selective difference in frontoparietal connections, particularly in the left-hemisphere backward effective connectivity. Using the LOSOCV paradigm, we – for the first time – showed that not only did the left frontoparietal connectivity best explain the difference between the DoC patients, but also that the difference was large and consistent enough to distinguish the patient groups from each other. Crucially, we generalised this finding to an unseen group of potentially "covertly aware" patients distinguished by their underlying PET metabolism.

In what follows, we will examine the robustness of these results, reflect on them in light of previous related findings and theories, and speculate about the theoretical and practical implications. Further, we discuss the limitations of this research and how this work may inform the current state of consciousness research and suggest some potential directions for moving forward. Overall, by

modelling and contrasting data from two altered states of consciousness, we have provided novel insights into the brain dynamics underlying transitions between altered states of consciousness and highlighted the value of tracking these dynamics in a clinical context. DCM, though computationally more expensive, can accurately predict states of consciousness and provide causal explanations of the brain dynamics that cannot be inferred from functional connectivity alone. Functional connectivity, though correlational, is still an accurate predictive tool of altered states of consciousness. With clinically challenging, ambiguous cases like potentially covertly aware patients, we propose that the causal explanations and accurate predictions of DCM modelling could outweigh the computational complexity.

7.1   Robustness of the results – a 2-by-2 approach

Before considering our results in the wider context of the field, it is worth spending some time considering the robustness of our results. When studying complex cognitive processing, it is crucial to take into account not only research reflecting different perspectives, but also studies conducted with different methods. Critically, a null finding obtained with a particular method at a particular time – an absence of evidence for an effect – does not necessarily imply there is evidence for an absence of the effect, and thus, needs to be treated with caution. This is, of course, also true in science more generally.

Relevant to our results, the ongoing debates of the role of attention (Cohen, Cavanagh, Chun, & Nakayama, 2012; Koch & Tsuchiya, 2007; Pitts, Lutsyshyna, & Hillyard, 2018) and the frontal cortex in consciousness (Koch et

al., 2016a; Odegaard, Knight, & Lau, 2017) pinpoint a multitude of examples highlighting the importance of using multi-scale and multi-modal approaches to produce more integrated accounts of the underlying neural mechanisms. For example, "classic" frontal lesion case studies (e.g. Brickner, 1952; Fulton, 1949; Hebb & Penfield, 1940) are often (indirectly) cited as evidence when arguing against the role of prefrontal cortex (PFC) in consciousness (as, for example, in Koch et al., 2016a). However, as discussed in Odegaard et al. (2017), not only are important details about the extent of lesions often overlooked, more recent investigations of prefrontal lesions (Del Cul et al., 2009; Fleming, Ryu, Golfinos, & Blackmon, 2014) and short-term prefrontal impairments induced with transcranial magnetic stimulation (Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010) have suggested that even incomplete PFC lesions can impact subjective perceptual experiences. Hence, the argument goes, making strong claims about consciousness based on individual case studies can be problematic. Therefore, integrating information from multi-scale and multi-modal approaches should be the aim, for example, when investigating the neural mechanisms in DoC (Luppi et al., 2021).

We fully agree with these views and in our attempt to include heterogeneity in this research, we adopted a 2-by-2 approach, estimating both functional (FC) and effective connectivity (EC) in altered states of consciousness induced by propofol-anaesthesia and DoC. Here, FC is defined as undirected statistical dependence among distant neurophysiological events (Razi & Friston, 2016). Although correlational, FC can track how information and functional properties are shared between distant brain regions, and hence, it can be used as a biomarker

based on which the group from which a particular subject was sampled can be predicted or classified (as indeed we have done in chapters 3 and 4). The resulting classification, however, does not test any hypothesis about differences in the underlying brain coupling, as such classification problems establish a mapping from physiological consequences to a categorical cause (i.e. to a diagnostic class). This is in contrast to generative models, such as DCM, mapping from causes to consequences through hidden neurophysiological states (Friston, 2011). By comparing models of coupling among hidden brain states, generative models of effective connectivity can provide causal explanations of the brain dynamics, such that cannot be inferred from functional connectivity alone.

By focusing on the possible overlaps and dissociations in the results between the two measures of connectivity on the one hand, and between the two altered states of consciousness on the other, the level of confidence we can ascribe to our results increases. To this end, in chapter 3, we found baseline FC especially in the posterior electrodes best predicted propofol-induced loss of consciousness as measured with slow-wave activation saturation and loss of behavioural responsiveness. Correspondingly, at the level of neuronal sources in chapter 5, we found largest reductions in effective connectivity between normal wakefulness and LOC between sources in the posterior areas. It should be noted, though, that the measured scalp-level signal in posterior electrodes does not necessarily reflect underlying activity in posterior cortical areas, as the accuracy of the source localisation is affected by a number of factors (such as head-modelling errors, source-modelling errors, and EEG noise; Grech et al., 2008).

Nonetheless, anaesthetic-induced LOC was associated especially with connectivity changes in the posterior electrodes and sources. Taken together, these results support the suggested pivotal role for the posterior hot-zone of consciousness, previously observed in sleep (Siclari et al., 2017; Lee et al., 2019) and in patients with brain damage (Vanhaudenhuyse et al., 2010; Wu et al., 2015).

Overlapping results were observed also in DoC patients. The results in chapter 4 suggested that left-hemisphere alpha-band FC, and not right, predicted DoC groups at the level of scalp electrodes. Complementary to this, between DoC groups, the largest reduction in EC was observed in the left-hemisphere backward connection from medial prefrontal cortex to left parietal cortex. Moreover, connectivity from frontal to left parietal cortex was the single best predictor for the DoC groups and this result further generalised to an unseen dataset of potentially covertly aware unresponsive wakefulness syndrome (UWS) patients. Taken together, these results identify specific alterations in brain networks after severe injury, and further, highlight the clinical utility of EEG-based measurement of effective connectivity for identifying covert consciousness.

We also observed disconnects in the results within the methodological paradigms. Between the results in chapters 3 and 5, although the propofol-induced LOC was associated specifically with changes in connectivity in the posterior electrodes and sources, the best predictor for the state of consciousness at the neuronal source level was not localised between the posterior nodes.

Rather, in chapter 5, frontoparietal connectivity was found to be the best predictor for the state of consciousness and this result generalised into an unseen dataset from a recovery-state. In chapter 3, at the level of electrodes, frontoparietal baseline connectivity did significantly predict some of the slow-wave activity related measures, but to a lesser degree than when predicted from posterior electrodes only.

In chapter 6, we observed the largest reduction in left-hemisphere frontoparietal effective connectivity when comparing UWS PET- patients to healthy controls or to patients in minimally conscious state (MCS+). However, when comparing MCS+ patients to healthy controls, the largest reductions were observed in the connectivity between the posterior nodes. This was true also for the predictive power: left frontoparietal connectivity best predicted the groups when comparing UWS PET- to either healthy controls or MCS+, but did not produce accurate predictions for MCS+ vs. healthy controls comparison.

This is potentially an interesting observation as the posterior hot zone has been argued to relate more with phenomenological consciousness – i.e. with the subjective content of consciousness (Boly et al., 2017) – rather than with cognitive consciousness (access-consciousness; Block, 1995). The neural correlates of access-consciousness have been argued to constitute a widespread global workspace including frontal, posterior, and temporal cortices (e.g. Dehaene & Changeux, 2004; Mashour et al., 2020). On the other hand, the posterior hot zone has been argued to consist of the key areas for phenomenological consciousness and frontal areas to reflect neural processes

subsequent to consciousness, such as selective attention, working memory, and task reporting (Boly et al., 2017; Koch et al., 2016a). One might speculate that the observed modulations when comparing clearly conscious patients (MCS+) to healthy controls does not reflect sufficient difference in access-consciousness to reliably distinguish the states based on frontoparietal connectivity. This is in contrast to comparisons between presumably fully unconscious patients (UWS PET-) and MCS+ or controls, where frontoparietal connections were modulated the most and produced the best predictions. We will come back to this point later in this chapter and consider these results in conjunction with the observations from propofol anaesthesia providing a prime example for the usefulness of integrating results from different methodologies and domains.

7.2   Considerations about the chosen methods

Parallel to such broad considerations of the robustness of the results in relation to the chosen methodological approach, there are also a number of more limited complications in relation to the used *methods*. For example, in chapters 3 and 4, we performed linear multivariate and linear multiple regressions, respectively, to predict the state of consciousness. However, it is not obvious that such an assumption of linearity is justified; this is probably more true for the DoC (chapter 4) than for anaesthetic-induced LOC in that it is far from clear what a linear "increase" in consciousness (when moving from coma to UWS to MCS to neurotypical wakefulness) is supposed to mean (Bayne et al., 2016). Nonetheless, progression in DoC is commonly characterised by alterations in

arousal and/or awareness in a quasi-linear fashion (Laureys, 2005). Previous studies have suggested that progression in DoC can – from a simple perspective – be characterised and classified in a linear fashion from EEG (Chennu et al., 2017; Sitt et al., 2014). Similarly, although assuming such linear relationships is a limitation here, we were able to predict state of consciousness both in anaesthesia and in DoC regardless of the assumption.

What about our DCM modelling results – can we trust that we have obtained the "true" underlying model in our observations? In short, no. To begin with, in DCM for EEG, the specification of the anatomical locations of the nodes/sources are required a priori. Performing source localisation first and DCM second, would increase the likelihood of estimating the connectivity between "true" sources of brain activity. Here, optimally, individual MRIs should be used as the basis for the head model, and as discussed in detail in chapter 6, this would be especially important in the case of patients with damaged brains. That said, we chose our methodology in chapters 5 and 6 to best mitigate these issues (for detailed description of the mitigating factors, see chapters 5 and 6). For example, we modelled only well-known resting state networks previously associated with alterations in consciousness, used parametric empirical Bayes with Bayesian model reduction (BMR) to avoid making strong, a priori, hypotheses about the

changes in the connectivity, and restricted our DoC sample to patients with traumatic brain injury.[6]

Second, it is important to understand that the best model identified by Bayesian model selection (here, BMR) is only the best model *among the models tested*. It may be that sources outside of our model space – and hence, missed in our analysis – contributed towards conscious awareness. Moreover, simulation studies have shown that Bayesian model selection may favour more complex models than the one actually generating the data (Litvak et al., 2019). This latter point is a concern especially in chapter 5 with the large model, as the implicit bias in DCM is towards non-sparse models (although, a model with small modulations of a large number of parameters may also be less complex than a model with large modulations of a few parameters. See Litvak et al. (2019) for details).

For these reasons, in conjunction with an appetite to obtain practically useful results, we applied a novel DCM-based cross-validation to establish the predictive validity of our models. The observed accurate prediction performance with cross-validation in both propofol anaesthesia and DoC data increased the level of confidence we can ascribe to our modelling results. Furthermore, we demonstrated generalisation of this predictive power in unseen data sets from the

---

[6] Traumatic brain injury (TBI) has been associated with more focal injury centred often in deeper brain areas rather than in the cortex (Guldenmund, Soddu, et al., 2016; Newcombe et al., 2010).

post-anaesthetic recovery state and potentially covertly aware patients. These results indicated that effective EEG connectivity could be used to track anaesthetic states and to potentially identify covertly aware patients who seem behaviourally unresponsive. Hence, the use of effective connectivity estimation in conjunction with cross-validation and behavioural assessments, could potentially reduce the diagnostic error with patients and provide more accurate estimations of the depth of anaesthesia.

The benefit of using cross-validation and especially demonstrating generalisation to unseen data becomes even more prominent when considering our relatively small sample sizes in chapters 5 and 6 (with $N = 10$ in propofol anaesthesia and $N = 6$, $N = 5$, and $N = 12$ in UWS PET-, UWS PET+ and MCS+ patients, respectively). The fact that the predictive power generalised to unseen data – if not anything else – speaks to the aphorism "all models are wrong but some are useful" (Box, 1976) and indicates practical efficacy of our modelling work that could potentially be used in the clinical context.

Nonetheless, to make stronger claims about the underlying brain networks, these results should first be reproduced and validated. This is particularly important under the current replicability crisis in which the results of many pivotal studies where sample sizes typical for e.g. neuroimaging studies, have turned out to be underpowered, and therefore difficult or indeed impossible (to

date) to reproduce.[7] Applying cross-validation with generalisation to unseen data can mitigate such concerns of replicability, but does not replace the key part of the scientific method that is reproducibility.

7.3   Validity in measuring consciousness

Another limitation worth mentioning is the problem of coordination (in relation to consciousness, see Michel, 2019). That is, whether the measurement procedure used actually measures what it is intended to measure. Here, with anaesthesia, we used responsiveness – or lack thereof – as a proxy for consciousness, although, for example, dreaming and dream-like states in anaesthesia indicate that unresponsiveness does not necessarily indicate unconsciousness (Leslie, 2017; Leslie et al., 2009; Nir & Tononi, 2010; Noreika et al., 2011; Radek et al., 2018). Similarly, the DoC diagnoses were based on the currently most sensitive behavioural scale to disentangle the DoC groups from each other (Coma Recovery Scale-Revised; Giacino, Kalmar, & Whyte, 2004; Seel et al., 2010). However, the extent to which these behavioural measures capture changes in consciousness is unknown – and unknowable until, of course, objective and robust measures of consciousness are established against which the behavioural outcomes can be reflected. And therein lies the problem; we cannot know if we measure what we intent to measure without access to a robust and

---

[7] This crisis has been most frequently discussed in relation to the fields of psychology and medicine, but concerns sciences more generally (Baker, 2016; Pashler & Wagenmakers, 2012).

reliable measure. This circularity is an issue in science more generally, and not just with the problem of consciousness. Luckily, this problem has been overcome in the past (Chang, 2004). One way – and perhaps the only way – to solve the issue is to accept that we need to start with some (reasonable) assumptions about the validity of the measures. Once more evidence is gathered, the precision of these measures can be increased, for example, via forming predictions based on the assumed validity, testing the measurement in an experimental setting, and by making adjustments to improve the accuracy of the predictions. Here, model recovery analyses can make an important contribution by providing a ground-truth, albeit a "simulated" one (Wilson & Collins, 2019).

We do acknowledge the problems in assuming ground truths about the state of consciousness based on behaviour. However, as discussed in more detail in section 2.6.5, assessments of conscious experiences are commonly inferred from behavioural activity and responsiveness (as we are only privy to our own subjective experience of consciousness). This is especially true in the clinical setting, where behavioural responsiveness is the first and most common assay of consciousness. Moreover, assessments of behavioural responsiveness approximate the intermediary stages by probing the particular stage where responsiveness is lost. Finally, behavioural markers of consciousness in general have been shown to be useful in practice; for example, the Coma Recovery Scale has been shown to correlate with the prognosis of DoC patients (Bruno et al., 2012; Giacino, Fins, Laureys, & Schiff, 2014). Therefore, our starting point was to accept the assumption that behavioural responsiveness approximates conscious processing and that it can be used as an objective measure to indicate the

presence or lack of consciousness (for more detailed discussion, see section 2.6.5).

Overall, both the integration of results crossing two altered states of consciousness with two analytical methods for measuring the associated changes in the brain, and the utilisation of cross-validation increases the level of confidence we can ascribe to our results. Even allowing that our modelling results are necessarily wrong (Box, 1976), our results indicate usefulness in practical (clinical) context. In the next section, we leave the consideration of the robustness of our results aside and instead, discuss our observations in relation to prominent theories of consciousness and to previous findings.

7.4 The main findings in relation to Global Neuronal Workspace and Integrated Information Theory

Assuming that our results are true, what new can we say about consciousness? With the risk of losing the interest of hard idealists, this work does not directly provide any new insights into the mechanisms of phenomenological consciousness (i.e. towards solving the hard problem; Chalmers, 1995). We do, however, provide novel insights towards the real problem of consciousness (Seth, 2016) and contribute to the debate about whether the anterior parts of the brain are necessary for conscious experience.

A straightforward interpretation of our effective connectivity results supports the globalist over the localist dogma of consciousness; long-range frontoparietal connectivity had the best predictive power of the state in anaesthetic-induced

LOC and DoC. Moreover, a key difference was observed between healthy controls or conscious patients and completely unresponsive patients in left-lateralised backward frontoparietal connectivity. In propofol anaesthesia, our results highlighted the relative importance of DMN over salience and central executive networks: we observed a selective breakdown of the DMN with decreases in effective connectivity to and from PCC/precuneus, including the medial prefrontal connectivity (although importantly, most of the reductions were located between nodes located within the posterior hot zone). These results – especially the modelling results with DoC and cross-validation results in both of the investigated altered states of consciousness – are consistent with the notion that conscious processing requires a brain-wide "global workspace" that constitute a widespread network in the frontal, parietal, and temporal cortices (Dehaene et al., 1998; Dehaene & Changeux, 2011; Dehaene et al., 2011).

According to the group of Dehaene and Changeaux, Global Neuronal Workspace Theory (GNWT) is limited to neural correlates of access consciousness (AC; Dehaene & Changeux, 2004; Mashour et al., 2020) instead of the more elusive phenomenological consciousness (PC; Block, 1995). Consciousness is postulated to closely relate to mechanisms of attention, working memory (Dehaene & Changeux, 2004), and metacognition (Dehaene, Lau, & Kouider, 2017). Although, according to a recent review, the involvement of consciousness in such mechanisms is considered indirect and the mechanisms distinct (Mashour et al., 2020).

On the other hand, Tononi's group is predominantly involved with identifying the neural correlates of PC (subjective consciousness); these consist of the cortical areas sustaining the neural mechanisms for any particular phenomenal contents within consciousness, "such as colors, faces, places, or thoughts" (Boly et al., 2017). According to Integrated Information Theory (IIT), consciousness requires the potential for widespread interactions between the neural correlates of PC, that form an integrated network in the cortex, complemented by *information* and *exclusion* (Oizumi et al., 2014). In IIT, no claims are made about the necessary brain areas for such an integration of information. Hence, the reductions we observed in frontoparietal connectivity can loosely be interpreted in support of IIT (i.e. reduced integration). Connectivity, and especially effective connectivity, can be viewed to reflect how information from different specialised systems is integrated in the cortex (Friston, 2011).

Nonetheless, despite the fact that the integrated network generating the qualia of wakeful consciousness at any given time is considered to be dynamic (Oizumi et al., 2014), Tononi and collaborators consider the posterior hot zone (temporo-parieto-occipital areas) as essential for consciousness. Crucially, the frontal areas are left outside of the neural correlates of consciousness (Boly et al., 2017; Koch et al., 2016a). Our results with functional (chapter 3) and effective (chapter 5) connectivity can be interpreted to support the key-role of the posterior hot zone. In chapter 3, the posterior electrodes (at the scalp level) best predicted various indirect measures of consciousness. Although without source localisation, we cannot know where the activity originated from (Michel et al., 2004), the posterior connectivity provided more reliable predictions than, for example,

connectivity between frontal and parietal electrodes. Moreover, in chapter 5, we observed most reductions in effective connectivity between posterior neuronal sources in the model best explaining the difference between normal wakefulness and unconsciousness.

Taken together, on the one hand our results support GNWT in that long-range frontoparietal connectivity was found to play a pivotal role in explaining and predicting the difference between wakeful consciousness and the two altered states. On the other, with propofol-anaesthesia and when contrasting conscious patients (MCS+) with healthy controls, most reductions in effective connectivity were located within the posterior hot zone. In what follows, we speculate how these observations might align with GNWT and IIT, and contribute to the debate about whether the neural correlates of consciousness have an anterior component (Bor & Seth, 2012; Del Cul et al., 2009; Laureys & Schiff, 2012) or are predominantly posterior (Boly et al., 2017; Koch et al., 2016a; Koch et al., 2016b; Siclari et al., 2017).

7.4.1 Access vs. phenomenological consciousness

For speculation's sake, let's assume, as has been argued, that GNWT indeed relates to neural correlates of AC (Dehaene & Changeux, 2004; Mashour et al., 2020) and the posterior hot zone to neural correlates of PC (Boly et al., 2017; Koch et al., 2016a). Under these assumptions, it is reasonable to think that effective connectivity reflecting AC necessitates involvement of prefrontal cortex, while PC would relate more to the connectivity between posterior nodes

(for more detailed specification and comparison of the related areas, see Frigato, 2021). If true, our results suggest that changes in the neural mechanisms that relate to access-consciousness are better for predicting the state of consciousness in both, DoC and propofol-anaesthesia. In other words, with these data, there were consistent group-level changes (with large enough effect sizes) in cognitive processes that relate to consciously accessible information, such that the state of consciousness in unseen data could be reliably predicted. When comparing unconscious DoC patients to healthy controls or to conscious patients, we similarly observed the key-difference in AC-related connectivity between the frontoparietal nodes, while no reduction between the nodes in the posterior hot zone.[8] To continue with the speculation, assuming that PC indeed relates to the posterior hot zone, these findings are consistent with the notion of possible residual phenomenological consciousness in some seemingly unaware patients.

When comparing conscious patients to healthy controls, we did not observe reductions in effective connectivity involving nodes in the prefrontal cortex. In other words, we did not observe any connectivity-related markers of diminished

---

[8] Note that here we applied a relatively strict criterion (>= .99 posterior probability) for the connectivity modulations to be included in the most parsimonious model that best explains the data. Such a threshold helps focusing on the most probable effects in the model. Removing the threshold criterion, one reduction when comparing UWS PET- and controls, and two reductions when comparing UWS PET- and MCS+ were observed between the posterior nodes (see figure 6.6). However, in addition to the relatively lower probability of the effects being present in the most parsimonious model, the corresponding effect sizes were clearly smaller than in the frontoparietal reduction.

AC between MCS+ patients and healthy controls. This is consistent with the diagnostic criteria of MCS+ (e.g. capability for command following, intelligible verbalisation or gestures with intentional communication) and with the evidence of less functional disability differentiating these patients from UWS and MCS-patients (Thibaut, Bodien, Laureys, & Giacino, 2020).

To finish with the speculation, with propofol anaesthesia, we observed most of the reductions within the posterior hot zone, which – following the above reasoning – could be interpreted to reflect diminished PC, and hence, possibly abolished subjective experience all together. However, we similarly observed reductions in posterior hot zone connectivity when comparing conscious patients and healthy controls. This observation, coupled with not observing reduced frontoparietal connectivity, would be difficult to explain under the assumptions of relating AC directly to an involved anterior component and PC to posterior hot zone. If we take decreases in posterior effective connectivity to reflect diminished PC, we would not expect to see such connectivity reductions in conscious patients compared to healthy controls, but not with unconscious patients as arguably, conscious patients do have subjective experiences. Hence, it may be more useful to withhold from drawing such a confining distinction between AC and PC here, and instead follow the suggestion of Dehaene & Naccache (2001) in that "global availability of information (…) is what we subjectively experience as a conscious state". Defining consciousness this way, conscious access and phenomenological experience appear together.

7.4.2 Is the anterior cortex needed for consciousness?

The effective connectivity models in chapters 5 and 6 indicated that loss of consciousness in anaesthesia and in DoC modulated long-range frontoparietal connectivity suggesting a role for an anterior component in the most parsimonious model. Regardless of the distinction of consciousness into AC and PC, it has been argued that the anterior parts of the brain do not contribute towards consciousness per se, but rather, reflect other, subsequent cognitive processes such as attention and working memory (Boly et al., 2017; Koch et al., 2016a). Consequently, one might critically evaluate our interpretations of the modelling results and suggest that the involvement of an anterior component relates to cognitive processes such as attention, and not directly to consciousness. We recognise that this may indeed be the case, as discussed in chapter 5. However, it is not clear what consciousness without any such cognitive processes might look like. Indeed, subjectively, loss of consciousness does not seem to be a binary event, but rather a gradual process with intermediary stages including, for example, narrowed attention, decrease of memory, impaired cognition, and lower

self-estimated level of consciousness (Esaki & Mashour, 2009; Vaitl et al., 2005).[9]

Evidence suggests that lack of attention can lead to absence of awareness (e.g. inattentional blindness; Mack & Rock, 1998; Mack, 2003; change blindness; Jensen, Yao, Street, & Simons, 2011; the attentional blink; Dux & Reńmarois, 2009; Shapiro, Raymond, & Arnell, 2009). Similarly, evidence obtained with a variety of paradigms have indicated that attention can operate or be drawn to stimuli without the information ever reaching consciousness (Bola, Paź, Doradzińska, & Nowicka, 2021; Bussche, Hughes, Humbeeck, & Reynvoet, 2010; Finkbeiner & Palermo, 2009; Kentridge, Nijboer, & Heywood, 2008; Koch & Tsuchiya, 2007). Going further, numerous authors have claimed a true double dissociation between attention and consciousness, suggesting that one may be conscious without attention (Block, 2011; Van Boxtel, Tsuchiya, & Koch, 2010; Koch, 2004; Koch & Tsuchiya, 2007; Tsuchiya & Koch, 2015). Hypotheses,

---

[9] Until the beginning of the 21st century, attention and consciousness were commonly viewed as one and the same process or at least directly linked with attention providing a gateway to consciousness (e.g. Broadbent, 1958; Posner, 1994). In recent decades, attention and its relationship to consciousness has provoked more discussion among the cognitive processes argued as subsequent to consciousness. Thus, here, we limit our discussion to attention. That is not to say that attention is the only cognitive process potentially linked with consciousness. For example, for recent discussions about the relationship between consciousness and working memory, see Persuh, Larock, & Berger (2018) and Velichkovsky (2017).

such as the overflow hypothesis (Block, 2007, 2011), suggest that information can "overflow" working memory and attention, and that essentially when observing a complex scene, we may be conscious of more than what we can think, know, or report about. If we do accept a strong dissociation between attention and consciousness, it is indeed possible that the role of frontoparietal effective connectivity modulations we observed relates to attention, or some other subsequent cognitive processes rather than consciousness.

It is not clear, however, how conscious processing *without the subject having any knowledge about that conscious processing* differs from truly unconscious processing. Certainly, there can be *behaviour* without any awareness of the conscious state. However, if no attention is paid to the state of experience and no memory is allocated for this information, the concept of conscious processing becomes more obscured. After all, if the subject does not know that they are in a conscious state, what then – if anything – does it mean to be conscious? Hence, it seems a plausible starting point to assume that for full conscious experience, functions such as working memory and attention are needed. To this end, increasing amounts of evidence support the view that at least *some type* of attention is required for consciousness (Cohen et al., 2012; Pitts et al., 2018). This renders the frontoparietal connections a necessary part of the mechanisms required for consciousness (Bor & Seth, 2012; Dehaene & Changeux, 2011; Mashour et al., 2020). Therefore, we can conclude that, even if the anterior parts of the brain relate to attention, the observed reductions in frontoparietal effective connectivity in conjunction with modulations in the posterior hot zone indeed seem to have a causal role in explaining the loss of consciousness.

224

7.5 Research findings and clinical utility

Our results with functional connectivity suggested that baseline alpha- and beta-dwPLI, especially when measured from the posterior electrodes, can reliably predict loss of consciousness caused by propofol-anaesthesia. Similarly, the results hinted predictive power for left hemisphere dwPLI in the alpha-band when classifying DoC states. Many factors have been shown to have an effect on the patient's response to anaesthesia: for example, with propofol, demographic characteristics such as weight, age, and sex have been associated with differences in the response to the drug (Gambús & Trocõniz, 2015; Schnider et al., 1999). Indeed, administering the right amount of anaesthetics can be difficult and requires continuous attention and care from the clinicians. Hence, a reliable, relatively easy to use, cost-effective, and above all, objective, tool that could indicate the state of the patient with regards to the amount of anaesthetic drugs required *before* any drugs are administered, would make the clinicians' task easier and could potentially prove to be even lifesaving.

The results presented in this thesis, in chapter 3 especially, are congruent with previous results that have suggested a relationship between baseline functional connectivity and individual susceptibility to anaesthesia (Chennu et al., 2016; Deng et al., 2019; Zhang et al., 2020). Based on the earlier research findings and on the results in this thesis, it is plausible that clinicians could get a real-life benefit by profiling the patients based on functional EEG-connectivity prior to administering anaesthetics. In order to increase the potential utility, future

research should further focus and expand the current knowledge about the
relationship between baseline connectivity and the depth of anaesthesia and to
investigate how – if at all – other factors such as anxiety or alertness interact with
that relationship. Moreover, validating from where, and at which frequency, the
susceptibility to anaesthesia can be best predicted would be worthwhile in near
future.

It could also prove to be worthwhile to expand this approach beyond mere
functional connectivity alone. For example, previous research has suggested that
other phase-synchrony based measures, including multiscale- and phase-lag
entropy, can not only discriminate DoC states from each other (coma from quasi-
brain-death; Li et al., 2012) but also provide novel insight into the intra-cortical
information flow within the underlying signal generating system (Ahmed, Li,
Cao, & Mandic, 2011) and even show promise as an anaesthetic-depth indicator
similar to what was discussed in chapter 3 (Ki et al., 2019; Shin et al., 2020).
Moreover, phase-synchrony based measures have been shown to fluctuate in
people due to, for example, heart failure (Costa, Goldberger, & Peng, 2002),
Alzheimer's disease (Hornero, Abásolo, Escudero, & Gómez, 2009), and ageing
(Takahashi et al., 2009), suggesting that there may be a complexity-loss in
systems "under stress" (Goldberger et al., 2002).

As discussed in chapter 3, the patient's baseline functional state – and by
extension the patient's individual susceptibility to anaesthesia – can also vary due
to similar "stress" (Gambús & Trocõniz, 2015; Hong, Jee, & Luthardt, 2005; Kil
et al., 2012; Laalou et al., 2010). Hence, it is reasonable to hypothesise a possible

correlation between baseline complexity and individual susceptibility to anaesthesia, especially as such a relationship between complexity (as measured with multi-scale entropy; MSE) and functional connectivity (Pearson correlations between spike train time series) exists (Wang et al., 2018). However, as far as we can tell, whether anaesthetic depth can be predicted based on baseline MSE is still an open question.

With the DCM models, the observed accurate prediction performance with cross-validation in both propofol anaesthesia and DoC data increased the level of confidence we could ascribe to our results. The modelling results may potentially help researchers to identify the underlying mechanisms causing these states. Furthermore, we demonstrated generalisation of the predictive power in unseen data sets from the post-anaesthetic recovery state and potentially covertly aware patients. These results indicated that effective EEG connectivity could be used to track anaesthetic states and to identify potentially covertly aware patients who seem behaviourally unresponsive. Hence, the use of effective connectivity estimation in conjunction with cross-validation and behavioural assessments, could potentially reduce the diagnostic error with patients and provide more accurate estimations of the depth of anaesthesia. First, these results should be replicated and validated in larger datasets.

DCM, however, is computationally quite expensive process and furthermore, is not straightforward to perform. In other words, carrying out this type of analysis with the LOSOCV predictions requires the researcher to invest time and effort to master the process and the model fitting process itself can take a lot of

time; the procedure may run days or even longer depending on the complexity of the specified models. Some of the steps in the process, however, could possibly be automated, and it could indeed be a worthwhile project to investigate the extent to which such automation is possible, for example, by following the steps taken in chapter 5.

Moreover, with such clinical datasets as used in this thesis, signal noise is a common issue (indeed as was the case especially with the DoC data used in chapters 4 and 6). This, of course, complicates usage of complex modelling procedures, such as DCM, and often careful manual calibration of the process is required in order to obtain robustness against noise. Nonetheless, assuming the results presented in this thesis are successfully replicated and validated, we do think that especially with clinically challenging, ambiguous cases like potentially covertly aware patients, producing causal models and performing accurate predictions via DCM modelling outweighs the complexity and computational cost of the process.

7.6 Future directions

In what follows, we suggest some future directions that can extend this work and the suggestions made in previous chapters.

7.6.1 State vs. trait in anaesthetic-induced loss of consciousness

In chapter 3, we provided evidence for predicting various indirect measures of anaesthetic-induced LOC from baseline EEG functional connectivity. We discussed how slow-wave activation saturation (SWAS) has been suggested to have a sound neurobiological basis, and hence, to potentially represent a phenotype of an underlying trait (Warnaby et al., 2017). On the other hand, when predicting individual propofol concentration levels, previous studies have shown that observed differences in the baseline alpha-network robustness are abolished at recovery (Chennu et al., 2016). This suggests that the differences are depending on a latent alpha-state, rather than any individual trait.

Nonetheless, our work in chapter 3 suggests that SWAS, as a proxy for unconsciousness, is distinct from direct propofol concentration levels, and hence, may indeed reflect a phenotype of a trait. To further investigate whether SWAS and its association with baseline alpha-connectivity is state- or trait-based, future studies could explore to what extent, if any, the predictive power and the individual differences observed are abolished at recovery. For example, does baseline alpha also predict the concentration of propofol and time needed to reach recovery of behavioural responsiveness, the SWA-power at the end of the saturation period, or at recovery of behavioural responsiveness? Such work could further illuminate the apparent distinction between SWAS on the one hand, and other proxies for unconsciousness, such as the individual propofol concentration levels on the other.

7.6.2 More advanced DCM modelling in anaesthesia and in disorders of consciousness

In chapter 4, we provided preliminary evidence for a hemispheric difference in DoC. Specifically, left-hemispheric alpha-band functional connectivity predicted the DoC state of the patients, while the right-hemispheric connectivity did not. The opposite result was observed with the network hubs: right-hemispheric participation coefficient in the alpha-band statistically significantly predicted the DoC states, while left-hemispheric participation coefficient did not. The left alpha-connectivity furthermore statistically significantly differentiated MCS- from MCS+.

These results were extended in chapter 6, where we showed that specifically left frontoparietal effective connectivity was reduced in unconscious patients when compared to conscious patients or healthy controls. Moreover, the left frontoparietal connectivity predicted DoC groups and generalised to an unseen data set of potentially covertly aware patients. Previous studies have indicated higher left hemisphere metabolism for MCS+ over MCS- patients, specifically in cortical areas associated with language, such as Broca's and Wernicke's areas (Aubinet et al., 2020; Bruno et al., 2012) and functional connectivity differences in the left hemisphere between UWS and MCS patients (with the latter showing higher connectivity levels; Lehembre et al., 2012).

However, at the neuronal level, we only modelled sources within the DMN. The hemispheric difference (specifically in relation to language functions) could be further investigated by the means of DCM by building additional models with

nodes in the temporal, language-related areas. For example, estimating two mirroring models for each hemisphere separately would enable one to use the estimated free energies to gain evidence for the model better explaining the data (Friston et al., 2007).

Moreover, by choosing a different neuronal model, the focus could be turned from network-level between-source effects, to within-source effects and to more fine-grained properties within the sources (e.g. synaptic properties). As discussed in chapter 5, this could be done, for example, by using the LFP model or the Canonical Microcircuits model (Bastos et al., 2012; Moran et al., 2007). Furthermore, we did not investigate the predictive power of the right lateralised network hubs at the level of neuronal sources. For this, DCM could potentially be used. As discussed in detail in chapter 6, with further DCM work, optimally, individual MRIs should be used as the basis for the head model. Furthermore, instead of assuming the neuronal sources, performing a source localisation first would increase the likelihood of modelling the connectivity between "true" sources.

Similar suggestions can be made also in relation to the "frontoposterior" vs. "posterior hot zone only" debate. With DCM, extensive models including neuronal sources only in anterior or posterior areas could be estimated and compared. A cross-validation paradigm could be then used to see which models (or combinations of models) produce the best predictions. A useful additional step in the paradigm applied in this thesis would be to validate the results of the model comparison and the parameter fitting with simulated data. That is, to

determine the extent to which data generated from the most parsimonious model is the best fit by that model, as opposed to other models and to data generated from other models (Wilson & Collins, 2019).

7.6.3 DCM modelling of contents of consciousness

Lastly, in this thesis we have provided evidence for DCM-based predictions of the level or state of consciousness. Similar DCMs with cross-validation could potentially be used to investigate the underlying neuronal mechanisms of *content* of consciousness. To elucidate the neural correlates of content of consciousness further, one could – in addition to estimating the models – see if DCM-based cross-validation can predict changes in the content within the participants. To do this, a paradigm including bistable perception (binocular rivalry) could be used to manipulate the perception in the participants (for a review, see Sterzer, Kleinschmidt, & Rees, 2009). As per the arguments presented in earlier in this chapter, it would be beneficial to integrate information from multiple modalities to increase the level of confidence attributed to the results.

It is worth noting that in this thesis, we cut off higher frequencies as most electromyographic noise was observed in these high frequencies. This is typically caused by involuntary movements and fidgeting characteristic more in DoC than under anaesthesia. Further, the DCM analysis was performed for the broadband and not for particular canonical frequency bands distinctively. Yet, previous research has suggested a relationship between loss of consciousness and gamma-activity; indeed, in the beginning of the systematic search for the neural

correlates of consciousness, synchronisation of neuronal populations via rhythmic discharges in the gamma range was proposed as the key mechanism for consciousness and as a mechanism for perceptual binding (Gray, König, Engel, & Singer, 1989; Crick & Koch, 1990; Varela, Lachaux, Rodriguez, & Martinerie, 2001). More recent research has suggested a relationship between gamma-band activity and loss of consciousness due to anaesthesia (Bola et al., 2018) and DoC (Cavinato et al., 2015; Papiernik, Binder, & Gawłowska, 2022). Moreover, an association between gamma-band phase synchrony and emergence of coherent conscious perception has been shown using binocular rivalry (Varela et al., 2001). Therefore, it would be interesting to see if the predictive power of the DCM models could be increased by limiting the analysis to higher frequency bands, instead of analysing the whole broadband.

7.7   In summary

In this thesis, we have modelled functional and causal (effective) connectivity of brain functioning at the network level scales in altered states of consciousness. By modelling changes at the level of electrodes and in key resting state networks, we have provided novel insights into the structures and functions involved in (loss of) consciousness. By accurately predicting states of consciousness based on the connectivity models, we provided evidence for clinical relevance and practical value of not only the estimated models but the applied methods as well. In doing so, we have enhanced the understanding of consciousness as a neuroanatomically specific state which maintenance seems to involve long-range

233

frontoparietal communication. Our work on applying more advanced computational modelling, especially in conjunction with cross-validation, provided insightful, causal explanations of the neuronal mechanisms underlying the loss of consciousness, and could help reduce diagnostic error in clinical context.

Bibliography

Achard, S., Delon-Martin, C., Vértes, P. E., Renard, F., Schenck, M., Schneider, F., … Bullmore, E. T. (2012). Hubs of brain functional networks are radically reorganized in comatose patients. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(50), 20608–20613. https://doi.org/10.1073/pnas.1208933109

Adamantidis, A. R., Gutierrez Herrera, C., & Gent, T. C. (2019). Oscillating circuitries in the sleeping brain. *Nature Reviews Neuroscience*, *20*(12), 746–762. https://doi.org/10.1038/s41583-019-0223-4

Adey, W. R., Walter, D. O., & Hendrix, C. E. (1961). Computer techniques in correlation and spectral analyses of cerebral slow waves during discriminative behavior. *Experimental Neurology*, *3*(6), 501–524. https://doi.org/10.1016/S0014-4886(61)80002-2

Ahmed, M. U., Li, L., Cao, J., & Mandic, D. P. (2011). Multivariate multiscale entropy for brain consciousness analysis. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 810–813. https://doi.org/10.1109/IEMBS.2011.6090185

Alkire, M. T., Haier, R. J., Barker, S. J., Shah, N. K., Wu, J. C., & Kao, J. Y. (1995). Cerebral metabolism during propofol anesthesia in humans studied with positron emission tomography. *The Journal of the American Society of Anesthesiologists*, *82*(2), 393–403.

Alkire, M. T., Hudetz, A. G., & Tononi, G. (2008). Consciousness and Anesthesia. *Science*, *322*(5903), 867–880. https://doi.org/10.1016/B978-0-12-800948-2.00009-1

Annen, J., Frasso, G., Crone, J. S., Heine, L., Di Perri, C., Martial, C., … Laureys, S. (2018). Regional brain volumetry and brain function in severely brain-injured patients. *Annals of Neurology*, *83*(4), 842–853. https://doi.org/10.1002/ana.25214

Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Moran, R., Henson, R., … Phillips, C. (2017). SPM12 Manual The FIL Methods Group (and honorary members). *Functional Imaging Laboratory*, 475–1. https://doi.org/10.1111/j.1365-294X.2006.02813.x

Aubinet, C., Cassol, H., Gosseries, O., Bahri, M. A., Larroque, S. K., Majerus, S., … Thibaut, A. (2020). Brain Metabolism but Not Gray Matter Volume Underlies the Presence of Language Function in the Minimally Conscious State (MCS): MCS+ Versus MCS− Neuroimaging Differences. *Neurorehabilitation and Neural Repair*, *34*(2), 172–184. https://doi.org/10.1177/1545968319899914

Baars, B.J. (1988). A Cognitive Theory of Consciousness. *NY: Cambridge University Press*.

Baars, Bernard J. (1997). In the theatre of consciousness. Global Workspace Theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4*(4), 292–309. https://doi.org/10.1093/acprof:oso/9780195102659.001.1

Baker, M. (2016). Is there a reproducibility crisis in science? *Nature*, *533*(7604),

452–454. https://doi.org/10.1038/d41586-019-00067-3

Baker, R., Gent, T. C., Yang, Q., Parker, S., Vyssotski, A. L., Wisden, W., …
Franks, N. P. (2014). Altered activity in the central medial thalamus precedes
changes in the neocortex during transitions into both sleep and propofol
anesthesia. *Journal of Neuroscience*, *34*(40), 13326–13335.
https://doi.org/10.1523/JNEUROSCI.1519-14.2014

Barrett, A. B., Murphy, M., Bruno, M. A., Noirhomme, Q., Boly, M., Laureys, S., &
Seth, A. K. (2012). Granger causality analysis of steady-state
electroencephalographic signals during propofol-induced anaesthesia. *PLoS
ONE*, *7*(1). https://doi.org/10.1371/journal.pone.0029072

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K.
J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–
711. https://doi.org/10.1016/j.neuron.2012.10.038

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in
syndrome. *Journal of Neurology*, *221*(2), 77–91.
https://doi.org/10.1007/BF00313105

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are There Levels of Consciousness?
Trends in Cognitive Sciences. https://doi.org/10.1016/j.tics.2016.03.009

Bell, A. J., & Sejnowski, T. J. (1995). A non linear information maximisation
algorithm that performs blind separation. In G. Tesauro, D. S. Touretzky, & T.
Leen (Eds.), *Advances in Neural Information Processing Systems 7* (p. 467).
The MIT Press, Cambridge, Massachusetts.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227–247. https://doi.org/10.1017/S0140525X00038188

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*(5–6), 481–548. https://doi.org/10.1017/S0140525X07002786

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, *15*(12), 567–575. https://doi.org/10.1016/j.tics.2011.11.001

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), 1–12. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bodart, O., Amico, E., Gómez, F., Casali, A. G., Wannez, S., Heine, L., … Gosseries, O. (2018). Global structural integrity and effective connectivity in patients with disorders of consciousness. *Brain Stimulation*, *11*, 358–365. https://doi.org/10.1016/j.brs.2017.11.006

Bodart, O., Gosseries, O., Wannez, S., Thibaut, A., Annen, J., Boly, M., … Laureys, S. (2017). Measures of metabolism and complexity in the brain of patients with disorders of consciousness. *NeuroImage: Clinical*, *14*, 354–362. https://doi.org/10.1016/j.nicl.2017.02.002

Bojak, I., & Liley, D. T. J. (2005). Modeling the effects of anesthesia on the electroencephalogram. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *71*(4), 1–22. https://doi.org/10.1103/PhysRevE.71.041902

Bola, M., Barrett, A. B., Pigorini, A., Nobili, L., Seth, A. K., & Marchewka, A. (2018). Loss of consciousness is related to hyper-correlated gamma-band activity in anesthetized macaques and sleeping humans. *NeuroImage*, *167*, 130–142. https://doi.org/10.1016/j.neuroimage.2017.11.030

Bola, M., Paź, M., Doradzińska, Ł., & Nowicka, A. (2021). The self-face captures attention without consciousness: Evidence from the N2pc ERP component analysis. *Psychophysiology*, *58*(4), 1–13. https://doi.org/10.1111/psyp.13759

Boly, M., Moran, R., Murphy, M., Boveroux, P., Bruno, M.-A., Noirhomme, Q., … Friston, K. (2012a). Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness. *Journal of Neuroscience*, *32*(20), 7082–7090. https://doi.org/10.1523/JNEUROSCI.3769-11.2012.Connectivity

Boly, M., Moran, R., Murphy, M., Boveroux, P., Bruno, M.-A., Noirhomme, Q., … Friston, K. (2012b). Connectivity Changes Underlying Spectral EEG Changes during Propofol-Induced Loss of Consciousness. *Journal of Neuroscience*, *32*(20), 7082–7090. https://doi.org/10.1523/JNEUROSCI.3769-11.2012

Boly, M., Phillips, L., Tshibanda, A. Vanhaudenhuyse, M. Schabus, D., Dang-Vu, T. T., Moonen, G., Hustinx, R., … Laureys, S. (2008). Intrinsic Brain Activity in Altered States of Consciousness: How Conscious Is the Default Mode of Brain Function? *Annals of the New York Academy of Sciences*, *1129*, 119–129. https://doi.org/10.1196/annals.1417.015.Intrinsic

Boly, Mélanie, Faymonville, M. E., Schnakers, C., Peigneux, P., Lambermont, B., Phillips, C., … Laureys, S. (2008). Perception of pain in the minimally

conscious state with PET activation: an observational study. *The Lancet Neurology*, *7*(11), 1013–1020. https://doi.org/10.1016/S1474-4422(08)70219-9

Boly, Melanie, Garrido, M. I., Gosseries, O., Bruno, M.-A., Boveroux, P., Schnakers, C., … Friston, K. (2011a). Response to Comment on "Preserved Feedforward But Impaired Top-Down Processes in the Vegetative State." *Science*, *334*(6060), 1203. https://doi.org/10.1126/science.1202043

Boly, Melanie, Garrido, M. I., Gosseries, O., Bruno, M., Boveroux, P., Schnakers, C., … Friston, K. (2011b). Preserved Feedforward But Impaired Top-Down Processes in the Vegetative State. *Science Reports*, *332*, 858–862.

Boly, Melanie, Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *Journal of Neuroscience*, *37*(40), 9603–9613. https://doi.org/10.1523/JNEUROSCI.3218-16.2017

Boly, Mélanie, Tshibanda, L., Vanhaudenhuyse, A., Noirhomme, Q., Schnakers, C., Ledoux, D., … Laureys, S. (2009). Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. *Human Brain Mapping*, *30*(8), 2393–2400. https://doi.org/10.1002/hbm.20672

Bonhomme, V., Staquet, C., Montupil, J., Defresne, A., Kirsch, M., Martial, C., … Gosseries, O. (2019). General Anesthesia: A Probe to Explore Consciousness. *Frontiers in Systems Neuroscience*, *13*(August). https://doi.org/10.3389/fnsys.2019.00036

Bor, D. (2012). *The ravenous brain: How the new science of consciousness explains our insatiable search for meaning.* New York: Basic Books.

Bor, Daniel, & Seth, A. K. (2012). Consciousness and the prefrontal parietal network: Insights from attention, working memory, and chunking. *Frontiers in Psychology*, *3*(MAR), 1–14. https://doi.org/10.3389/fpsyg.2012.00063

Boveroux, P., Vanhaudenhuyse, A., Bruno, M.-A., Noirhomme, Q., Lauwick, S., Luxen, A., … Boly, M. (2010). Breakdown of within- and between-network Resting State during Propofol-induced Loss of Consciousness. *Anesthesiology*, *113*(5), 1038–1053.

Box, G. E. P. (1976). Science and Statistics. *Journal Ofthe American Statistical Association*, *71*(356), 791–799. https://doi.org/10.1641/B570910

Brain Products GmbH. (2008). BrainVision Analyzer. Gilching, Germany: Brain Products GmbH.

BrainAmp MRplus, & Brain Products GmbH. (n.d.). BrainAmp MRplus. Gilching, Germany: Brain Products GmbH.

BrainCap MR, & Easycap GmbH. (n.d.). BrainCap MR. Gilching, Germany: Easycap GmbH.

Brickner, R. M. (1952). BRAIN OF PATIENT A. AFTER BILATERAL FRONTAL LOBECTOMY; STATUS OF FONTAL-LOBE PROBLEM. *AMA Archives of Neurology & Psychiatry*, *68*(3), 293-313.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Brown, E. N., Lydic, R., & Schiff, N. D. (2010). General Anesthesia, Sleep, and
Coma. *New England Journal of Medicine*, *363*(27), 2638–2650.
https://doi.org/10.1001/jama.306.20.2283

Bruno, M. A., Ledoux, D., Vanhaudenhuyse, A., Gosseries, O., Thibaut, A., &
Laureys, S. (2012). Prognosis of Patients with Altered State of Consciousness.
In *Coma and Disorders of Consciousness* (pp. 11–23). Springer, London.
https://doi.org/10.1007/978-1-4471-2440-5

Bruno, Marie Aurélie, Majerus, S., Boly, M., Vanhaudenhuyse, A., Schnakers, C.,
Gosseries, O., … Laureys, S. (2012). Functional neuroanatomy underlying the
clinical subcategorization of minimally conscious state patients. *Journal of
Neurology*, *259*(6), 1087–1098. https://doi.org/10.1007/s00415-011-6303-7

Bruno, Marie Aurélie, Vanhaudenhuyse, A., Thibaut, A., Moonen, G., & Laureys, S.
(2011). From unresponsive wakefulness to minimally conscious PLUS and
functional locked-in syndromes: Recent advances in our understanding of
disorders of consciousness. *Journal of Neurology*, *258*(7), 1373–1384.
https://doi.org/10.1007/s00415-011-6114-x

Bruns, A. (2004). Fourier-, Hilbert- and wavelet-based signal analysis: Are they
really different approaches? *Journal of Neuroscience Methods*, *137*(2), 321–
332. https://doi.org/10.1016/j.jneumeth.2004.03.002

Buckley, T., Crippen, D., DeWitt, A. L., Fisher, M., Liolios, A., Scheetz, C. L., &
Whetstine, L. M. (2004). Ethics roundtable debate: Withdrawal of tube feeding
in a patient with persistent vegetative state where the patient's wishes are
unclear and there is family dissension. *Critical Care*, *8*(2), 79–84.

https://doi.org/10.1186/cc2451

Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., & Vidal, F. (2015). Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view. *International Journal of Psychophysiology*, *97*(3), 210–220. https://doi.org/10.1016/j.ijpsycho.2015.05.004

Bussche, E. Van den, Hughes, G., Humbeeck, N. Van, & Reynvoet, B. (2010). The relation between consciousness and attention: An empirical study using the priming paradigm. *Consciousness and Cognition*, *19*(1), 86–97. https://doi.org/10.1016/j.concog.2009.12.019

C. W. J. Granger. (1969). Investigating Causal Relations by Economic Models and Cross spectral Methods. *Econometrica*.

Cantero, J. L., Atienza, M., Stickgold, R., Kahana, M. J., Madsen, J. R., & Kocsis, B. (2003). Sleep-Dependent θ Oscillations in the Human Hippocampus and Neocortex. *Journal of Neuroscience*, *23*(34), 10897–10903. https://doi.org/10.1523/jneurosci.23-34-10897.2003

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., … Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, *5*(198). https://doi.org/10.1126/scitranslmed.3006294

Cavinato, M., Genna, C., Manganotti, P., Formaggio, E., Storti, S. F., Campostrini, S., … Piccione, F. (2015). Coherence and Consciousness: Study of Fronto-Parietal Gamma Synchrony in Patients with Disorders of Consciousness. *Brain*

Topography, *28*(4), 570–579. https://doi.org/10.1007/s10548-014-0383-5

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

Chalmers, D. J. (1995). Facing Up to the Problem of Consciousness. *Journal of Conscious Studies*, *2*(3), 200–219. https://doi.org/10.1093/acprof

Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press. https://doi.org/10.1023/A:1008650325798

Changeux, J. P. G. (2012). Conscious processing: Implications for general anesthesia. *Current Opinion in Anaesthesiology*, *25*(4), 397–404. https://doi.org/10.1097/ACO.0b013e32835561de

Chatelle, C., Thibaut, A., Whyte, J., De Val, M. D., Laureys, S., & Schnakers, C. (2014). Pain issues in disorders of consciousness. *Brain Injury*, *28*(9), 1202–1208. https://doi.org/10.3109/02699052.2014.920518

Chen, P., Xie, Q., Wu, X., Huang, H., Lv, W., Chen, L., … Huang, R. (2018). Abnormal Effective Connectivity of the Anterior Forebrain Regions in Disorders of Consciousness. *Neuroscience Bulletin*, *34*, 647–658. https://doi.org/10.1007/s12264-018-0250-6

Chennu, S. (2018). MOHAWK.

Chennu, S., Annen, J., Wannez, S., Thibaut, A., Chatelle, C., Cassol, H., … Laureys, S. (2017). Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. *Brain*, *140*(8), 2120–2132. https://doi.org/10.1093/brain/awx163

Chennu, Srivas, Finoia, P., Kamau, E., Allanson, J., Williams, G. B., Monti, M. M.,

    … Bekinschtein, T. A. (2014). Spectral Signatures of Reorganised Brain

    Networks in Disorders of Consciousness. *PLoS Computational Biology*, *10*(10).

    https://doi.org/10.1371/journal.pcbi.1003887

Chennu, Srivas, O'Connor, S., Adapa, R., Menon, D. K., & Bekinschtein, T. A.

    (2016). Brain Connectivity Dissociates Responsiveness from Drug Exposure

    during Propofol-Induced Transitions of Consciousness. *PLoS Computational*

    Biology, *12*(1), 1–17. https://doi.org/10.1371/journal.pcbi.1004669

Churchland, P. S. (2005). A neurophilosophical slant on consciousness research.

    Progress in Brain Research, *149*, 285–293. https://doi.org/10.1016/S0079-

    6123(05)49020-2

Cimenser, A., Purdon, P. L., Pierce, E. T., Walsh, J. L., Salazar-Gomez, A. F.,

    Harrell, P. G., … Brown, E. N. (2011). Tracking brain states under general

    anesthesia by using global coherence analysis. *Proceedings of the National*

    Academy of Sciences of the United States of America, *108*(21), 8832–8837.

    https://doi.org/10.1073/pnas.1017041108

Claassen, J., Doyle, K., Matory, A., Couch, C., Burger, K. M., Velazquez, A., …

    Rohaut, B. (2019). Detection of Brain Activation in Unresponsive Patients with

    Acute Brain Injury. *New England Journal of Medicine*, *380*(26), 2497–2505.

    https://doi.org/10.1056/nejmoa1812757

Cocchi, L., Zalesky, A., Fornito, A., & Mattingley, J. B. (2013). Dynamic

    cooperation and competition between brain systems during cognitive control.

    Trends in Cognitive Sciences, *17*(10), 493–501.

https://doi.org/10.1016/j.tics.2013.08.006

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, *16*(8), 411–417. https://doi.org/10.1016/j.tics.2012.06.013

Cohen, M. X. (2015). Effects of time lag and frequency matching on phase-based connectivity. *Journal of Neuroscience Methods*, *250*, 137–146. https://doi.org/10.1016/j.jneumeth.2014.09.005

Cohen, M. X. (2017). Where Does EEG Come From and What Does It Mean? *Trends in Neurosciences*, *40*(4), 208–218. https://doi.org/10.1016/j.tins.2017.02.004

Costa, M., Goldberger, A. L., & Peng, C. K. (2002). Multiscale Entropy Analysis of Complex Physiologic Time Series. *Physical Review Letters*, *89*(6), 6–9. https://doi.org/10.1103/PhysRevLett.89.068102

Coulborn, S., Taylor, C., Naci, L., Owen, A. M., & Fernández-Espejo, D. (2021). Disruptions in effective connectivity within and between default mode network and anterior forebrain mesocircuit in prolonged disorders of consciousness. *Brain Sciences*, *11*(6). https://doi.org/10.3390/brainsci11060749

Cowey, A. (2010). The blindsight saga. *Experimental Brain Research*, *200*(1), 3–24. https://doi.org/10.1007/s00221-009-1914-2

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275. Retrieved from http://papers.klab.caltech.edu/22/1/148.pdf

Crone, J. S., Ladurner, G., Höller, Y., Golaszewski, S., Trinka, E., & Kronbichler, M. (2011). Deactivation of the default mode network as a marker of impaired consciousness: An fmri study. *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0026373

Crone, J. S., Lutkenhoff, E. S., Bio, B. J., Laureys, S., & Monti, M. M. (2017). Testing proposed neuronal models of effective connectivity within the cortico-basal gangliathalamo-cortical loop during loss of consciousness. *Cerebral Cortex*, *27*(4), 2727–2738. https://doi.org/10.1093/cercor/bhw112

Crone, J. S., Schurz, M., Höller, Y., Bergmann, J., Monti, M., Schmid, E., … Kronbichler, M. (2015). Impaired consciousness is linked to changes in effective connectivity of the posterior cingulate cortex within the default mode network. *NeuroImage*, *110*, 101–109. https://doi.org/10.1016/j.neuroimage.2015.01.037

Crone, J. S., Soddu, A., Höller, Y., Vanhaudenhuyse, A., Schurz, M., Bergmann, J., … Kronbichler, M. (2014). Altered network properties of the fronto-parietal network and the thalamus in impaired consciousness. *NeuroImage: Clinical*, *4*, 240–248. https://doi.org/10.1016/j.nicl.2013.12.005

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., … Owen, A. M. (2011). Bedside detection of awareness in the vegetative state: A cohort study. *The Lancet*, *378*(9809), 2088–2094. https://doi.org/10.1016/S0140-6736(11)61224-5

Damasio, A. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace.

Daunizeau, J., David, O., & Stephan, K. E. (2011). Dynamic causal modelling: A

    critical review of the biophysical and statistical foundations. *NeuroImage*,

    58(2), 312–322. https://doi.org/10.1016/j.neuroimage.2009.11.062

Daunizeau, Jean, Kiebel, S. J., & Friston, K. J. (2009). Dynamic causal modelling of

    distributed electromagnetic responses. *NeuroImage*, *47*(2), 590–601.

    https://doi.org/10.1016/j.neuroimage.2009.04.062

David, O., Harrison, L., & Friston, K. J. (2005). Modelling event-related responses in

    the brain. *NeuroImage*, *25*(3), 756–770.

    https://doi.org/10.1016/j.neuroimage.2004.12.030

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., & Friston, K. J.

    (2006). Dynamic causal modeling of evoked responses in EEG and MEG.

    NeuroImage, *30*(4), 1255–1272.

    https://doi.org/10.1016/j.neuroimage.2005.10.045

Dehaene, S, & Changeux, J. P. (2004). Neural Mechanisms for Access to

    Consciousness. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (Vol. 3).

    MIT press.

Dehaene, Stanislas, Changeux, J.-P., & Naccache, L. (2011). The Global Neuronal

    Workspace Model of Conscious Access: From Neuronal Architectures to

    Clinical Applications. In S. Dehaene & Y. Christen (Eds.), *Characterizing*

    Consciousness: From Cognition to the Clinic?. Research and Perspectives in

    Neurosciences. (Vol. 18, pp. 55–84). Springer, Berlin, Heidelberg.

    https://doi.org/10.1007/978-3-642-18015-6

Dehaene, Stanislas, & Changeux, J. P. (2011). Experimental and Theoretical
Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227.
https://doi.org/10.1016/j.neuron.2011.03.018

Dehaene, Stanislas, Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a
global workspace in effortful cognitive tasks. *Proceedings of the National
Academy of Sciences*, *95*(24), 14529–14534. https://doi.org/10.1111/j.1749-
6632.2001.tb05714.x

Dehaene, Stanislas, Lau, H., & Kouider, S. (2017). What is consciousness, and could
machines have it? *Science*, *358*(6362), 486–492.
https://doi.org/10.1126/science.aan8871

Dehaene, Stanislas, & Naccache, L. (2001). Towards a cognitive neuroscience of
consciousness: Basic evidence and a workspace framework. *Cognition*, *79*(1–2),
1–37. https://doi.org/10.1016/S0010-0277(00)00123-2

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role
of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132*(9),
2531–2540. https://doi.org/10.1093/brain/awp111

Delorme, A., & Makeig, S. (2004). EEGLAB: an open sorce toolbox for analysis of
single-trail EEG dynamics including independent component anlaysis. *Journal
of Neuroscience Methods*, *134*, 9–21.
https://doi.org/10.1016/j.jneumeth.2003.10.009

Demertzi, A., Racine, E., Bruno, M. A., Ledoux, D., Gosseries, O., Vanhaudenhuyse,
A., … Laureys, S. (2013). Pain perception in disorders of consciousness:

Neuroscience, clinical care, and ethics in dialogue. *Neuroethics*, *6*(1), 37–50.

https://doi.org/10.1007/s12152-011-9149-x

Demertzi, Athena, Antonopoulos, G., Heine, L., Voss, H. U., Crone, J. S., De Los

Angeles, C., … Laureys, S. (2015). Intrinsic functional connectivity

differentiates minimally conscious from unresponsive patients. *Brain*, *138*(9),

2619–2631. https://doi.org/10.1093/brain/awv169

Deng, F., Cusack, R., & Naci, L. (2019). Brain Functional Connectivity in

Wakefulness Predicts Susceptibility to Anaesthesia. *Conference on Cognitive

Computational Neuroscience*, 1075–1078.

https://doi.org/10.32470/ccn.2019.1294-0

Dennett, D. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (Eds.),

Consciousness in modern science (pp. 42–77). Oxford University Press.

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

consciousness*. *The New York state dental journal* (Vol. 79). Cambridge,

Massachusetts: MIT Press.

Dennett, D. C. (2014). Why and How Does Consciousness Seem the Way it Seems?

Open MIND, *10*, 1–11. https://doi.org/10.15502/9783958570245

Di Perri, C., Bahri, M. A., Amico, E., Thibaut, A., Heine, L., Antonopoulos, G., …

Laureys, S. (2016). Neural correlates of consciousness in patients who have

emerged from a minimally conscious state: A cross-sectional multimodal

imaging study. *The Lancet Neurology*, *15*(8), 830–842.

https://doi.org/10.1016/S1474-4422(16)00111-3

Driesen, N. R., McCarthy, G., Bhagwagar, Z., Bloch, M., Calhoun, V., D'Souza, D.
C., … Krystal, J. H. (2013). Relationship of resting brain hyperconnectivity and
schizophrenia-like symptoms produced by the NMDA receptor antagonist
ketamine in humans. *Molecular Psychiatry*, *18*(11), 1199–1204.
https://doi.org/10.1038/mp.2012.194

Dux, P. E., & Renťmarois. (2009). The attentional blink: A review of data and theory.
Attention, Perception, and Psychophysics, *71*(8), 1683–1700.
https://doi.org/10.3758/APP.71.8.1683

Edlow, B. L. (2018). Covert Consciousness: Searching for Volitional Brain Activity
in the Unresponsive. *Current Biology*, *28*(23), R1345–R1348.
https://doi.org/10.1016/j.cub.2018.10.022

Edlow, B. L., Claassen, J., Schiff, N. D., & Greer, D. M. (2021). Recovery from
disorders of consciousness: mechanisms, prognosis and emerging therapies.
Nature Reviews Neurology, *17*(3), 135–156. https://doi.org/10.1038/s41582-
020-00428-x

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic Predictions: Oscillations and
Synchrony in Top–Down Processing. *Nature Reviews Neuroscience*, *2*(10),
704–716.

Engemann, D. A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F.,
… Sitt, J. D. (2018). Robust EEG-based cross-site and cross-protocol
classification of states of consciousness. *Brain*, 3179–3192.
https://doi.org/10.1093/brain/awy251

Esaki, R. K., & Mashour, G. A. (2009). Levels of consciousness during regional

    anesthesia and monitored anesthesia care: Patient expectations and experiences.

    Anesthesia and Analgesia, *108*(5), 1560–1563.

    https://doi.org/10.1213/ane.0b013e31819c2aa3

Fastenrath, M., Friston, K. J., & Kiebel, S. J. (2009). Dynamical causal modelling for

    M/EEG: Spatial and temporal symmetry constraints. *NeuroImage*, *44*(1), 154–

    163. https://doi.org/10.1016/j.neuroimage.2008.07.041

Fernández-Espejo, D., Soddu, A., Cruse, D., Palacios, E. M., Junque, C.,

    Vanhaudenhuyse, A., … Owen, A. M. (2012). A role for the default mode

    network in the bases of disorders of consciousness. *Annals of Neurology*, *72*(3),

    335–343. https://doi.org/10.1002/ana.23635

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious

    processing: A comparison of face and nonface stimuli. *Psychological Science*,

    20(1), 42–51. https://doi.org/10.1111/j.1467-9280.2008.02256.x

Fisch, B. J., & Spehlmann, R. (1991). *Spehlmann's EEG primer*. Amsterdam:

    Elsevier.

Fiset, P., Paus, T., Daloze, T., Plourde, G., Meuret, P., Bonhomme, V., … Evans, A.

    C. (1999). Brain mechanisms of propofol-induced loss of consciousness in

    humans: A positron emission tomographic study. *Journal of Neuroscience*,

    19(13), 5506–5513. https://doi.org/10.1523/jneurosci.19-13-05506.1999

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific

    impairment in metacognitive accuracy following anterior prefrontal lesions.

Brain, *137*(10), 2811–2822. https://doi.org/10.1093/brain/awu221

Fridman, E. A., Beattie, B. J., Broft, A., Laureys, S., & Schif, N. D. (2014). Regional cerebral metabolic patterns demonstrate the role of anterior forebrain mesocircuit dysfunction in the severely injured brain. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(17), 6473–6478. https://doi.org/10.1073/pnas.1320969111

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Frigato, G. (2021). The Neural Correlates of Access Consciousness and Phenomenal Consciousness Seem to Coincide and Would Correspond to a Memory Center, an Activation Center and Eight Parallel Convergence Centers. *Frontiers in Psychology*, *12*(September). https://doi.org/10.3389/fpsyg.2021.749610

Friston, K. (2009). Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, *7*(2), 0220–0225. https://doi.org/10.1371/journal.pbio.1000033

Friston, K. J., Bastos, A., Litvak, V., Stephan, K. E., Fries, P., & Moran, R. J. (2012). DCM for complex-valued data: Cross-spectra, coherence and phase-delays. *NeuroImage*, *59*(1), 439–455. https://doi.org/10.1016/j.neuroimage.2011.07.048

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., & Penny, W. (2007). Variational free energy and the Laplace approximation. *NeuroImage*, *34*(1), 220–234. https://doi.org/10.1016/j.neuroimage.2006.08.035

Friston, K., & Penny, W. (2011). Post hoc Bayesian model selection. *NeuroImage*, *56*(4), 2089–2099. https://doi.org/10.1016/j.neuroimage.2011.03.062

Friston, Karl J. (2011). Functional and Effective Connectivity: A Review. *Brain Connectivity*, *1*(1), 13–36. https://doi.org/10.1089/brain.2011.0008

Friston, Karl J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., Van Wijk, B. C. M., … Zeidman, P. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage*, *128*, 413–431. https://doi.org/10.1016/j.neuroimage.2015.11.015

Fulton, J. F. (1949). Functional Localization in Relation to Frontal Lobotomy. *Oxford University Press*, 2015.

Gambús, P. L., & Trocõniz, I. F. (2015). Pharmacokinetic-pharmacodynamic modelling in anaesthesia. *British Journal of Clinical Pharmacology*, *79*(1), 72–84. https://doi.org/10.1111/bcp.12286

Gaskell, A. L., Hight, D. F., Winders, J., Tran, G., Defresne, A., Bonhomme, V., … Sanders, R. D. (2017). Frontal alpha-delta EEG does not preclude volitional response during anaesthesia: Prospective cohort study of the isolated forearm technique. *British Journal of Anaesthesia*, *119*(4), 664–673. https://doi.org/10.1093/bja/aex170

Gazzaniga, M. S., Bogen, J. E., & Sperry, R. W. (1962). Some functional effects of

sectioning the cerebral commissures in man. *Proceedings of the National Academy of Sciences*, *48*(10), 1765–1769.

Giacino, J T, Ashwal, S., Childs, N., Cranford, R., Jennett, B., & Katz, D. I. (2002). The minimally conscious state. *Neurology*, *58*(3), 349–353.

Giacino, Joseph T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., … Zasler, N. D. (2002). The minimally conscious state: Definition and diagnostic criteria. *Neurology*, *58*(3), 349–353. https://doi.org/10.1212/WNL.58.3.349

Giacino, Joseph T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: The state of the science. *Nature Reviews Neurology*, *10*(2), 99–114. https://doi.org/10.1038/nrneurol.2013.279

Giacino, Joseph T., Kalmar, K., & Whyte, J. (2004). The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility. *Archives of Physical Medicine and Rehabilitation*, *85*(12), 2020–2029. https://doi.org/10.1016/j.apmr.2004.02.033

Giacino, Joseph T., Schnakers, C., Rodriguez-Moreno, D., Kalmar, K., Schiff, N., & Hirsch, J. (2009). *Behavioral assessment in patients with disorders of consciousness: gold standard or fool's gold? Progress in Brain Research* (Vol. 177). Elsevier. https://doi.org/10.1016/S0079-6123(09)17704-X

Gibbs, F. ., Gibbs, E. ., & Lennox, W. G. (1937). EFFECT ON THE ELECTRO-ENCEPHALOGRAM OF CERTAIN DRUGS WHICH INFLUENCE NERVOUS ACTIVITY. *Archives of Internal Medicine*, *60*(1), 154–166.

Glover, G. H. (2011). Overview of functional magnetic resonance imaging.

Neurosurgery Clinics of North America, *22*(2), 133–139.

https://doi.org/10.1016/j.nec.2010.11.001

Goddard, N., & Smith, D. (2013). Unintended awareness and monitoring of depth of anaesthesia. *Continuing Education in Anaesthesia, Critical Care and Pain*, *13*(6), 213–217. https://doi.org/10.1093/bjaceaccp/mkt016

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C. K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(SUPPL. 1), 2466–2472. https://doi.org/10.1073/pnas.012579499

Goldenholz, D. M., Ahlfors, S. P., Hämäläinen, M. S., Sharon, D., Ishitobi, M., Vaina, L. M., & Stufflebeam, S. M. (2009). Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human Brain Mapping*, *30*(4), 1077–1086. https://doi.org/10.1002/hbm.20571

Gosseries, O, Schnakers, C., Ledoux, D., Vanhaudenhuyse, A., Bruno, M. A., Demertzi, A., … Laureys, S. (2011). Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. *Functional Neurology*, *26*(1), 25–30.

Gosseries, Olivia, Di, H., Laureys, S., & Boly, M. (2014). Measuring Consciousness in Severely Damaged Brains. *Annual Review of Neuroscience*, *37*(1), 457–478. https://doi.org/10.1146/annurev-neuro-062012-170339

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in

cat visual cortex exhibit inter-columnar synchronization which reflects global

stimulus properties. *Nature*. https://doi.org/10.1038/338334a0

Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., …

Vanrumste, B. (2008). Review on solving the inverse problem in EEG source

analysis. *Journal of NeuroEngineering and Rehabilitation*, *5*, 1–33.

https://doi.org/10.1186/1743-0003-5-25

Greicius, M. D., Kiviniemi, V., Tervonen, O., Vainionpää, V., Reiss, A. L., &

Menon, V. (2008). Persistent Default- Mode Network Connectivity During

Light Sedation. *Human Brain Mapping*, *29*(7), 839–847.

https://doi.org/10.1002/hbm.20537.Persistent

Guimera, R., & Amaral, L. (2005). Functional catography of complex metabolic

networks. *Nature*, *433*(7028), 895–900.

Guldenmund, J. P., Vanhaudenhuyse, A., Boly, M., Laureys, S., & Soddu, A. (2012).

A default mode of brain function in altered states of consciousness. *Archives

Italiennes de Biologie*, *150*, 107–121. https://doi.org/10.4449/aib.v150i2.1373

Guldenmund, P., Demertzi, A., Boveroux, P., Boly, M., Vanhaudenhuyse, A., Bruno,

M. A., … Soddu, A. (2013). Thalamus, Brainstem and Salience Network

Connectivity Changes During Propofol-Induced Sedation and Unconsciousness.

Brain Connectivity, *3*(3), 273–285. https://doi.org/10.1089/brain.2012.0117

Guldenmund, P., Gantner, I. S., Baquero, K., Das, T., Demertzi, A., Boveroux, P., …

Soddu, A. (2016). Propofol-Induced Frontal Cortex Disconnection: A Study of

Resting-State Networks, Total Brain Connectivity, and Mean BOLD Signal

Oscillation Frequencies. *Brain Connectivity*, *6*(3), 225–237. https://doi.org/10.1089/brain.2015.0369

Guldenmund, P., Soddu, A., Baquero, K., Vanhaudenhuyse, A., Bruno, M. A., Gosseries, O., … Gómez, F. (2016). Structural brain injury in patients with disorders of consciousness: A voxel-based morphometry study. *Brain Injury*, *30*(3), 343–352. https://doi.org/10.3109/02699052.2015.1118765

Gurland, J., & Tripathi, R. C. (1971). A simple approximation for unbiased estimation of the standard deviation. *American Statistician*, *25*(4), 30–32. https://doi.org/10.1080/00031305.1971.10477279

Gusnard, D. a, & Raichle, M. E. (2001). Reviews Searching for a Baseline : Functional Imaging and the. *Neuroscience*, *2*(October), 685–694. https://doi.org/10.1038/35094500

Harmony, T. (2013). The functional significance of delta oscillations in cognitive processing. *Frontiers in Integrative Neuroscience*, *7*(DEC), 1–10. https://doi.org/10.3389/fnint.2013.00083

Hebb, D. O., & Penfield, W. (1940). Human behavior after extensive bilateral removal from the frontal lobes. *Arch. Neurol. Psychiatry*, (c), 73–87.

Heine, L., Soddu, A., Gómez, F., Vanhaudenhuyse, A., Tshibanda, L., Thonnard, M., … Demertzi, A. (2012). Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states. *Frontiers in Psychology*, *3*(AUG), 1–12. https://doi.org/10.3389/fpsyg.2012.00295

258

Hohwy, J. (2009). The neural correlates of consciousness: New experimental approaches needed? *Consciousness and Cognition*, *18*(2), 428–438. https://doi.org/10.1016/j.concog.2009.02.006

Hong, J. Y., Jee, Y. S., & Luthardt, F. W. (2005). Comparison of conscious sedation for oocyte retrieval between low-anxiety and high-anxiety patients. *Journal of Clinical Anesthesia*, *17*(7), 549–553. https://doi.org/10.1016/j.jclinane.2005.01.008

Hornero, R., Abásolo, D., Escudero, J., & Gómez, C. (2009). Non-linear analysis of EEG and MEG in patients with Alzheimer's disease. *Royal Soc Philos T R Soc A*, *367*(1887), 317–336.

Horovitz, S. G., Braun, A. R., Carr, W. S., Picchioni, D., Balkin, T. J., Fukunaga, M., & Duyn, J. H. (2009). Decoupling of the brain's default mode network during deep sleep. *Proceedings of the National Academy of Sciences*, *106*(27), 11376–11381. https://doi.org/10.1073/pnas.0901435106

Hutt, A., & Longtin, A. (2010). Effects of the anesthetic agent propofol on neural populations. *Cognitive Neurodynamics*, *4*(1), 37–59. https://doi.org/10.1007/s11571-009-9092-2

Ihalainen, R., Gosseries, O., de Steen, F. Van, Raimondo, F., Panda, R., Bonhomme, V., … Chennu, S. (2021). How hot is the hot zone? Computational modelling clarifies the role of parietal and frontoparietal connectivity during anaesthetic-induced loss of consciousness. *NeuroImage*, *231*(January), 117841. https://doi.org/10.1016/j.neuroimage.2021.117841

Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, *32*(127), 127–136.

Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, *73*(4), 357–366. https://doi.org/10.1007/BF00199471

Jennett, B., & Plum, F. (1972). Persistent vegetative state after brain damage. *The Lancet*, *299*(7753), 734–737. https://doi.org/10.1016/s0140-6736(72)90242-5

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 529–546. https://doi.org/10.1002/wcs.130

Jevtovic-Todorovic, V. (2016). General Anesthetics and Neurotoxicity: How Much Do We Know? *Anesthesiology Clinics*, *34*(3), 439–451. https://doi.org/10.1016/j.anclin.2016.04.001

Kalmar, K., & Giacino, J. T. (2005). The JFK Coma Recovery Scale - Revised. *Neuropsychological Rehabilitation*, *15*(3–4), 454–460. https://doi.org/10.1080/09602010443000425

Karakaş, S. (2020). A review of theta oscillation and its functional correlates. *International Journal of Psychophysiology*, *157*, 82–99. https://doi.org/10.1016/j.ijpsycho.2020.04.008

Kaskinoro, K., Maksimow, A., Långsjö, J., Aantaa, R., Jääskeläinen, S., Kaisti, K., … Scheinin, H. (2011). Wide inter-individual variability of bispectral index and spectral entropy at loss of consciousness during increasing concentrations of

dexmedetomidine, propofol, and sevoflurane. *British Journal of Anaesthesia*, *107*(4), 573–580. https://doi.org/10.1093/bja/aer196

Kassubek, J., Juengling, F. D., Els, T., Spreer, J., Herpers, M., Krause, T., … Lücking, C. H. (2003). Activation of a residual cortical network during painful stimulation in long-term postanoxic vegetative state: A 15O-H2O PET study. *Journal of the Neurological Sciences*, *212*(1–2), 85–91. https://doi.org/10.1016/S0022-510X(03)00106-0

Kentridge, R. W., Nijboer, T. C. W., & Heywood, C. A. (2008). Attended but unseen : visual attention is not sufficient for visual awareness. *Neuropsychologia*, *46*(3), 864–869. Retrieved from http://dx.doi.org/10.1037/xge0000076

Ki, S., Kim, K. M., Lee, Y. H., Bang, J. Y., Choi, B. M., & Noh, G. J. (2019). Phase lag entropy as a hypnotic depth indicator during propofol sedation. *Anaesthesia*, *74*(8), 1033–1040. https://doi.org/10.1111/anae.14704

Kiebel, S. J., Garrido, M. I., Moran, R., Chen, C.-C., & Friston, K. J. (2009). Dynamic causal modeling for EEG and MEG. *Human Brain Mapping*, *30*(6), 1866–1876. https://doi.org/10.1002/hbm.20775

Kiebel, S. J., Garrido, M. I., Moran, R. J., & Friston, K. J. (2008). Dynamic causal modelling for EEG and MEG. *Cognitive Neurodynamics*, *2*(2), 121–136. https://doi.org/10.1007/s11571-008-9038-0

Kil, H. K., Kim, W. O., Chung, W. Y., Kim, G. H., Seo, H., & Hong, J. Y. (2012). Preoperative anxiety and pain sensitivity are independent predictors of propofol

and sevoflurane requirements in general anaesthesia. *British Journal of Anaesthesia*, *108*(1), 119–125. https://doi.org/10.1093/bja/aer305

Kim, H., Hudetz, A. G., Lee, J., Mashour, G. A., Lee, U., Group, the R. S., … Vlisides, P. E. (2018). Estimating the Integrated Information Measure Phi from High-Density Electroencephalography during States of Consciousness in Humans. *Frontiers in Human Neuroscience*, *12*(February), 42. https://doi.org/10.3389/fnhum.2018.00042

Kim, M., Mashour, G. A., Moraes, S. B., Vanini, G., Tamal, V., Janke, E., … Lee, U. (2016). Functional and topological conditions for explosive synchronization develop in human brain networks with the onset of anesthetic-induced unconsciousness. *Frontiers in Computational Neuroscience*, *10*(JAN), 1–15. https://doi.org/10.3389/fncom.2016.00001

Kim, S.-P. (2018). Preprocessing of EEG. In C.-H. Im (Ed.), *Computational EEG Analysis* (pp. 15–33). Springer Singapore.

King, J.-R., Bekinschtein, T., & Dahaene, S. (2011). Comment on "Preserved Feedforward But Impaired Top-Down Processes in the Vegetative State." *Science*, *334*(6060), 1203. https://doi.org/10.1126/science.1202043

King, J. R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., … Dehaene, S. (2013). Information sharing in the brain indexes consciousness in noncommunicative patients. *Current Biology*, *23*(19), 1914–1919. https://doi.org/10.1016/j.cub.2013.07.075

Klein, C. (2014). The brain at rest: What it is doing and why that matters. *Philosophy*

of Science, *81*(5), 974–985. https://doi.org/10.1086/677692

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, *29*(2–3), 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Koch, C. (2004). *The quest for consciousness: A neurobiological approach* (18th ed.). Denver, Colorado: Roberts and Company.

Koch, C. (2019). *The feeling of life itself: Why consciousness is widespread but can't be computed*. Cambridge, Massachusetts: MIT Press.

Koch, Christof. (2018). What Is Consciousness? *Scientific American*, *318*(6), 60–64. https://doi.org/10.1038/scientificamerican0618-60

Koch, Christof, Massimini, M., Boly, M., & Tononi, G. (2016a). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

Koch, Christof, Massimini, M., Boly, M., & Tononi, G. (2016b). Posterior and anterior cortex — where is the difference that makes the difference? *Nature Reviews Neuroscience*, *17*(10), 666–666. https://doi.org/10.1038/nrn.2016.105

Koch, Christof, & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012

Krishnaswamy, P., Obregon-Henao, G., Ahveninen, J., Khan, S., Babadi, B., Iglesias, J. E., … Purdon, P. L. (2017). Sparsity Enables Estimation of both Subcortical and Cortical Activity from MEG and EEG.

https://doi.org/10.1073/pnas.1705414114

Ku, S. W., Lee, U. C., Noh, G. J., Jun, I. G., & Mashour, G. A. (2011). Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. *PLoS ONE*, *6*(10), 1–9. https://doi.org/10.1371/journal.pone.0025155

Kuehlmeyer, K., Palmour, N., Riopelle, R. J., Bernat, J. L., Jox, R. J., & Racine, E. (2014). Physicians' attitudes toward medical and ethical challenges for patients in the vegetative state: Comparing Canadian and German perspectives in a vignette survey. *BMC Neurology*, *14*(1). https://doi.org/10.1186/1471-2377-14-119

Laalou, F. Z., Egard, M., Guillot, M., Noll, E., Taglang, G., & Pain, L. (2010). Influence of preoperative cognitive status on propofol requirement to maintain hypnosis in the elderly. *British Journal of Anaesthesia*, *105*(3), 342–346. https://doi.org/10.1093/bja/aeq160

Lachaux, J.-P., Rodriguez, E., Martinerie, J., & Varela, F. (1999). Measuring Phase Synchrony in Brain Signals. *Human Brain Mapping*, *8*, 194–208. https://doi.org/10.1162/jocn

Laureys, S., Faymonville, M. E., & Lamy, M. (2000). Cerebral Function in Vegetative State studied by Positron Emission Tomography. In J. L. Vincent (Ed.), *Yearbook of Intensive Care and Emergency Medicine 2000. Yearbook of Intensive Care and Emergency Medicine* (pp. 587–597). Springer Berlin Heidelberg.

Laureys, S., Faymonville, M. E., Peigneux, P., Damas, P., Lambermont, B., Del

    Fiore, G., … Maquet, P. (2002). Cortical Processing of Noxious Somatosensory

    Stimuli in the Persistent Vegetative State. *NeuroImage*, *17*(2), 732–741.

    https://doi.org/10.1006/nimg.2002.1236

Laureys, S., Goldman, S., Phillips, C., Van Bogaert, P., Aerts, J., Luxen, A., …

    Maquet, P. (1999). Impaired effective cortical connectivity in vegetative state:

    Preliminary investigation using PET. *NeuroImage*, *9*(4), 377–382.

    https://doi.org/10.1006/nimg.1998.0414

Laureys, Steven. (2005). The neural correlate of (un)awareness: Lessons from the

    vegetative state. *Trends in Cognitive Sciences*, *9*(12), 556–559.

    https://doi.org/10.1016/j.tics.2005.10.010

Laureys, Steven, Celesia, G. G., Cohadon, F., Lavrijsen, J., León-Carrión, J., Sannita,

    W. G., … Dolce, G. (2010). Unresponsive wakefulness syndrome: A new name

    for the vegetative state or apallic syndrome. *BMC Medicine*, *8*, 2–5.

    https://doi.org/10.1186/1741-7015-8-68

Laureys, Steven, Owen, A. M., & Schiff, N. D. (2004a). Brain function in brain

    death, coma, vegetative state, minimally conscious state and locked-in

    syndrome. *Lancet Neurology*, *3*(9), 537–546.

Laureys, Steven, Owen, A. M., & Schiff, N. D. (2004b). Brain function in coma,

    vegetative state, and related disorders. *Lancet Neurology*, *3*(9), 537–546.

    https://doi.org/10.1016/S1474-4422(04)00852-X

Laureys, Steven, & Schiff, N. D. (2012). Coma and consciousness: Paradigms

(re)framed by neuroimaging. *NeuroImage*, *61*(2), 478–491.

https://doi.org/10.1016/j.neuroimage.2011.12.041

Lechinger, J., Chwala-Schlegel, N., Fellinger, R., Donis, J., Michitsch, G., Pichler,

G., & Schabus, M. (2013). Mirroring of a simple motor behavior in Disorders of

Consciousness. *Clinical Neurophysiology*, *124*(1), 27–34.

https://doi.org/10.1016/j.clinph.2012.05.016

Lee, H., Mashour, G. A., Noh, G. J., Kim, S., & Lee, U. C. (2013). Reconfiguration

of network hub structure after propofol-induced unconsciousness.

Anesthesiology, *119*(6), 1347–1359.

Lee, M., Baird, B., Gosseries, O., Nieminen, J. O., Boly, M., Postle, B. R., … Lee, S.

W. (2019). Connectivity differences between consciousness and

unconsciousness in non-rapid eye movement sleep: a TMS–EEG study.

Scientific Reports, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-41274-2

Lee, M., Sanders, R. D., Yeom, S. K., Won, D. O., Seo, K. S., Kim, H. J., … Lee, S.

W. (2017). Network Properties in Transitions of Consciousness during

Propofol-induced Sedation. *Scientific Reports*, *7*(1), 1–13.

https://doi.org/10.1038/s41598-017-15082-5

Lee, UnCheol, Kim, S., Noh, G. J., Choi, B. M., Hwang, E., & Mashour, G. A.

(2009). The directionality and functional organization of frontoparietal

connectivity during consciousness and anesthesia in humans. *Consciousness

and Cognition*, *18*(4), 1069–1078. https://doi.org/10.1016/j.concog.2009.04.004

Lee, Uncheol, Ku, S., Noh, G., Baek, S., Choi, B., & Mashour, G. A. (2015).

Disruption of Frontal-Parietal Communication by Ketamine, Propofol, and

Sevoflurane. *Anesthesiology*, *118*(6), 1264–1275.

https://doi.org/10.1097/ALN.0b013e31829103f5.Disruption

Leech, R., Braga, R., & Sharp, D. J. (2012). Echoes of the brain within the posterior

cingulate cortex. *Journal of Neuroscience*, *32*(1), 215–222.

https://doi.org/10.1523/JNEUROSCI.3689-11.2012

Lehembre, R., Bruno, M.-A., Vanhaudenhuyse, A., Chatelle, C., Cologan, V.,

Leclercq, Y., … Noirhomme, Q. (2012). Resting-state EEG study of comatose

patients: a connectivity and frequency analysis to find differences between

vegetative and minimally conscious states. *Functional Neurology*, *27*(1), 41–47.

https://doi.org/10.13838/j.cnki.kjgc.2014.02.041

Leslie, K. (2017). Awareness and dreaming during TIVA. In A. R. Absalom & K. P.

Mason (Eds.), *Total Intravenous Anesthesia and Target Controlled Infusions*

(pp. 783–796). Springer, Cham. https://doi.org/10.1213/ane.0000000000002666

Leslie, Kate, Sleigh, J., Paech, M. J., Voss, L., Lim, C. W., & Sleigh, C. (2009).

Dreaming and electroencephalographic changes during anesthesia maintained

with propofol or desflurane. *Anesthesiology*, *111*(3), 547–555.

https://doi.org/10.1097/ALN.0b013e3181adf768

Li, L., Looney, D., Park, C., Tanaka, T., Cao, J., & Mandic, D. P. (2012). Phase-

based brain consciousness analysis. *Proceedings of the Annual International

Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*,

1032–1035. https://doi.org/10.1109/EMBC.2012.6346110

Lin, P., Yang, Y., Gao, J., De Pisapia, N., Ge, S., Wang, X., … Niu, C. (2017).
Dynamic Default Mode Network across Different Brain States. *Scientific
Reports*, *7*(March), 1–13. https://doi.org/10.1038/srep46088

Litvak, V., Jafarian, A., Zeidman, P., Tibon, R., Henson, R. N., & Friston, K. (2019).
There's no such thing as a "true" model: The challenge of assessing face
validity. *Conference Proceedings - IEEE International Conference on Systems,
Man and Cybernetics*, *2019-Octob*, 4403–4408.
https://doi.org/10.1109/SMC.2019.8914255

Liu, R. S. N., Lemieux, L., Bell, G. S., Sisodiya, S. M., Shorvon, S. D., Sander, J. W.
A. S., & Duncan, J. S. (2003). A longitudinal study of brain morphometrics
using quantitative magnetic resonance imaging and difference image analysis.
NeuroImage, *20*(1), 22–33. https://doi.org/10.1016/S1053-8119(03)00219-2

Liu, X., Lauer, K., Ward, D., Li, S.-J., & Hudetz, A. (2013). Differential Effects of
Deep Sedation with Propofol on the Specific and Nonspecific Thalamocortical
Systems. *Anesthesiology*, *118*(1), 59–69. https://doi.org/10.1038/jid.2014.371

Luck, S.J. (2014). *An introduction to the event-related potential technique.* MIT
press.

Luck, Steven J., & Kappenman, E. S. (Eds.). (2011). *The Oxford Handbook of Event-
Related Potential Components*. Oxford University Press.
https://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Luppi, A. I., Cain, J., Spindler, L. R. B., Górska, U. J., Toker, D., Hudson, A. E., …
Boly, M. (2021). Mechanisms Underlying Disorders of Consciousness:

Bridging Gaps to Move Toward an Integrated Translational Science. *Neurocritical Care*, *35*(s1), 37–54. https://doi.org/10.1007/s12028-021-01281-6

Luppi, A. I., Craig, M. M., Pappas, I., Finoia, P., Williams, G. B., Allanson, J., … Stamakis, E. A. (2019). Consciousness-specific dynamic interactions of brain integration and functional diversity. *Nature Communications*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-12658-9

Lutkenhoff, E. S., Johnson, M. A., Casarotto, S., Massimini, M., & Monti, M. M. (2020). Subcortical atrophy correlates with the perturbational complexity index in patients with disorders of consciousness. *Brain Stimulation*, *13*(5), 1426–1435. https://doi.org/10.1016/j.brs.2020.07.012

MacDonald, A. A., Naci, L., MacDonald, P. A., & Owen, A. M. (2015). Anesthesia and neuroimaging: investigating the neural correlates of unconsciousness. *Trends in Cognitive Sciences*, *19*(2), 100–107. https://doi.org/10.1016/J.TICS.2014.12.005

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. In R. D. Wright (Ed.), *Visual Attention* (pp. 55–76). New York: Oxford University Press.

Mack, Arien. (2003). Inattentional Blindness: Looking Without Seeing. *Current Directions in Psychological Science*, *12*(5), 180–184. https://doi.org/10.1111/1467-8721.01256

Maksimow, A., Silfverhuth, M., Långsjö, J., Kaskinoro, K., Georgiadis, S., Jääskeläinen, S., & Scheinin, H. (2014). Directional connectivity between

frontal and posterior brain regions is altered with increasing concentrations of propofol. *PLoS ONE*, *9*(11), 1–16. https://doi.org/10.1371/journal.pone.0113616

Malekmohammadi, M., AuYong, N., Price, C. M., Tsolaki, E., Hudson, A. E., & Pouratian, N. (2018). Propofol-induced Changes in α-β Sensorimotor Cortical Connectivity. *Anesthesiology*, *128*(2), 305–316. https://doi.org/10.1097/ALN.0000000000001940

Malekmohammadi, M., Price, C. M., Hudson, A. E., DiCesare, J. A. T., & Pouratian, N. (2019). Propofol-induced loss of consciousness is associated with a decrease in thalamocortical connectivity in humans. *Brain*, *142*(8), 2288–2302. https://doi.org/10.1093/brain/awz169

Mashour, G.A. (2004). Consciousness unbound. *Anesthesiology*, *100*(2), 428–433. https://doi.org/10.1097/00000542-200402000-00035

Mashour, George A., Roelfsema, P., Changeux, J. P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of Cortical Effective Connectivity During Sleep. *Science Reports*, *309*(September), 2228–2233.

Mhuircheartaigh, R. N., Rosenorn-Lanng, D., Wise, R., Jbabdi, S., Rogers, R., & Tracey, I. (2010). Cortical and subcortical connectivity changes during decreasing levels of consciousness in humans: A functional magnetic resonance imaging study using propofol. *Journal of Neuroscience*, *30*(27), 9095–9102.

https://doi.org/10.1523/JNEUROSCI.5516-09.2010

Mhuircheartaigh, R. N., Warnaby, C., Rogers, R., Jbabdi, S., & Tracey, I. (2013).
Slow-wave activity saturation and thalamocortical isolation during propofol
anesthesia in humans. *Science Translational Medicine*, *5*(208).
https://doi.org/10.1126/scitranslmed.3006007

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave De
Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*(10),
2195–2222. https://doi.org/10.1016/j.clinph.2004.06.001

Michel, M. (2019). The mismeasure of consciousness: A problem of coordination for
the perceptual awareness scale. *Philosophy of Science*, *86*(5), 1239–1249.
https://doi.org/10.1086/705509

Monti, M. M., Lutkenhoff, E. S., Rubinov, M., Boveroux, P., Vanhaudenhuyse, A.,
Gosseries, O., … Laureys, S. (2013). Dynamic Change of Global and Local
Information Processing in Propofol-Induced Loss and Recovery of
Consciousness. *PLoS Computational Biology*, *9*(10).
https://doi.org/10.1371/journal.pcbi.1003271

Monti, M. M., Vanhaudenhuyse, A., Coleman, M., Boly, M., Pickard, J. D.,
Tshibanda, L., … Laureys, S. (2010). Willful modulation of brain activity in
disorders of consciousness. *New England Journal of Medicine*.
https://doi.org/10.1056/NEJMc1003229

Moran, R. J., Kiebel, S. J., Stephan, K. E., Reilly, R. B., Daunizeau, J., & Friston, K.
J. (2007). A neural mass model of spectral responses in electrophysiology.

271

NeuroImage, *37*(3), 706–720. https://doi.org/10.1016/j.neuroimage.2007.05.032

Moran, R. J., Stephan, K. E., Seidenbecher, T., Pape, H. C., Dolan, R. J., & Friston, K. J. (2009). Dynamic causal models of steady-state responses. *NeuroImage*, *44*(3), 796–811. https://doi.org/10.1016/j.neuroimage.2008.09.048

Moran, R., Pinotsis, D. A., & Friston, K. (2013). Neural masses and fields in dynamic causal modelling. *Frontiers in Computational Neuroscience*, *7*(APR 2013), 1–12. https://doi.org/10.3389/fncom.2013.00057

Murphy, M., Bruno, M. A., Riedner, B. A., Boveroux, P., Noirhomme, Q., Landsness, E. C., … Boly, M. (2011a). Propofol anesthesia and sleep: A high-density EEG study. *Sleep*, *34*(3). https://doi.org/10.1093/sleep/34.3.283

Murphy, M., Bruno, M. A., Riedner, B. A., Boveroux, P., Noirhomme, Q., Landsness, E. C., … Boly, M. (2011b). Propofol anesthesia and sleep: A high-density EEG study. *Sleep*, *34*(3). https://doi.org/10.1093/sleep/34.3.283

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435. https://doi.org/10.2307/2183914

Nakayama, N., Okumura, A., Shinoda, J., Nakashima, T., & Iwama, T. (2006). Relationship between regional cerebral metabolism and consciousness disturbance in traumatic diffuse brain injury without large focal lesions: An FDG-PET study with statistical parametric mapping analysis. *Journal of Neurology, Neurosurgery and Psychiatry*, *77*(7), 856–862. https://doi.org/10.1136/jnnp.2005.080523

Newcombe, V., Williams, G., Scoffings, D., Cross, J., Carpenter, T. A., Pickard, J.,

& Menon, D. (2010). Aetiological differences in neuroanatomy of the vegetative state: Insights from diffusion tensor imaging and functional implications. *Journal of Neurology, Neurosurgery and Psychiatry*, *81*(5), 552–561. https://doi.org/10.1136/jnnp.2010.205815

Nicolaou, N., Hourris, S., Alexandrou, P., & Georgiou, J. (2012). EEG-based automatic classification of "awake" versus "anesthetized" state in general anesthesia using granger causality. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0033869

Nir, Y., & Tononi, G. (2010). Dreaming and the brain: from phenomenology to neurophysiology. *Trends in Cognitive Sciences*, *14*(2), 88–100. https://doi.org/10.1016/j.tics.2009.12.001

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., & Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology*, *115*(10), 2292–2307. https://doi.org/10.1016/j.clinph.2004.04.029

Noreika, V., Jylhänkangas, L., Móró, L., Valli, K., Kaskinoro, K., Aantaa, R., … Revonsuo, A. (2011). Consciousness lost and found: Subjective experiences in an unresponsive state. *Brain and Cognition*, *77*(3), 327–334. https://doi.org/10.1016/j.bandc.2011.09.002

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1023–1035. https://doi.org/10.1016/j.neubiorev.2009.12.014

273

Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a Few Null Findings Falsify

   Prefrontal Theories of Conscious Perception? *The Journal of Neuroscience*,

   37(40), 9593–9602. https://doi.org/10.1523/jneurosci.3217-16.2017

Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the Phenomenology to the

   Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS*

   Computational Biology, *10*(5). https://doi.org/10.1371/journal.pcbi.1003588

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open

   source software for advanced analysis of MEG, EEG, and invasive

   electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*.

   https://doi.org/10.1155/2011/156869

Owen, A. M., & Coleman, M. R. (2008). Functional neuroimaging of the vegetative

   state. *Nature Reviews Neuroscience*, *9*(3), 235–243.

   https://doi.org/10.1038/nrn2330

Owen, A. M., Coleman, M. R., Boly, M., Matthew, H. D., Laureys, S., & Pickard, J.

   D. (2006). Detecting awareness in the vegetative state. *Science*, *313*(5792),

   1402–1402. https://doi.org/10.1196/annals.1417.018

Palanca, B. J. A., Mashour, G. A., & Avidan, M. S. (2009). Processed

   electroencephalogram in depth of anesthesia monitoring. *Current Opinion in*

   Anaesthesiology, *22*(5), 553–559.

   https://doi.org/10.1097/ACO.0b013e3283304032

Papiernik, J., Binder, M., & Gawłowska, M. (2022). Auditory steady-state potentials

   in gamma band as a diagnostic biomarker in disorders of consciousness.

274

Parvizi, J., & Damasio, A. R. (2003). Neuroanatomical correlates of brainstem coma. *Brain*, *126*(7), 1524–1536. https://doi.org/10.1093/brain/awg166

Pashler, H., & Wagenmakers, E. J. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science*, *7*(6), 528–530. https://doi.org/10.1177/1745691612465253

Persuh, M., Larock, E., & Berger, J. (2018). Working memory and consciousness: The current state of play. *Frontiers in Human Neuroscience*, *12*(March), 1–11. https://doi.org/10.3389/fnhum.2018.00078

Peterson, A., Cruse, D., Naci, L., Weijer, C., & Owen, A. M. (2015). Risk, diagnostic error, and the clinical science of consciousness. *NeuroImage: Clinical*, *7*, 588–597. https://doi.org/10.1016/j.nicl.2015.02.008

Phillips, I. (2021). Blindsight is qualitatively degraded conscious vision. *Psychological Review*, *128*(3), 558–584. https://doi.org/10.1037/rev0000254

Pinto, Y., Neville, D. A., Otten, M., Corballis, P. M., Lamme, V. A. F., De Haan, E. H. F., … Fabri, M. (2017). Split brain: Divided perception but undivided consciousness. *Brain*, *140*(5), 1231–1237. https://doi.org/10.1093/brain/aww358

Pitts, M. A., Lutsyshyna, L. A., & Hillyard, S. A. (2018). The relationship between attention and consciousness: An expanded taxonomy and implications for no-report paradigms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755). https://doi.org/10.1098/rstb.2017.0348

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(16), 7398–7403. https://doi.org/10.1073/pnas.91.16.7398

Purdon, P. L., Pavone, K. J., Akeju, O., Smith, A. C., Sampson, A. L., Lee, J., … Brown, E. N. (2015). The Ageing Brain: Age-dependent changes in the electroencephalogram during propofol and sevofluranegeneral anaesthesia. *British Journal of Anaesthesia*, *115*, i46–i57. https://doi.org/10.1093/bja/aev213

Purdon, P. L., Pierce, E. T., Mukamel, E. A., Prerau, M. J., Walsh, J. L., Wong, K. F. K., … Brown, E. N. (2013). Electroencephalogram signatures of loss and recovery of consciousness from propofol. *Proceedings of the National Academy of Sciences*, *110*(12), E1142–E1151. https://doi.org/10.1073/pnas.1221180110

Radek, L., Kallionpää, R. E., Karvonen, M., Scheinin, A., Maksimow, A., Långsjö, J., … Valli, K. (2018). Dreaming and awareness during dexmedetomidine- and propofol-induced unresponsiveness. *British Journal of Anaesthesia*, *121*(1), 260–269. https://doi.org/10.1016/j.bja.2018.03.014

Raichle, M E, MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Raichle, Marcus E., & Snyder, A. Z. (2007). A default mode of brain function: A brief history of an evolving idea. *NeuroImage*, *37*(4), 1083–1090. https://doi.org/10.1016/j.neuroimage.2007.02.041

Ramachandran, V. S., & Hirstein, W. (1997). Three Laws of Qualia. *Journal of Consciousness Studies*, *4*(5), 429–458.

Ramsay, M. A. E., Savege, T. M., Simpson, B. R. J., & Goodwin, R. (1974). Hospital Topics: Controlled Sedation with Alphaxalone-Alphadolone. *British Medical Journal*, *22*, 656–659.

Razi, A., & Friston, K. J. (2016). The Connected Brain: Causality, models, and intrinsic dynamics. *IEEE Signal Processing Magazine*, *33*(3), 14–55. https://doi.org/10.1109/MSP.2015.2482121

Razi, A., Seghier, M. L., Zhou, Y., McColgan, P., Zeidman, P., Park, H.-J., … Friston, K. J. (2017). Large-scale DCMs for resting-state fMRI. *Network Neuroscience*. https://doi.org/10.1162/NETN_a_00015

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, *3*(4), 261–270. https://doi.org/10.1038/nrn783

Ringli, M., & Huber, R. (2011). Developmental aspects of sleep slow waves: Linking sleep, brain maturation and behavior. *Progress in Brain Research*, *193*, 63–82. https://doi.org/10.1016/B978-0-444-53839-0.00005-3

Rosa, M. J., Friston, K., & Penny, W. (2012). Post-hoc selection of dynamic causal models. *Journal of Neuroscience Methods*, *208*(1), 66–78. https://doi.org/10.1016/j.jneumeth.2012.04.013

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs

metacognitive visual awareness. *Cognitive Neuroscience*, *1*(3), 165–175. https://doi.org/10.1080/17588921003632529

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Sahraie, A., Hibbard, P. B., Trevethan, C. T., Ritchie, K. L., & Weiskrantz, L. (2010). Consciousness of the first order in blindsight. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 21217–21222. https://doi.org/10.1073/pnas.1015652107

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Computers in Biology and Medicine*, *41*(12), 1110–1117. https://doi.org/10.1016/j.compbiomed.2011.06.020

Salin, P. A., & Bullier, J. (1995). Corticocortical connections in the visual system: Structure and function. *Physiological Reviews*, *75*(1), 107–154. https://doi.org/10.1152/physrev.1995.75.1.107

Sanders, M. D., Warrington, E. K., Marshall, J., & Wieskrantz, L. (1974). "Blindsight": Vision in a Field Defect. *The Lancet*, *303*(7860), 707–708. https://doi.org/10.1016/S0140-6736(74)92907-9

Sanders, R. D., Banks, M. I., Darracq, M., Moran, R., Sleigh, J., Gosseries, O., … Boly, M. (2018). Propofol-induced unresponsiveness is associated with impaired feedforward connectivity in cortical hierarchy. *British Journal of Anaesthesia*, *121*(5), 1084–1096. https://doi.org/10.1016/j.bja.2018.07.006

Sanders, Robert D., Tononi, G., Laureys, S., & Sleigh, J. W. (2012).

Unresponsiveness ≠ unconsciousness. *Anesthesiology*, *116*(4), 946–959.

https://doi.org/10.1097/ALN.0b013e318249d0a7

Sanz, L. R., Thibaut, A., Edlow, B. L., Laureys, S., & Gosseries, O. (2021). Update

on neuroimaging in disorders of consciousness. *Current Opinion in Neurology*,

34(4), 488–496.

Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V.,

Casarotto, S., … Massimini, M. (2015). Consciousness and complexity during

unresponsiveness induced by propofol, xenon, and ketamine. *Current Biology*,

25(23), 3099–3105. https://doi.org/10.1016/j.cub.2015.10.014

Schiff, N. D. (2010). Recovery of consciousness after brain injury: a mesocircuit

hypothesis. *Trends in Neurosciences*, *33*(1), 1–9.

https://doi.org/10.1016/j.tins.2009.11.002

Schiff, N. D. (2015). Cognitive motor dissociation following severe brain injuries.

JAMA Neurology, *72*(12), 1413–1415.

https://doi.org/10.1001/jamaneurol.2015.2899

Schnakers, C., Perrin, F., Schabus, M., Majerus, S., Ledoux, D., Damas, P., …

Laureys, S. (2008). Voluntary brain processing in disorders of consciousness.

Neurology, *71*(20), 1614–1620.

https://doi.org/10.1212/WNL.0b013e3181bd68bc

Schnakers, Caroline, Giacino, J. T., Løvstad, M., Habbal, D., Boly, M., Di, H., …

Laureys, S. (2015). Preserved covert cognition in noncommunicative patients

with severe brain injury? *Neurorehabilitation and Neural Repair*, *29*(4), 308–
317. https://doi.org/10.1177/1545968314547767

Schnakers, Caroline, Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M.,
Majerus, S., … Laureys, S. (2009). Diagnostic accuracy of the vegetative and
minimally conscious state: Clinical consensus versus standardized
neurobehavioral assessment. *BMC Neurology*, *9*, 1–5.
https://doi.org/10.1186/1471-2377-9-35

Schnider, T. W., Minto, C. F., Shafer, S. L., Gambus, P. L., Andersen, C., Goodale,
D. B., & Youngs, E. J. (1999). The Influence of Age on Propofol
Pharmacodynamics. *The Journal of the American Society of Anesthesiologists*,
90(6), 1502–1516.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, *85*(2),
461–464. https://doi.org/10.1103/PhysRevLett.85.461

Schröter, M. S., Spoormaker, V. I., Schorer, A., Wohlschläger, A., Czisch, M.,
Kochs, E. F., … Ilg, R. (2012). Spatiotemporal reconfiguration of large-scale
brain functional networks during propofol-induced loss of consciousness.
Journal of Neuroscience, *32*(37), 12832–12840.
https://doi.org/10.1523/JNEUROSCI.6046-11.2012

Schrouff, J., Perlbarg, V., Boly, M., Marrelec, G., Boveroux, P., Vanhaudenhuyse,
A., … Benali, H. (2011). Brain functional integration decreases during
propofol-induced loss of consciousness. *NeuroImage*, *57*(1), 198–205.
https://doi.org/10.1016/j.neuroimage.2011.04.020

Seeber, M., Cantonas, L. M., Hoevels, M., Sesia, T., Visser-Vandewalle, V., & Michel, C. M. (2019). Subcortical electrophysiological activity is detectable with high-density EEG source imaging. *Nature Communications*, *10*(1), 1–7. https://doi.org/10.1038/s41467-019-08725-w

Seel, R. T., Sherer, M., Whyte, J., Katz, D. I., Giacino, J. T., Rosenbaum, A. M., … Zasler, N. (2010). Assessment scales for disorders of consciousness: Evidence-based recommendations for clinical practice and research. *Archives of Physical Medicine and Rehabilitation*, *91*(12), 1795–1813. https://doi.org/10.1016/j.apmr.2010.07.218

Seth, A. (2016). The real problem. *AEON Magazine*.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321. https://doi.org/10.1016/j.tics.2008.04.008

Shapiro, K., Raymond, J., & Arnell, K. (2009). Attentional blink. *Scholarpedia*, *4*(6), 3320. https://doi.org/10.4249/scholarpedia.3320

Sherman, S. M., & Guillery, R. W. (1998). On the actions that one nerve cell can have on another: Distinguishing "drivers" from "modulators." *Proceedings of the National Academy of Sciences of the United States of America*, *95*(12), 7121–7126. https://doi.org/10.1073/pnas.95.12.7121

Shin, H. W., Kim, H. J., Jang, Y. K., You, H. S., Huh, H., Choi, Y. J., … Hong, J. S. (2020). Monitoring of anesthetic depth and EEG band power using phase lag

entropy during propofol anesthesia. *BMC Anesthesiology*, *20*(1), 1–10.

https://doi.org/10.1186/s12871-020-00964-5

Siclari, F., Baird, B., Perogamvros, L., Bernardi, G., LaRocque, J. J., Riedner, B., …

Tononi, G. (2017). The neural correlates of dreaming. *Nature Neuroscience*,

20(6), 872–878. https://doi.org/10.1038/nn.4545

Sitt, J. D., King, J. R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., …

Naccache, L. (2014). Large scale screening of neural signatures of

consciousness in patients in a vegetative or minimally conscious state. *Brain*,

137(8), 2258–2270. https://doi.org/10.1093/brain/awu141

Sleigh, J., Warnaby, C., & Tracey, I. (2018). General anaesthesia as fragmentation of

selfhood: insights from electroencephalography and neuroimaging. *British*

Journal of Anaesthesia, *121*(1), 233–240.

https://doi.org/10.1016/j.bja.2017.12.038

Soddu, A., Vanhaudenhuyse, A., Bahri, M. A., Bruno, M. A., Boly, M., Demertzi,

A., … Noirhomme, Q. (2012). Identifying the default-mode component in

spatial IC analyses of patients with disorders of consciousness. *Human Brain*

Mapping, *33*(4), 778–796. https://doi.org/10.1002/hbm.21249

Solms, M. (2019). The hard problem of consciousness and the free energy principle.

Frontiers in Psychology, *9*(JAN), 1–16.

https://doi.org/10.3389/fpsyg.2018.02714

Solms, M., & Friston, K. (2018). How and why consciousness arises: Some

considerations from physics and physiology. *Journal of Consciousness Studies*,

25(5–6), 202–238.

Sperry, R. (1984). Consciousness, personal identity, and the divided brain.
Neuropsychologia, *22*(6), 661–673.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural
description of the human brain. *PLoS Computational Biology*, *1*(4), 0245–0251.
https://doi.org/10.1371/journal.pcbi.0010042

Stam, C. J., Nolte, G., & Daffertshofer, A. (2007). Phase lag index: Assessment of
functional connectivity from multi channel EEG and MEG with diminished bias
from common sources. *Human Brain Mapping*, *28*(11), 1178–1193.
https://doi.org/10.1002/hbm.20346

Stamatakis, E. A., Adapa, R. M., Absalom, A. R., & Menon, D. K. (2010). Changes
in resting neural connectivity during propofol sedation. *PLoS ONE*, *5*(12).
https://doi.org/10.1371/journal.pone.0014224

Stender, J., Gosseries, O., Bruno, M. A., Charland-Verville, V., Vanhaudenhuyse,
A., Demertzi, A., … Laureys, S. (2014). Diagnostic precision of PET imaging
and functional MRI in disorders of consciousness: A clinical validation study.
The Lancet, *384*(9942), 514–522. https://doi.org/10.1016/S0140-
6736(14)60042-8

Stender, J., Kupers, R., Rodell, A., Thibaut, A., Chatelle, C., Bruno, M. A., …
Gjedde, A. (2015). Quantitative rates of brain glucose metabolism distinguish
minimally conscious from vegetative state patients. *Journal of Cerebral Blood
Flow and Metabolism*, *35*(1), 58–65. https://doi.org/10.1038/jcbfm.2014.169

Stender, J., Mortensen, K. N. N., Thibaut, A., Darkner, S., Laureys, S., Gjedde, A., &
Kupers, R. (2016). The Minimal Energetic Requirement of Sustained
Awareness after Brain Injury. *Current Biology*, *26*(11), 1494–1499.
https://doi.org/10.1016/j.cub.2016.04.024

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., &
Friston, K. J. (2010). Ten simple rules for dynamic causal modeling.
NeuroImage, *49*(4), 3099–3109.
https://doi.org/10.1016/j.neuroimage.2009.11.015

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable
perception. *Trends in Cognitive Sciences*, *13*(7), 310–318.
https://doi.org/10.1016/j.tics.2009.04.006

Tagliazucchi, E., & van Someren, E. J. W. (2017). The large-scale functional
connectivity correlates of consciousness and arousal during the healthy and
pathological human sleep cycle. *NeuroImage*, *160*, 55–72.
https://doi.org/10.1016/j.neuroimage.2017.06.026

Tagliazucchi, E., Von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K., &
Laufs, H. (2013). Breakdown of long-range temporal dependence in default
mode and attention networks during deep sleep. *Proceedings of the National
Academy of Sciences of the United States of America*, *110*(38), 15419–15424.
https://doi.org/10.1073/pnas.1312848110

Takahashi, T., Cho, R. Y., Murata, T., Mizuno, T., Kikuchi, M., Mizukami, K., …
Wada, Y. (2009). Age-related variation in EEG complexity to photic
stimulation: A multiscale entropy analysis. *Clinical Neurophysiology*, *120*(3),

476–483. https://doi.org/10.1016/j.clinph.2008.12.043

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired

consciuousness: A practical scale. *The Lancet*, *304*(7872), 81–84.

https://doi.org/10.1080/14635240.1999.10806094

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science*

Review, *2*(2), 1–11. https://doi.org/10.1021/pr070350l

The Mathworks Inc. (2017). MATLAB. Natick, Massachusetts: The MathWorks Inc.

Thibaut, A., Bodien, Y. G., Laureys, S., & Giacino, J. T. (2020). Minimally

conscious state "plus": diagnostic criteria and relation to functional recovery.

Journal of Neurology, *267*(5), 1245–1254. https://doi.org/10.1007/s00415-019-

09628-y

Thibaut, A., Bruno, M. A., Chatelle, C., Gosseries, O., Vanhaudenhuyse, A.,

Demertzi, A., … Laureys, S. (2012). Metabolic activity in external and internal

awareness networks in severely brain-damaged patients. *Journal of*

Rehabilitation Medicine, *44*(6), 487–494. https://doi.org/10.2340/16501977-

0940

Thibaut, A., Panda, R., Annen, J., Sanz, L. R. D., Naccache, L., Martial, C., …

Gosseries, O. (2021). Preservation of Brain Activity in Unresponsive Patients

Identifies MCS Star. *Annals of Neurology*, *90*(1), 89–100.

https://doi.org/10.1002/ana.26095

Tononi, G. (2004). An information integration theory of consciousness. *BMC*

Neuroscience, *5*, 1–22. https://doi.org/10.1186/1471-2202-5-42

Tononi, G. (2012). Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie*, *150*(4), 293–329. https://doi.org/10.4449/aib.v149i5.1388

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience*, *17*(7), 450–461. https://doi.org/10.1038/nrn.2016.44

Tsuchiya, N., & Koch, C. (2015). *The Relationship Between Consciousness and Top-Down Attention. The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*. Elsevier Ltd. https://doi.org/10.1016/B978-0-12-800948-2.00005-4

Vaitl, D., Gruzelier, J., Jamieson, G. A., Lehmann, D., Ott, U., Sammer, G., … Weiss, T. (2005). Psychobiology of altered states of consciousness. *Psychological Bulletin*, *131*(1), 98–127. https://doi.org/10.1037/0033-2909.131.1.98

Van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Psychology*, *1*(DEC), 1–13. https://doi.org/10.3389/fpsyg.2010.00217

van den Heuvel, M. P., & Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.3539-11.2011

van Diessen, E., Numan, T., van Dellen, E., van der Kooi, A. W., Boersma, M., Hofman, D., … Stam, C. J. (2015). Opportunities and methodological

challenges in EEG and MEG resting state functional brain network research. *Clinical Neurophysiology*, *126*(8), 1468–1481. https://doi.org/10.1016/j.clinph.2014.11.018

van Erp, W. S., Lavrijsen, J. C. M., Vos, P. E., Bor, H., Laureys, S., & Koopmans, R. T. C. M. (2015). The vegetative state: Prevalence, misdiagnosis, and treatment limitations. *Journal of the American Medical Directors Association*, *16*(1), 85.e9-85.e14. https://doi.org/10.1016/j.jamda.2014.10.014

van Vugt, M. K., Sederberg, P. B., & Kahana, M. J. (2007). Comparison of spectral analysis methods for characterizing brain oscillations. *Journal of Neuroscience Methods*, *162*(1–2), 49–63. https://doi.org/10.1016/j.jneumeth.2006.12.004

van Wijk, B. C. M., Cagnan, H., Litvak, V., Kühn, A. A., & Friston, K. J. (2018). Generic dynamic causal modelling: An illustrative application to Parkinson's disease. *NeuroImage*, *181*(January), 818–830. https://doi.org/10.1016/j.neuroimage.2018.08.039

Vanhaudenhuyse, A., Charland-Verville, V., Thibaut, A., Chatelle, C., Tshibanda, J. F. L., Maudoux, A., … Gosseries, O. (2018). Conscious while being considered in an unresponsive wakefulness syndrome for 20 years. *Frontiers in Neurology*, *9*(AUG), 1–9. https://doi.org/10.3389/fneur.2018.00671

Vanhaudenhuyse, A., Demertzi, A., Schabus, M., Noirhomme, Q., Bredart, S., Boly, M., … Laureys, S. (2011). Two Distinct Neuronal Networks Mediate the Awareness of Environment and of Self. *Journal of Cognitive Neuroscience*, *23*(3), 570–578.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J. F., Bruno, M. A., Boveroux,

    P., Schnakers, C., … Boly, M. (2010). Default network connectivity reflects the

    level of consciousness in non-communicative brain-damaged patients. *Brain*,

    133(1), 161–171. https://doi.org/10.1093/brain/awp313

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb:

    Phase synchronization and large-scale integration. *Nature Reviews*

    Neuroscience, *2*(4), 229–239. https://doi.org/10.1038/35067550

Velichkovsky, B. B. (2017). Consciousness and working memory: Current trends

    and research perspectives. *Consciousness and Cognition*, *55*(July), 35–45.

    https://doi.org/10.1016/j.concog.2017.07.005

Vinck, M., Oostenveld, R., Van Wingerden, M., Battaglia, F., & Pennartz, C. M. A.

    (2011). An improved index of phase-synchronization for electrophysiological

    data in the presence of volume-conduction, noise and sample-size bias.

    NeuroImage, *55*(4), 1548–1565.

    https://doi.org/10.1016/j.neuroimage.2011.01.055

Wang, D. J. J., Jann, K., Fan, C., Qiao, Y., Zang, Y. F., Lu, H., & Yang, Y. (2018).

    Neurophysiological basis of multi-scale entropy of brain complexity and its

    relationship with functional connectivity. *Frontiers in Neuroscience*, *12*(MAY),

    1–14. https://doi.org/10.3389/fnins.2018.00352

Warnaby, Catherine E., Sleigh, J. W., Hight, D., Jbabdi, S., & Tracey, I. (2017).

    Investigation of Slow-wave Activity Saturation during Surgical Anesthesia

    Reveals a Signature of Neural Inertia in Humans. *Anesthesiology*, *127*(4), 645–

    657. https://doi.org/10.1097/ALN.0000000000001759

Warnaby, Catherine Elizabeth, Seretny, M., Mhuircheartaigh, R. N., Rogers, R.,

    Jbabdi, S., Sleigh, J., & Tracey, I. (2016). Anesthesia-induced suppression of

    human dorsal anterior insula responsivity at loss of volitional behavioral

    response. *Anesthesiology*, *124*(4), 766–778.

    https://doi.org/10.1097/ALN.0000000000001027

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world"

    networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

White, N. S., & Alkire, M. T. (2003). Impaired thalamocortical connectivity in

    humans during general-anesthetic- induced unconsciousness. *NeuroImage*,

    19(2), 402–411. https://doi.org/10.1016/S1053-8119(03)00103-4

Whitfield-Gabrieli, S., Moran, J. M., Nieto-Castañón, A., Triantafyllou, C., Saxe, R.,

    & Gabrieli, J. D. E. (2011). Associations and dissociations between default and

    self-reference networks in the human brain. *NeuroImage*, *55*(1), 225–232.

    https://doi.org/10.1016/j.neuroimage.2010.11.048

Wijdicks, E. F. (2001). The diagnosis of brain death. *New England Journal of*

    Medicine, *344*(16), 1215–1221.

Wilson, R. C., & Collins, A. G. E. (2019). Ten simple rules for the computational

    modeling of behavioral data. *ELife*, *8*(e49547).

    https://doi.org/10.7554/eLife.49547

Wolman, D. (2012). The split brain: a tale of two halves. *Nature*, *483*(7389), 260–

    263.

Wu, X., Zou, Q., Hu, J., Tang, W., Mao, Y., Gao, L., … Yang, Y. (2015). Intrinsic

functional connectivity patterns predict consciousness level and recovery outcome in acquired brain injury. *Journal of Neuroscience*, *35*(37), 12932–12946. https://doi.org/10.1523/JNEUROSCI.0415-15.2015

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, *8: e68910*.

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *NeuroImage*, *200*(May), 12–25. https://doi.org/10.1016/j.neuroimage.2019.06.032

Zhang, Y., Wang, Y., Yan, F., Song, D., Wang, H., Wang, Q., & Huang, L. (2020). Influence of pre-anesthesia dynamic frontal-parietal communication on individual susceptibility to propofol. *Clinical Neurophysiology*, *131*(11), 2566–2577. https://doi.org/10.1016/j.clinph.2020.07.018