TOWARDS EEG-BASED BIOMARKERS OF LARGE SCALE BRAIN NETWORKS

TOWARDS EEG-BASED BIOMARKERS OF LARGE SCALE BRAIN NETWORKS

BY

SAURABH BHASKAR SHAW, M.A.Sc, B.Eng.

A THESIS

SUBMITTED TO THE DEPARTMENT OF NEUROSCIENCE GRADUATE PROGRAM

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY (PH.D.)

Doctor of Philosophy (Ph.D.) (2021) McMaster University

(Neuroscience Graduate Program) Hamilton, Ontario, Canada

TITLE:	Towards EEG-based biomarkers of large scale brain networks
AUTHOR:	Saurabh Bhaskar Shaw M.A.Sc. (Biomedical Engineering), McMaster University, Hamilton, Canada B.Eng. (Electrical & Biomedical Engineering), McMaster University, Hamilton, Canada
SUPERVISOR:	Dr. Suzanna Becker
NUMBER OF PAGES:	xiii, 173

Lay Abstract

Synergistic activity in specific brain regions gives rise to large-scale brain networks, linked to specific cognitive tasks. Interactions between three such brain networks are believed to underlie healthy behavior and cognition, and these are found to be disrupted in those with mental health disorders. The ability to cheaply and effectively detect these networks can enable routine network-based clinical assessments, improving diagnosis of mental health disorders and tracking their response to treatment. The first study in this thesis found major flaws in a popular method to assess these networks using a suitably cheap imaging method called electroencephalography(EEG). The remainder of the thesis addressed these issues by first identifying healthy patterns of network activity, followed by designing a novel method to identify network activity using EEG. The final study validates the developed method by tracking network changes after lifestyle interventions. In sum, this thesis takes a step towards improving the clinical accessibility of such brain network-based biomarkers.

Abstract

Several major functional networks in the brain have been identified, based on sub-regions in the brain that display functionally correlated, synchronous activity and perform common cognitive functions. Three such brain networks (default mode network - DMN, central executive network - CEN, and salience network - SN) form a tri-network model of higher cognitive functioning and are found to be dysregulated in a number of psychopathologies, such as PTSD, autism, schizophrenia, anxiety, depression, bipolar disorder and fronto-temporal dementia (FTD). Current therapies that improve the patient's cognitive and behavioural states are also found to re-normalize these dysregulated networks, suggesting a correlation between network dysfunction and behavioural dysregulation. Hence, assessing tri-network activity and its dynamics can be a powerful tool to objectively assess treatment response in such psychopathologies. Doing so would most likely rely on functional magnetic resonance imaging (fMRI), as one of the most commonly used modalities for studying such brain networks. While fMRI allows for superior spatial resolution, it poses serious challenges to widespread clinical adoption due to MRI's high operational costs and poor temporal resolution of the acquired signal. One potential strategy to overcome this shortcoming is by identifying the activity of these networks using their EEG-based temporal signatures, greatly reducing the cost and increasing accessibility of

using such measures. This thesis takes a step towards improving the clinical accessibility of such brain network-based biomarkers.

Doing so first required the exploration of a popular EEG-based method currently being used to study brain networks in mental health disorders - Microstates. This work uncovered flaws in the core assumptions made in assessing Microstates, necessitating the development of an alternate method to detect such network activity using EEG. To accomplish this, it was important to understand the healthy dynamics between the three brain networks constituting the tri-network model and test one of the core predictions of this model, i.e. the SN gates the DMN and CEN activation based on interoceptive and exteroceptive task demands. Probing this question next uncovered mechanistic details of this process, discovering that the SN co-activates with the task-relevant network. Using this information, a novel machine learning pipeline was developed that used simultaneous EEG-fMRI data to identify EEG-based signatures of the three networks within the tri-network model, and could use these signatures to predict network activation. Finally, the novel machine learning pipeline was trialed in a study investigating the effects of lifestyle interventions on the network dynamics, showing that CEN-SN synchrony can predict response to intervention, while DMN-SN synchrony can develop in those that fail to respond. The understanding of healthy network dynamics gathered from the earlier study helps interpret these results, suggesting that the non-responders persistently activated DMN as a maladaptive strategy.

In conclusion, the studies discussed in this thesis have improved our understanding of healthy network dynamics, uncovered critical flaws in currently popular methods

of EEG-based network analysis, provided an alternative methodology to assess network dynamics using EEG, and also validated its use in tracking changes in network synchrony. The identified EEG signatures of widely used functional networks, will greatly increase the clinical accessibility of such brain network measures as biomarkers for neuropathologies. Monitoring the level of network activity in affected subjects may also lead to the development of novel individualized treatments such as brain network-based neurofeedback interventions.

"Research is what I'm doing when I don't know what I'm doing."

- Wernher von Braun

Acknowledgements

I would first like to acknowledge the immense contributions of my supervisor, Dr. Sue Becker. Her unwavering mentorship, guidance and support over the course of my PhD degree has not only improved my critical thinking and research acumen, making me a better researcher, but has also taught me numerous other soft skills, such as work-life balance, and supervisory skills, that I find useful on a daily basis. This was especially relevant in navigating the research challenges posed by the COVID-19 pandemic. My experience has collectively allowed me to understand the kind of scientist I would like to be and have greatly influenced my decision to pursue a career in the academic sciences.

I would also like to thank my committee members Dr. Margaret Mckinnon and Dr. Jim Reilly for their sharing their valuable time, expertise and research insights on the various aspects of my research, greatly improving the interdisciplinary impact of my research. Their guidance has helped me decide the direction of my future research program.

Starting from mentoring my undergraduate research project, to attending research conferences together, Dr. Kiret Dhindsa has also played a pivotal role in helping me mature into a seasoned researcher. I deeply cherish this friendship and thank him for his continuing support.

I extend my heartiest thanks to Yarden Levy, Allison Mizzi, and Isaac Kinley for their collaborative roles in conducting this research and their research kinship.

Majority of this time-intensive research would have been impractical without the help of numerous undergraduate students over the course of my PhD - Thank you Gabrielle Herman, Imasha Perera, Tyler Brown, Antoaneta Todorova, Tatiana Gayowski, Shawniya Alageswaran, Selena Singh, Michael Balas, Helen Ma, Daniel Merlano, Connor O'Donnell, Christy Au Yeung, Christiana Kiervin, Bradley Karat, Araniya Selvarajah, Amna Saleem, Ali Howidi and Alexander Pleava. I am extremely grateful to these amazing future researchers.

I would like to acknowledge my girlfriend Maria, and my closest friends Daniel, Kyle, Jaki, Neel and Dwip for sticking by me during the ups and downs of the research process. We can finally meet up now that I am done my PhD!

Most importantly, I would like to deeply and whole-heartedly thank my family Srinidhi Shaw, Suchitra Shaw & Sanjay Shaw. They have been a rock throughout my life and have helped me grow into the person I am today. I am eternally grateful for everything they have done and continue to do. They have always put a lot of faith in me and have sacrificed a lot for me to reach this milestone.

Finally, at this culmination of two and half decades of formal education, I would also like to extend my thanks to all of my amazing teachers, educators, librarians, and accessory staff that have enabled my learning and facilitated this achievement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In recent years, Canada has faced an increasing prevalence of mental health disorders. With approximately one in ten Canadians developing either post traumatic stress disorder (PTSD) or major depressive disorder (MDD) in their lifetime, this country has one of the highest rates of mental health disorders among developed nations (Goldstein et al., 2016). Current clinical practices have to rely on unreliable self-reported measures of symptom severity to develop treatment plans and track treatment response. Often, this approach does not adequately characterize the symptom profile of each individual patient, especially for those with heterogeneous psychopathologies such as PTSD. This leads to challenges in adequately personalizing therapies, resulting in sub-optimal treatment outcomes, longer hospital stays and increased burdens on the Canadian healthcare infrastructure. Thus, there is an urgent need to develop objective measures that can be used to adequately characterize each patient's disease profile and develop personalized treatment plans. One promising, yet underutilized category of measures is brain network-based biomarkers (Parkes et al., 2020).

Several major functional networks in the brain have been identified, based on

sub-regions in the brain that display functionally correlated, synchronous activity and perform common cognitive functions (Relión et al., 2017). Three such brain networks form a tri-network model of higher cognitive functioning and are found to be dysregulated in a number of psychopathologies, such as PTSD (Connolly et al., 2014; Menon, 2011), autism, schizophrenia, anxiety, depression, bipolar disorder and fronto-temporal dementia (FTD). Current therapies that improve the patient's cognitive and behavioural states are also found to re-normalize these dysregulated networks (Kluetsch et al., 2014), suggesting a correlation between network dysfunction and behavioural dysregulation. Hence, assessing tri-network activity and its dynamics can be a powerful tool to objectively assess treatment response in such psychopathologies.

Doing so would typically rely on functional magnetic resonance imaging (fMRI), the most commonly used imaging modality for studying such brain networks. While fMRI allows for excellent spatial resolution, it presents serious barriers to widespread clinical adoption for a number of reasons. First, MRI's high operational costs and limited availability mean it is not widely used for routine health assessments. Furthermore, fMRI scanning can be claustrophobic and stressful, and therefore not well tolerated by many patients with severe mood disorders. Finally, because fMRI measures blood flow, an indirect correlate of neural activity, it has relatively poor temporal resolution; it can resolve brain activity on the timescale of several seconds, but not the millisecond level time-scale of actual underlying neural events. This calls for a cheaper, less invasive modality that is capable of acquiring neural signals at much higher temporal resolutions. Electroencephalography (EEG) is one such modality that uses surface electrodes on the scalp to acquire neural signals from the underlying

cortex (Gloor, 1969). Despite its benefits, however, EEG has not been widely used for probing network dynamics due to its poor spatial resolution.

One potential strategy to overcome this shortcoming is by identifying the activity of intrinsic connectivity networks (ICNs) using their EEG-based temporal signatures. There is some evidence that the discussed brain networks have EEG signatures such as specific frequency band power activity (Sauseng et al., 2005; Scheeringa et al., 2008; Shah et al., 2017), phase signatures (Dimitriadis et al., 2015) and other spatiotemporal patterns (Charbonnier et al., 2016) that could potentially be used to identify the level of activity in each of the three networks of the tri-network model. Identification and development of such EEG signatures of ICNs as brain network-based biomarkers will be particularly useful for tracking progress in the treatment of neuropathologies in which network dynamics within and between these three networks is dysregulated. Monitoring the level of network activity in affected subjects may also lead to the development of novel individualized treatments such as brain network-based neurofeedback interventions.

This thesis takes a step towards improving the clinical accessibility of such EEG-based ICN biomarkers by 1) exploring currently used EEG-based methods to study brain networks, 2) understanding the healthy dynamics between the three brain networks constituting the tri-network model, 3) creating a novel EEG-based method to detect such network activity and dynamics of the three networks, and finally 4) showing its utility in tracking intervention-linked changes in brain network dynamics. Accomplishing this involved four major steps constituting four studies, as discussed below.

1. The first study explored the adequacy of core assumptions made by EEG Microstates, a popular EEG-based method currently used to study network dynamics in mental health disorders.

2. The second study investigated the network dynamics of the healthy brain during task switching, as predicted by the tri-network model, using fMRI, the current standard for studying brain networks.

3. Using information gathered from study 2, the third study identified the EEG signatures of the three networks by using simultaneously acquired EEG-fMRI data, and created a machine learning model capable of predicting ICN activation based on EEG data alone.

4. Finally, the fourth study validated the use of the machine learning model developed in study 3 by using it to investigate the ICN changes due to two lifestyle interventions, aerobic exercise and neurofeedback, in a sample of healthy undergraduate students.

These studies have collectively identified unique task-linked ICN dynamics that can be identified using the newly developed machine learning model, and have shown its proof-of-concept application in studying intervention-linked changes in ICN dynamics. Designed to be used clinically in characterizing and tracking treatment of individuals with psychopathologies such as PTSD, the EEG-based ICN analysis tools developed by these studies have greatly increased the clinical accessibility of brain network measures as biomarkers for neuropathologies.

This thesis is organized into 8 Chapters. Chapter 2 presents a detailed review of large-scale brain networks and general network structure in the human brain. Chapter

3 introduces the challenges of collecting simultaneous EEG-fMRI data, and presents a collection of newly developed solutions that resolve these challenges. Chapter 4 presents the results of study 1, scrutinizing the assumptions underlying EEG microstate analyses, while chapter 5 discusses the results of study 2 – the fMRI-based assessment of ICN dynamics. Chapter 6 and 7 present the results of studies 3 and 4 respectively, discussing the developed machine learning model capable of detecting ICN activation using EEG data alone, and its application in assessing ICN changes after lifestyle interventions in a population of healthy undergraduate students. Finally, chapter 9 concludes the thesis by discussing the future possibility of applying the created framework for diagnostic purposes, summarizing the major contributions of this thesis.

Bibliography

Charbonnier, S., Roy, R., Doležalová, R., Campagne, A., and Bonnet, S. (2016). Estimation of Working Memory Load using EEG Connectivity Measures. *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies*, 4(Biostec):122–128.

Connolly, C. G., Wu, J., Ho, T. C., Hoeft, F., Wolkowitz, O., Frank, G., Hendren, R., Max, J. E., Paulus, M. P., Susan, F., Banerjee, D., Simmons, A. N., and Yang, T. T. (2014). Resting-State Functional Connectivity of Subgenual Anterior Cingulate Cortex in Depressed Adolescents. 74(12):898–907.

Dimitriadis, S. I., Sun, Y., Kwok, K., Laskaris, N. A., Thakor, N., and Bezerianos, A. (2015). Cognitive Workload Assessment Based on the Tensorial Treatment

of EEG Estimates of Cross-Frequency Phase Interactions. *Annals of Biomedical Engineering*, 43(4):977–989.

Gloor, P. (1969). Hans Berger on Electroencephalography. *American Journal of EEG Technology*.

Goldstein, R. B., Smith, S. M., Chou, S. P., Saha, T. D., Jung, J., Zhang, H., Pickering, R. P., Ruan, W. J., Huang, B., and Grant, B. F. (2016). The epidemiology of DSM-5 posttraumatic stress disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *Social Psychiatry and Psychiatric Epidemiology*, 51(8):1137–1148.

Kluetsch, R. C., Ros, T., Théberge, J., Frewen, P. A., Calhoun, V. D., Schmahl, C., Jetly, R., and Lanius, R. A. (2014). Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback. *Acta Psychiatrica Scandinavica*, 130(2):123–136.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, 15(10):483–506.

Parkes, L., Satterthwaite, T. D., and Bassett, D. S. (2020). Towards precise resting-state fMRI biomarkers in psychiatry: synthesizing developments in transdiagnostic research, dimensional models of psychopathology, and normative neurodevelopment. *Current Opinion in Neurobiology*, 65:120–128.

Relión, J. D. A., Kessler, D., Levina, E., and Taylor, S. F. (2017). Network classification with applications to brain connectomics. pages 1–28.

Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, 57(2):97–103.

Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., and Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International Journal of Psychophysiology*, 67(3):242–251.

Shah, N. J., Arrubla, J., Rajkumar, R., Farrher, E., Mauler, J., Kops, E. R., Tellmann, L., Scheins, J., Boers, F., Dammers, J., Sripad, P., Lerche, C., Langen, K. J., Herzog, H., and Neuner, I. (2017). Multimodal Fingerprints of Resting State Networks as assessed by Simultaneous Trimodal MR-PET-EEG Imaging. *Scientific Reports*, 7(1):6452.

Chapter 2

Functional Brain Networks

2.1 Introduction

The brain is a complex interconnected system that determines and modulates an individual's personality, behavior, and their abilities to perform simple and complex tasks. Features of this complicated system have been studied at varying scales ranging from individual synapses, to single neurons, to aggregates of a few hundred neurons, and finally to large scale brain regions (Varela et al., 2001).

Over the past two decades, the study of the neural correlates of human behaviour and cognition have transitioned from a modular localist approach to a more distributed, networked systems approach (Bressler and Menon, 2010). Termed intrinsically connected brain networks (ICN) or simply, functional networks (FN), each networked region forms a cluster of nodes that activate synchronously and perform a cohesive functional role. These ICNs cover a wide range of cortical brain regions and span numerous functions required for performing everyday actions and cognitive processing. A subset of these ICNs are shown in Figure 2.1 (see Thomas Yeo et al.

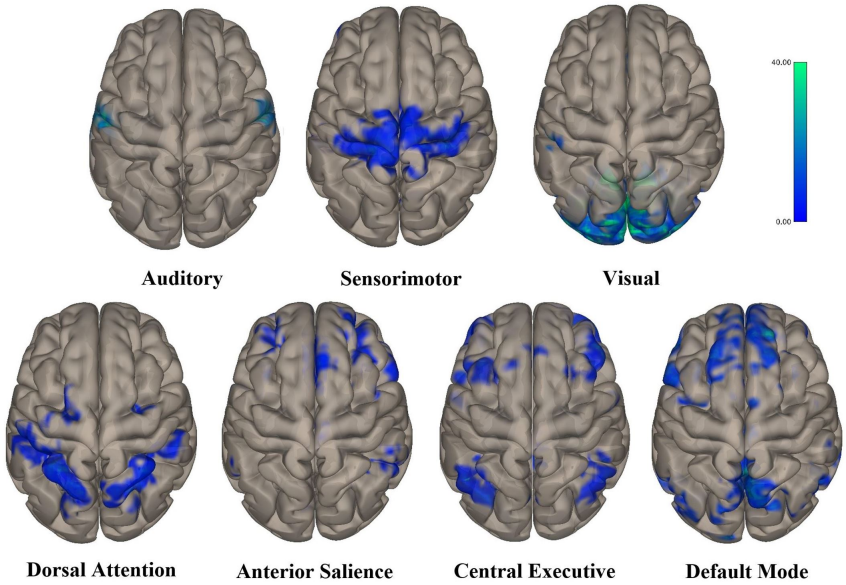

Figure 2.1: Seven major intrinsic connectivity networks (ICNs) found using spatiotemporal ICA analyses on fMRI BOLD signal, labeled based on results from Shirer et al. (2012a). Activity shown is thresholded at FDR-corrected $p < 0.05$.

(2011) for more ICNs). In addition to task-linked roles, some of these networks are also essential for the thought processes ongoing in the absence of overt tasks (Raichle and Mintun, 2006) and are involved in the integration of previous experiences and potential future experiences into a cohesive experiential timeline, creating the concept of the "self" (Spreng et al., 2010).

Apart from cortical nodes, recent advances in high-resolution neuroimaging have also uncovered the involvement of sub-cortical structures such as the cerebellum in such cognitive processes (Diedrichsen et al., 2009), suggesting their inclusion within these distributed large-scale brain networks (Dosenbach et al., 2008).

This chapter provides an introductory exposition of ICNs by discussing key network-based characteristics of ICNs and popular ICN-based models of cognition.

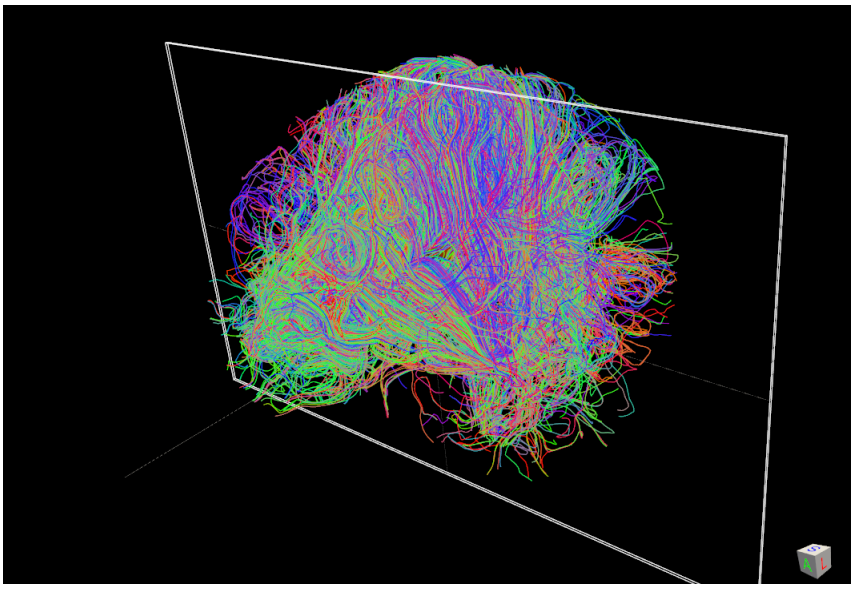

Figure 2.2: Inter-hemispheric white matter tracts connecting the two cerebral hemispheres via the Corpus Callosum. Such tracts form the edges of the structural connectome. See Oishi et al. (2008) for more details.

2.2 The Connectome

Mirroring the "genome" that represents the collection of genes subserving all phenotypes, the connectome refers to the collection of connectivity patterns between brain regions that underlie complex behaviour, emotions and cognition [SBS:cite Sporns, 2005 and Hagmann, 2005]. The connectome is often described by the functional connectivity (functional connectome) that arises between brain regions that are also structurally connected through white-matter tracts (structural connectome). As expected, the structural connections remain relatively stable over time (Bullmore and Sporns, 2009), while the functional connections are found to dynamically change with task-specific demands.

2.2.1 Terminology

The connectome has been popularly characterized using graph theory, where each brain region forms a *node*, and each connection between a pair of brain regions forms an *edge* with an associated *edge weight* that corresponds to the strength of the connection between these brain regions. The edge represents a physical connection between brain regions via white-matter tracts for structural networks, and a functional link in the case of a functional network. Additionally, each edge can also have an associated direction representing the direction of information flow, making the graph a *directed graph*.

Constructing graphs using these principles gives rise to natural node-based network measures that can characterize local and global brain network structure. One such measure is the *node degree*, which refers to the number of other nodes connected to a given node. Highly connected nodes (high degree) can be critical for information transfer, forming a *hub* (van den Heuvel et al., 2010; Power et al., 2013). Such hubs are often also associated with low *average path length*, referring to the average number of nodes that need to be traversed before reaching the current node. Hence, the hubs can be reached quickly from other nodes in the network, and conversely, the hub node can reach other network nodes without needing many intermediate nodes. Hubs are also characterized by low *clustering*, which represents the extent to which the current node forms isolated clumps of nodes. In other words, hubs are connected to a wide range of nodes in the network and rarely form isolated clusters of connections with a limited number of nodes. Finally, hubs also portray high *betweeness centrality*, which is the number of shortest paths connecting other pairs of nodes passing through the current node. This measure represents the "connector" property of the node and a

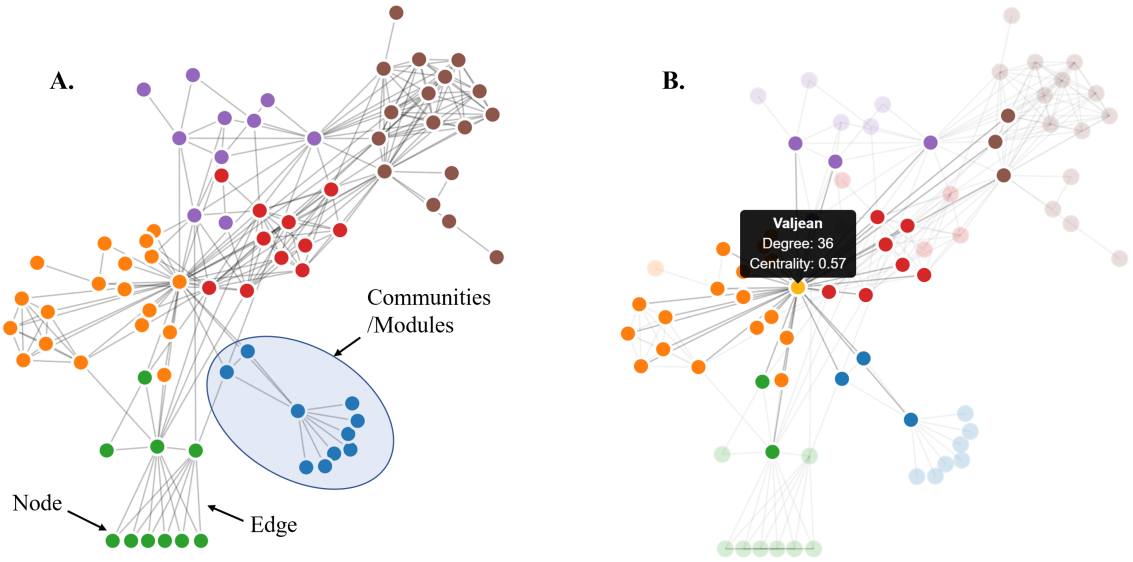

Figure 2.3: A graph of major characters in Les Miserables. A. Each node represents a character, while each edge represents a scene in which the characters appear together. The communities/modules are shown in different colors. B. The node with the largest degree and betweeness centrality is highlighted, showing its hub-like properties. The data was sourced from The Stanford Graphbase, and visualized by Connectthedots.

high value means that other nodes can use the current node as an intermediate node to quickly to get to other nodes in the network.

These brain networks are analogous to social networks, where each person is a represented by a node and some relationships between people form the edges of the social graph. This is portrayed in Figure 2.3, where a graph is constructed based on the on-screen appearance of actors in the Les Miserables movie. In this example, Valjean is well connected to other actors (high degree), and can reach most other actors through his current acquaintances (low average path length). Furthermore, Valjean interacts with a wide range of actors, not limiting himself to a select few (low clustering). Finally, if some actor wanted to convey a message to another actor,

they could do so quickly by asking Valjean to convey the message, either directly, or through his primary contacts (high betweeness centrality). Hence, Valjean would be considered a hub in this network.

Study of the brain-derived graph has uncovered numerous hubs within the brain (Power et al., 2013), some of which are even connected to other high degree nodes, forming *"rich clubs"*. These clusters of high degree nodes are critical for information transfer (van den Heuvel and Sporns, 2011), and are found at highly connected brain regions such as the hippocampus, precuneus, thalamus, and key fronto-parietal regions (van den Heuvel and Sporns, 2011). Interestingly, these regions also are key hubs within major ICNs and are essential nodes for connecting different sub-regions within the network, as discussed below.

In addition to hub-driven connectivity seen in the human brain, study of the local and global graph structure of the human connectome has revealed a "small-world" organization, referring to a structure that has numerous local connections with few long distance connections. Continuing with the social network analogy, this phrase is derived from the sensation of a "small-world" that one may feel after finding common acquaintances with a new person that they may meet. This is a by-product of most social networks usually being limited to numerous close people (short-distance/local connections) with a few long-distance acquaintances (long-distance connections), that allows disconnected clusters of people to be connected. Such an organization is considered efficient, since people perform most tasks with their local social circles, while occasionally connecting with new people in other social circles via their acquaintances. In the context of neural processing, a small-world network organization allows for extensive lower level processing within local clusters, followed by processed information

being transferred between distant brain regions over longer-range connections. The local connections are known to be much more stable under normal cognition, compared to the longer range connections, which can change based on task demands (Park et al., 2012). Such a small-world organization can also be disrupted by psychopathologies such as schizophrenia (Bassett et al., 2008), or even psychedelic drugs such as psilocybin (Petri et al., 2014). In fact, schizophrenia is known to affect some of the hubs discussed above, affecting information flow between major sections of the ICNs (Van Den Heuvel et al., 2013). Interestingly, the small-world structure seems critical for extremely basic cognitive functions such as consciousness, with propofol-induced loss of consciousness leading to a breakdown of longer-range connections and reorganization of hub structure, which re-emerge once the effects of the anesthetic wear off and the patients gain consciousness (Lee et al., 2013).

This small-world organization results in an efficient hierarchical system that shows high clustering, and short path lengths, that are repeated at multiple scales (scale-free). The dense local connections form sub-modules/sub-networks, that in turn connect with other sub-modules/sub-networks to form the modules/ICNs described earlier, which communicate across the longer range global connections (Park and Friston, 2013). Given that the intrinsically connected networks (ICNs) are a byproduct of heirachical organization of the small-world brain architecture, disruptions in the small-world architecture lead to dysregulated ICN activity and dynamics. Consequently, a wide range of psychopathologies that disrupt the graph structure of the human connectome (Van Den Heuvel et al., 2013) also show disruptions in ICN connectivity and dynamics (Menon, 2011). Hence, ICN changes can be a promising indicator of changes in the underlying connectome, and can be used as biomarkers

for the characterization of psychopathologies.

Around the turn of the previous decade, researchers started synthesizing these links between network structure, cognition and behaviour into network-based cognitive frameworks that could explain a wide range of network, behavioural and cognitive dysfunctions observed in such psychopathologies (Dosenbach et al., 2008; Bressler and Menon, 2010; Menon, 2011). Two such influential network-based models of cognition were largely adopted to describe behaviour, attention and cognitive control, and are discussed in detail in the following sections.

2.2.2 Dual-networks model

Over the past few decades, numerous studies have attempted to understand the neural underpinnings of top-down cognitive control and develop practically relevant models (Hammond and Summers, 1972; MacDonald et al., 2000; Koechlin et al., 2003). Many of the earliest findings of focal neural activation underlying cognitive control found that the dorso-lateral prefrontal cortex (dlPFC) and the anterior cingulate cortex (ACC) were both activated during top-down control and working memory tasks (D'Esposito et al., 1995; MacDonald et al., 2000; Koechlin et al., 2003). Interestingly, this was often accompanied by coherent activation in some parietal (posterior parietal cortex - PPC) and opercular (insula) brain regions (D'Esposito et al., 1995; Cohen et al., 1997). These connected brain regions were eventually found to exhibit "small-world" architecture, and constituted two functionally distinct brain networks that worked together to accomplish working memory and top-down cognitive control (Dosenbach et al., 2008). Under this model, the frontal and parietal

brain regions form a fronto-parietal control network (FPN), while the anterior cingulate and insulo-opercular regions form a secondary cingulo-opercular control network (CON), with some cerebellar nodes bridging these two networks. This model posits that both networks maintain task-relevant information, albeit for different purposes. The FPN is ascribed an adaptive control role that rapidly adapts to changes in the task, while the CON performs stable set-maintenance. These related yet distinct roles are thought to run in parallel, leading to a "reactionary" FPN network that can quickly react to changes/errors in the environment, which is then integrated over a longer time period within the CON to create a more stable representation of the relevant items in working memory. Follow-up work has shown that these networks are comprised of multiple sub-networks, much like the small-world architecture discussed in the previous section.

The FPN is found to have numerous distinct sub-networks (Thomas Yeo et al., 2011; Dixon et al., 2018), that have distinct functional roles (Nee, 2021). For instance, a gradient of functional hierarchy exists within the frontal and parietal brain regions, such that regions most distal to the sensory-motor cortices process the most abstract levels of information that are relevant for future decisions (internal), while regions that are proximal to the sensory-motor cortices process progressively less complex information, which are based on the current inputs (external) (Nee, 2021). In the context of this gradient, the dlPFC and the PPC of the FPN seem to be located in an "intermediary" zone that is thought to facilitate integration of FPN-based cognition with goal-directed behaviour and lower-level attentional control. The former is achieved by integrating with default mode network (DMN) sub-regions located

distally along the sensory-motor axis, while the latter involves integration with dorsal attention network (DAN) nodes located proximal to the sensory-motor cortices. Hence, integration between various large-scale brain networks seems necessary for appropriately processing and interacting with events in everyday life.

An alternate model was proposed based on this principle of network interactions that modeled cognitive and behavioural dysfunctions as disruptions in network interactions. This model is described in detail next.

2.2.3 Tri-network model

Out of the large number of previously identified FNs spanning a wide range of brain functions, the dynamics of three large-scale brain networks are thought to underlie the majority of behavioural and cognitive processes while performing everyday tasks. This tri-network model (Menon, 2011) extends the Dosenbach et al. (2008) model described above by including the FPN and CON as the central executive network (CEN) and the salience network (SN), respectively, in addition to the default mode network (DMN). The tri-network model posits that the CEN is involved in extroceptive processing, working memory and other executive functions; the default mode network (DMN) is involved in interoceptive processing, autobiographical memory retrieval, imagining the future, spatial planning and navigation and self-reflection; and the salience network (SN) modulates switching attention between exteroceptive and interoceptive cognitive processes by switching between the engagement of the CEN and the DMN, respectively. Recent evidence suggests that performing tasks reliant on these networks, such as recalling previous autobiographical memories and remembering a word seen in a stream of words, requires dynamic interaction between the three

networks and that task switching is modulated by the SN, that co-activates with the task-appropriate network (Shaw et al., 2021). Appropriate network switching dynamics between these three core networks is thought to be critical for healthy cognitive and behavioural processing, and disruptions in normal inter- and intra-network activity between these networks have been observed in numerous neuropsychological conditions affecting emotion and cognition. For instance, in Major Depressive Disorder (MDD) (Connolly et al., 2014), patients have trouble down-regulating activity within the DMN resulting in symptoms such as persistent rumination (Nejad et al., 2013). Furthermore, increased connectivity of the dorsal mid-insular cortex has been shown to be associated with increased depressive symptom severity. Another example of aberrant interaction of these brain networks associated with neuropathology is seen in Post-Traumatic Stress Disorder (PTSD), where disruptions in the integration of various DMN, CEN and SN components have been observed, including the posterior cingulate cortex (PCC)/Precuenus (Bluhm et al., 2009; Rabellino et al., 2015; Lanius et al., 2015), ventrolateral prefrontal and anterior cingulate cortex (Rabellino et al., 2015), ventromedial prefrontal cortex (vmPFC) and hippocampus (Sripada et al., 2012). Abnormal activation of the DMN has also been observed in PTSD patients while switching to a working memory task that normally recruits the CEN (Daniels et al., 2010). The neural substrates of numerous other psychological disorders, such as bipolar disorder, schizophrenia and mild cognitive impairment (MCI) (Menon, 2011), also overlap with critical DMN, CEN and SN nodes and it is therefore clinically relevant to effectively study and track the dynamics of these three networks. The nodes constituting these networks, and their functional roles, are further discussed in the following sub-sections.

Default Mode Network (DMN)

The default mode network (DMN) consists of some key nodes within the pre-frontal cortex (PFC), posterior parietal cortex (PPC) and medial and lateral temporal cortices (MTC/LTC) (Raichle, 2015) and is the predominantly active network when the person is not performing any particular task (Raichle, 2015). For this reason, this network was initially mis-characterized as simply reflecting the brain's idling state and was also referred to as the default network or task-negative network. However, a large body of evidence shows that the DMN subserves a wide range of cognitive tasks involved in internally-directed thoughts, spanning numerous distinct functional domains, which map onto three distinct sub-networks (Andrews-Hanna et al., 2014; Thomas Yeo et al., 2011).

The dorso-medial sub-network is comprised of the temporal pole (TP), lateral temporal cortex (LTC), temporo-parietal junction (TPJ) and the dorso-medial PFC (dmPFC). This sub-network has been linked to semantic comprehension and social cognition or "Theory of Mind" tasks such as associating emotions and desires to oneself and/or others and processing of social and non-social touch interactions (Lee Masson et al., 2020).

While the dmPFC is heavily involved in knowledge of others, the ventromedial prefrontal cortex (vmPFC), discussed in the next sub-section, is more responsive to self-referential information (Raichle, 2015).

The medio-temporal sub-network is comprised of the hippocampus, parahippocampus, retrosplenial cortex, posterior inferior parietal lobule (pIPL) or angular

gyrus (AG) and the ventromedial PFC (vmPFC). This sub-network has been associated with autobiographical memory retrieval, episodic future thinking and contextual memory retrieval.

The vmPFC is a key hub of this sub-network that receives external sensory input from the orbitofrontal cortex and relays it to deeper structures such as the hippocampus, amygdala and even the periaqueductal-grey (PAG) (Nicholson et al., 2017) that are important for episodic memory, scene construction (Hassabis and Maguire, 2007) and emotion regulation (Raichle, 2015). Reciprocal connections with these deeper structures can also modulate the activity of the vmPFC (Raichle, 2015; Lanius et al., 2006).

The core network, also known as the cortical-midline sub-network (Kim, 2012), first mentioned byNorthoff and Bermpohl (2004), integrates information from the two sub-networks discussed above, and is composed of the posterior cingulate cortex (PCC) and the antero-medial PFC (amPFC). It is primarily involved in self-referential processing and emotion evaluation, while also co-activating with the medio-temporal and dorso-medial sub-networks during autobiographical memory and social cognitive processes respectively.

More specifically, the ventral PCC (vPCC) correlates with the rest of the network during most of the previously mentioned tasks, while the dorsal PCC (dPCC) is associated with salient environmental cue detection and monitoring for behaviourally relevant , and is connected with key cognitive control centres in the brain (Leech et al., 2011). PCC is also heavily connected to the precuneus and this link is essential for the DMN's involvement with spatial navigation (Byrne et al., 2007).

The amPFC is heavily connected to other DMN nodes and is involved in autobiographical memory retrieval, episodic future thinking, affect and arousal regulation, positive and negative emotional reward and perceived value (Andrews-Hanna et al., 2014). Hence the amPFC is considered an integrative hub for salient external information, prior episodic memories and current affective state. Given its proximity to the vmPFC, the amPFC and vmPFC share some common affect processing functions, however, they may have unique roles in self related and situational processing (Lieberman et al., 2019).

The amPFC and PCC hubs are thought to drive active and passive self-referential thoughts, respectively, where a self-reference task activates the amPFC, while passive resting state activates the PCC (Benjamin et al., 2010). Additionally, Self-related memory encoding and retrieval processes were originally thought to be left and right lateralized, respectively (Craik et al., 1999), however, there is now considerable evidence surrounding the left lateralization of self-referential processing (Axelrod et al., 2017) and autobiographical memory retrieval (Kim, 2012).

The dorso-medial and the medio-temporal sub-networks are often pooled into a larger parieto-temporal subnetwork (PTS) involved in memory retrieval tasks (Kim, 2012), containing the inferior parietal lobule (containing the temporo-parietal junction and angular gyrus), medial temporal lobe and lateral temporal cortex nodes (Kim, 2012).

Interestingly, the dynamics of DMN activation is known to change depending on the dynamics of the task being performed. For example, the DMN is heavily involved in the retrieval of autobiographical memories and shows different temporal activation dynamics depending on the nature of the AM. General AMs that have sparser details

and are more generalized have a shorter time-to-retrieval, compared to specific AMs that have a greater amount of detail about the specifics of the memory (Addis et al., 2004).

Central Executive Network (CEN)

The central executive network (CEN) is a complex attentional control system that is anchored in the dorso-lateral prefrontal cortex (dlPFC) and some parietal regions. For this reason, it is sometimes referred to as the frontoparietal attention, frontoparietal control network, or task-positive network and is responsible for a host of executive functions such as updating working memory, inhibitory control or selective attention, multiple task coordination (Collette and Van der Linden, 2002; Baddeley, 1996), and random sequence generation (Baddeley et al., 1998).

Norman et al. (1986)'s two-pronged model suggests that executive functioning is a part of the supervisory activating system (SAS) that takes over attentional control from an implicit habit-based system when schema-based actions are no longer appropriate. This requires multiple component systems working in concert with one another to perform the task at hand, as described by the three-component Baddeley and Hitch (1974) model. Under this model, a modality-invariant central executive system receives information from the visuospatial and verbal subsystems through the visuospatial scratchpad and the phonological loop respectively. Each of these subsystems has distinct storage-rehearsal substructures, with rehearsal processes maintaining the stored objects in the respective short term memory (Smith and Jonides, 1997; Barton et al., 1995; Washburn and Astur, 1998).

The visuospatial scratchpad maintains and manipulates visuospatial representations of task relevant objects (Vallar and Pagano, 2002). The storage of these visuospatial representations is subserved by the right posterior parietal cortex (PPC) (Smith and Jonides, 1997; Byrne et al., 2007; Dhindsa et al., 2014), while the rehearsal-based maintenance of these representations involves right premotor areas (Smith and Jonides, 1997).

The phonological loop maintains verbal/phonological short term memory, which is refreshed by an articulatory rehearsal process (Vallar and Pagano, 2002). The phonological storage is responsible for verbal/phonological short term memory and is centered around the left inferior parietal cortex (BA 40) (Smith and Jonides, 1997), whereas the loci for sub-vocal rehearsal that maintain these objects in storage are the left Broca's area/inferior frontal gyrus, left premotor area and left supplementary motor areas (Smith and Jonides, 1997; Fegen et al., 2015). The medial frontal gyrus (MFG) and superior parietal lobule (SPL) are also involved in the rehearsal process albeit in a general supervisory role (Fegen et al., 2015).

Representations of task relevant objects can exist in either of these two systems as visual or verbal objects, and can transform from one to another via a fourth component - the episodic buffer (Baddeley, 2000). This buffer integrates information across the sensory subsystems and long-term episodic, visual and linguistic memory, forming a modality-invariant representation and is known to have a limited capacity of around 7 ± 2 objects (Kamiński et al., 2011). Such distributed patterns of modality-specific information localized in respective sensory cortices that are integrated at the parietal and prefrontal structures over increasing time lags has led to a more modern distributive view of working memory (Christophel et al., 2017).

The central executive processes that function at the level of modality-invariant working memory, such as temporal coding, integrate across the sensory subsystems and are located in the dorsolateral prefrontal cortex (dlPFC). Unlike the visuospatial and the verbal subsystems that are right and left lateralized respectively, dlPFC is found to be bilaterally activated (Smith and Jonides, 1997; Baddeley, 2003).

Numerous tasks have been used to probe executive functioning, with each task activating different executive subprocesses. For example, while the Wisconsin card sorting test (WCST), the go-no-go task and stroop interference trials all rely on inhibitory control processes, the WCST additionally relies on the executive function of shifting task attention (Collette and Van der Linden, 2002). Towers of Hanoi/London is an example of a task that relies on frontal lobe-linked planning and strategy (Sullivan et al., 2009) in addition to some inhibitory control (Collette and Van der Linden, 2002). N-back tasks are another popular task that relies on the multi-component executive process of temporal coding (Collette and Van der Linden, 2002; Smith and Jonides, 1997) and reliably activates frontal regions such as frontal poles, dlPFC, vlPFC, and latero-medial premotor cortices, in addition to some parietal regions such as medio-lateral posterior parietal cortices, dorsal cingulate (Owen et al., 2005; Yaple et al., 2019), and the precuneus (Yaple et al., 2019).

Interestingly, some brain regions such as IFG/vlPFC and the inferior parietal cortex (BA 40), that constitute some CEN sub-networks, are also found to integrate with the DMN, and are consequently thought to also be a part of DMN sub-networks. In fact, recent work in dynamic functional connectivity find that such association cortices can dynamically switch network membership, and that this switching behaviour is desirable for optimal cognitive functioning (Pedersen et al., 2018).

Salience (SN)

The salience network is composed primarily of two major cortical nodes, the anterior insula (AI) and the dorsal anterior cingulate cortex (dACC) (Menon, 2011), and some non-cortical structures such as the amygdala, ventral striatum and the ventral tegmental area/substantia nigra (VTA). It is responsible for detecting higher-level emotional and reward-linked salience of incoming stimuli, and driving goal-directed behaviours (Sridharan et al., 2008; Menon and Uddin, 2010; Menon, 2015).

The anterior insula (AI)/frontoinsular cortex (fIC) receives a wide range of sensory input, along with emotional and reward-linked salience information from the sub-cortical nodes of the SN. Consequently, AI is thought to integrate across these inputs to detect stimuli coherent with behaviourally relevant actions in a goal-directed manner.

The dorsal anterior cingulate cortex, unlike the afferent-heavy connection structure of AI, has extensive efferent connections to somatomotor and primary motor areas (Menon, 2015), along with some outputs to hypothalamus and peri-aqueductal grey (PAG) Menon and Uddin (2010). The dACC node is heavily involved with response selection, introceptive and autonomic processing (Mesulam, 1998).

The sub-cortical structures that constitute the SN include key nodes of the limbic system and reward circuitry - the amygdala and the ventral striatum/VTA respectively, which detect negative valence and reward salience of received stimuli. These feed into the AI to be integrated with other sensory input.

The two primary nodes of the SN have a very characteristic cytoarchitecture

involving single spindle neurons called Von Economo neurons, which seem to be primarily localized at the AI and dACC in the primate brain (Menon, 2015). This provides an extremely efficient communication pathway between the AI and dACC via the white matter tracts of the uncinate fasciculus (Menon, 2015).

After detecting the emotional, social or task-linked salience of an input, the SN acts as a "gate" by activating the relevant functional network to further process and respond with an appropriate action, as described in the following section.

While the nodes described above constitute the canonical salience network included in the original tri-network model, recent evidence has shown that this network is in fact a smaller anterior component of a larger salience system (Shirer et al., 2012b). A secondary salience sub-network (named the posterior salience sub-network) is found to closely interact with the above discussed nodes, performing multisensory integration in the context of emotions and processing embodiment (Harricharan et al., 2021). A core node within this sub-network is the posterior insula, which is thought to process thermosensory and pain information. In fact, the primary thermosensory cortex is believed to be located in the posterior insula, allowing it to cohesively integrate these experiences with other bottom-up sensory information reaching the posterior insula via thalamic, limbic and brainstem structures (, Bud). Hence, the posterior insula is able to identify internal bodily changes, and is found to be dysregulated in the dissociative subtype of PTSD, where processing of body-related sensations or embodiment is lacking (Harricharan et al., 2020).

2.2.4 Detecting Intrinsically Connected Networks (ICNs)

The connected nature of these network nodes poses a particularly difficult challenge in adequately detecting their activation. Firstly, non-invasively collecting signals from small nodes located deep within the brain requires a functional imaging modality, such as functional magnetic resonance imaging (fMRI), that can identify regions with high spatial resolution. However, this comes at the cost of poor temporal resolution since fMRI requires considerable time to collect signals from the entire brain at each time point, and is limited by the slow haemodynamic response (as described below). Conversely, functional imaging modalities with fine temporal resolution, such as electroencephalography (EEG) and magnetoencephalograpy (MREG), suffer from poor spatial resolution. This trade-off between spatial resolution and temporal resolution is at the heart of this thesis, as we use multi-modal methods to combine these two modalities, combining the strengths of the respective modalities. Each individual imaging modality is briefly described next, while the next chapter explores the challenges of combining these modalities together.

Functional Magnetic Resonance Imaging (fMRI)

First discovered by Seiji Ogawa in 1990, functional magnetic resonance imaging has become one of the most popular imaging modalities over the past three decades. FMRI senses neural activity by identifying changes in levels of blood oxygenation (*i.e.* Blood Oxygen Level Dependent (BOLD) signal), due to the phenomenon of increased blood flow in the local area of increased neural activity. This "haemodynamic response" results in an increase in the local ratio of oxygenated/deoxygenated blood, which in turn changes the local magnetic properties due to deoxyhaemoglobin's

paramagnetic nature, leading to an increase in the local T_2^* relaxation time. This T_2^* relaxation time refers to the time it takes the Hydrogen (H) atoms in the brain tissue to move out of phase with each other, after all the H atoms have been simultaneously excited by a pulse of energy (aligning their phases together). The rate at which this occurs depends on the local inhomogeneity of the magnetic field, since this inhomogeneity causes the H atoms at different locations to spin at slightly different speeds. Therefore, MRI imaging sequences that are T_2^* weighted (*i.e.* sensitive to the T_2^* relaxation time) show increased intensity in local areas of increased neural activity. Each voxel's intensity then represents the aggregate activity of the neurons contained within that voxel at a given time. Multiple whole-brain T_2^* weighted images detect the temporal progression of neural activity at each voxel, resulting in a BOLD time-series.

The activity-linked peak in blood flow (haemodynamic response) occurs around 3-6 seconds after activity onset (Weiskopf et al., 2004), leading to a significant delay in detecting neural activity using the BOLD signal. Additionally. the time required to image the whole brain at each time point often limits the sampling rates of this modality to $< 1Hz$. Together, these factors severely limit the temporal resolution of the fMRI-derived BOLD signal to several seconds. Yet another drawback of fMRI is the high cost of acquisition and operation, which limits the ability to use this modality for repeated assessments of neural activity.

Despite these shortcomings, fMRI is one of the most commonly used modalities for studying large-scale brain networks due to its excellent spatial resolution that allows for effective identification of spatially separated clusters.

Electroencephalography (EEG)

EEG is an alternate imaging modality that is considerably less expensive than MRI, and boasts much higher temporal resolution (on the order of microseconds). Electroencephalography (EEG) uses surface electrodes on the scalp to detect neural signals from the underlying cortex (Gloor, 1969).

The neural origin of the EEG signals are thought to be the graded postsynaptic potentials of the neural cell body (soma) and the apical dendrites of vertically orientated pyramidal cells in cortical layers three, four and five (Mayer and Bellgowan, 2014), caused by their synchronized polarization and depolarization. The EEG electrodes are also sensitive to vertical currents that are flowing perpendicular to the scalp, in the extracellular space (**?**). Since the source of these stray currents could be neurons other than those directly underneath the electrode, the EEG signal at an electrode is influenced by a fuzzy "ball" of neural tissue centered at the electrode. Furthermore, the impedance of the overlying CSF, skull and tissues further attenuate and spatially blur the neural signal (Rice et al., 2013), thereby reducing the spatial resolution of the EEG signal to an area of 10 cm^2 due to limited conductance of the neural tissue, and the distance-dependent attenuation of the electrical fields generated by the cortical apical dentrites (Buzsáki et al., 2012). Furthermore, since the EEG electrode is sensitive to all electric currents generated close to the scalp, it is often contaminated with other electrical activity such as cardiac activity, underlying muscle activity, eye movements, eye blinks, power line noise and motion artifacts. Hence, despite its advantages, EEG is not the preferred modality for probing network dynamics.

However, improvements in signal processing and the potential benefits of using

EEG have motivated the development of numerous EEG-based network measures in recent years.

One notable EEG-based measure being used to probe network dynamics is EEG Microstates. Thought to represent "atoms of thought" (Koenig and Lehmann, 1996), EEG microstates describe the continuous changes in EEG scalp topology as discontinuous transitions between quasi-stable "microstates" with fixed scalp topologies. This analysis method has been popularized due to its ease of use and the correlation between the identified microstates and some major ICNs (Britz et al., 2010). However, there is some disagreement in the network-based interpretation of some of these microstates, and recent work (some in chapter 3 of this thesis) shows that some major assumptions underlying EEG microstate analysis might not be appropriate for representing network dynamics. This is further discussed in chapter 3.

Other amplitude and phase-based EEG-based measures of connectivity have also shown promise in detecting ICN activity. These are discussed in detail in chapter 6.

Bibliography

Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., and McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: A partial least squares approach. *NeuroImage*, 23(4):1460–1471.

Andrews-Hanna, J. R., Smallwood, J., and Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, 1316(1):29–52.

Axelrod, V., Rees, G., and Bar, M. (2017). The default network and the combination of cognitive processes that mediate self-generated thought. *Nature Human Behaviour*, 1(12):896–910.

Baddeley, A. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section A*, 49(1):5–28.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11):417–423.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4(10):829–839.

Baddeley, A., Emslie, H., Kolodny, J., and Duncan, J. (1998). Random Generation and the Executive Control of Working Memory. *The Quarterly Journal of Experimental Psychology Section A*, 51(4):819–852.

Baddeley, A. D. and Hitch, G. (1974). Working Memory. In *ReCALL*, volume 255, pages 47–89.

Barton, A., Matthews, B., Farmer, E., and Belyavin, A. (1995). Revealing the basic properties of the visuospatial sketchpad: The use of complete spatial arrays. *Acta Psychologica*, 89(3):197–216.

Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., and Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and Schizophrenia. *Journal of Neuroscience*, 28(37):9239–9248.

Benjamin, C., Lieberman, D. A., Chang, M., Ofen, N., Whitfield-Gabrieli, S., Gabrieli, J. D. E., and Gaab, N. (2010). The Influence of Rest Period Instructions

on the Default Mode Network. *Frontiers in Human Neuroscience*, 4(December):1–9.

Bluhm, R. L., Williamson, P. C., Osuch, E. A., Frewen, P. A., Stevens, T. K., Boksman, K., Neufeld, R. W. J., Théberge, J., and Lanius, R. A. (2009). Alterations in default network connectivity in posttraumatic stress disorder related to early-life trauma. *Journal of psychiatry & neuroscience : JPN*, 34(3):187–94.

Bressler, S. L. and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences*, 14(6):277–290.

Britz, J., Van De Ville, D., and Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, 52(4):1162–1170.

(Bud) Craig, A. D. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10(1):59–70.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*, 10(3):186–198.

Buzsáki, G., Anastassiou, C. a., and Koch, C. (2012). The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nature reviews. Neuroscience*, 13(6):407–20.

Byrne, P., Becker, S., and Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, 114(2):340–375.

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., and Haynes, J. D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, 21(2):111–124.

Cohen, J. D., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., Smith, E. E., and Perlstein, W. M. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(April):604–608.

Collette, F. and Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews*, 26(2):105–125.

Connolly, C. G., Wu, J., Ho, T. C., Hoeft, F., Wolkowitz, O., Frank, G., Hendren, R., Max, J. E., Paulus, M. P., Susan, F., Banerjee, D., Simmons, A. N., and Yang, T. T. (2014). Resting-State Functional Connectivity of Subgenual Anterior Cingulate Cortex in Depressed Adolescents. 74(12):898–907.

Craik, F. I., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., and Kapur, S. (1999). In Search of the Self: A Positron Emission Tomography Study. *Psychological Science*, 10(1):26–34.

Daniels, J. K., McFarlane, A. C., Bluhm, R. L., Moores, K. A., Clark, C. R., Shaw, M. E., Williamson, P. C., Densmore, M., and Lanius, R. A. (2010). Switching between executive and default mode networks in posttraumatic stress disorder: alterations in functional connectivity. *Journal of psychiatry & neuroscience : JPN*, 35(4):258–66.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., and Grossman,

M. (1995). The neural basis of the central executive system of working memory. *Nature*, 378(6554):279–281.

Dhindsa, K., Drobinin, V., King, J., Hall, G. B., Burgess, N., and Becker, S. (2014). Examining the role of the temporo-parietal network in memory, imagery, and viewpoint transformations. *Frontiers in Human Neuroscience*, 8(September):1–13.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., and Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *NeuroImage*, 46(1):39–46.

Dixon, M. L., Vega, A. D. L., Mills, C., Andrews-hanna, J., Spreng, R. N., Cole, M. W., Dixon, M. L., La, A. D., Mills, C., Andrews-hanna, J., and Spreng, R. N. (2018). Correction for Dixon et al., Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proceedings of the National Academy of Sciences*, 115(13):E3068–E3068.

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3):99–105.

Fegen, D., Buchsbaum, B. R., and D'Esposito, M. (2015). The effect of rehearsal rate and memory load on verbal working memory. *NeuroImage*, 105:120–131.

Gloor, P. (1969). Hans Berger on Electroencephalography. *American Journal of EEG Technology*.

Hammond, K. R. and Summers, D. A. (1972). Cognitive control. *Psychological Review*, 79(1):58–67.

Harricharan, S., McKinnon, M. C., and Lanius, R. A. (2021). How Processing of Sensory Information From the Internal and External Worlds Shape the Perception and Engagement With the World in the Aftermath of Trauma: Implications for PTSD. *Frontiers in Neuroscience*, 15(April):1–20.

Harricharan, S., Nicholson, A. A., Thome, J., Densmore, M., McKinnon, M. C., Théberge, J., Frewen, P. A., Neufeld, R. W. J., and Lanius, R. A. (2020). PTSD and its dissociative subtype through the lens of the insula: Anterior and posterior insula resting-state functional connectivity and its predictive validity using machine learning. *Psychophysiology*, 57(1):1–23.

Hassabis, D. and Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, 11(7):299–306.

Kamiński, J., Brzezicka, A., and Wróbel, A. (2011). Short-term memory capacity (7 plusminus s2) predicted by theta to gamma cycle length ratio. *Neurobiology of Learning and Memory*, 95(1):19–23.

Kim, H. (2012). A dual-subsystem model of the brain's default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *NeuroImage*, 61(4):966–977.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, 302(5648):1181–1185.

Koenig, T. and Lehmann, D. (1996). Microstates in language-related brain potential maps show noun-verb differences. *Brain and language*, 53(2):169–182.

Lanius, R. A., Bluhm, R., Lanius, U., and Pain, C. (2006). A review of neuroimaging studies in PTSD: Heterogeneity of response to symptom provocation. *Journal of Psychiatric Research*, 40(8):709–729.

Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., and McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology*, 6(1):27313.

Lee, H., Mashour, G. A., Noh, G. J., Kim, S., and Lee, U. C. (2013). Reconfiguration of network hub structure after propofol-induced unconsciousness. *Anesthesiology*, 119(6):1347–1359.

Lee Masson, H., Pillet, I., Boets, B., and Op de Beeck, H. (2020). Task-dependent changes in functional connectivity during the observation of social and non-social touch interaction. *Cortex*, 125:73–89.

Leech, R., Kamourieh, S., Beckmann, C. F., and Sharp, D. J. (2011). Fractionating the default mode network: Distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *Journal of Neuroscience*, 31(9):3217–3224.

Lieberman, M. D., Straccia, M. A., Meyer, M. L., Du, M., and Tan, K. M. (2019). Social, self, (situational), and affective processes in medial prefrontal cortex (MPFC): Causal, multivariate, and reverse inference evidence. *Neuroscience and Biobehavioral Reviews*, 99(December 2018):311–328.

MacDonald, A. W., Cohen, J. D., Andrew Stenger, V., and Carter, C. S. (2000).

Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472):1835–1838.

Mayer, A. R. and Bellgowan, P. S. F. (2014). *Concussions in Athletics*.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, 15(10):483–506.

Menon, V. (2015). *Salience Network*, volume 2. Elsevier Inc.

Menon, V. and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, pages 1–13.

Mesulam, M. (1998). From sensation to cognition. pages 1013–1052.

Nee, D. E. (2021). Integrative frontal-parietal dynamics supporting cognitive control. *eLife*, 10:1–35.

Nejad, A. B., Fossati, P., and Lemogne, C. (2013). Self-referential processing, rumination, and cortical midline structures in major depression. *Frontiers in Human Neuroscience*, 7(OCT):1–9.

Nicholson, A. A., Friston, K. J., Zeidman, P., Harricharan, S., McKinnon, M. C., Densmore, M., Neufeld, R. W., Théberge, J., Corrigan, F., Jetly, R., Spiegel, D., and Lanius, R. A. (2017). Dynamic causal modeling in PTSD and its dissociative subtype: Bottom-up versus top-down processing within fear and emotion regulation circuitry. *Human Brain Mapping*, 38(11):5551–5561.

Norman, D., Shallice, T., Davidson, R., and Schwartz, G. (1986). Consciousness and self-regulation.

Northoff, G. and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, 8(3):102–107.

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A. W., Pike, G. B., Rosa-Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P. C., Mazziotta, J., and Mori, S. (2008). Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*, 43(3):447–457.

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1):46–59.

Park, B., Kim, J. I., Lee, D., Jeong, S. O., Lee, J. D., and Park, H. J. (2012). Are brain networks stable during a 24-hour period? *NeuroImage*, 59(1):456–466.

Park, H.-J. and Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science*, 342(6158):1238411–1238411.

Pedersen, M., Zalesky, A., Omidvarnia, A., and Jackson, G. D. (2018). Multilayer network switching rate predicts brain performance. *Proceedings of the National Academy of Sciences of the United States of America*, 115(52):13376–13381.

Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., and Vaccarino, F. (2014). Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101):20140873.

Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N., and Petersen, S. E. (2013). Evidence for hubs in human functional brain networks. *Neuron*, 79(4):798–813.

Rabellino, D., Tursich, M., Frewen, P. A., Daniels, J. K., Densmore, M., Théberge, J., and Lanius, R. A. (2015). Intrinsic Connectivity Networks in post-traumatic stress disorder during sub- and supraliminal processing of threat-related stimuli. *Acta Psychiatrica Scandinavica*, 132(5):365–378.

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience*, 38(1):433–447.

Raichle, M. E. and Mintun, M. A. (2006). Brain Work and Brain Imaging. *Annual Review of Neuroscience*, 29(1):449–476.

Rice, J. K., Rorden, C., Little, J. S., and Parra, L. C. (2013). Subject position affects EEG magnitudes. *NeuroImage*, 64(1):476–484.

Shaw, S. B., McKinnon, M. C., Heisz, J., and Becker, S. (2021). Dynamic task-linked switching between brain networks – A tri-network perspective. *Brain and Cognition*, 151(April):105725.

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012a). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, 22(1):158–165.

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012b). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral cortex (New York, N.Y. : 1991)*, 22(1):158–65.

Smith, E. E. and Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33(1):5–42.

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., and Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage*, 53(1):303–317.

Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, 105(34):12569–12574.

Sripada, R. K., King, A. P., Welsh, R. C., Garfinkel, S. N., Wang, X., Sripada, C. S., and Liberzon, I. (2012). Neural dysregulation in posttraumatic stress disorder: Evidence for disrupted equilibrium between salience and default mode brain networks. *Psychosomatic Medicine*, 74(9):904–911.

Sullivan, J. R., Riccio, C. A., and Castillo, C. L. (2009). Concurrent Validity of the Tower Tasks as Measures of Executive Function in Adults: A Meta-Analysis. *Applied Neuropsychology*, 16(1):62–75.

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., and Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165.

Vallar, G. and Pagano, C. (2002). In Handbook of memory Disorders (eds: Baddeley AD, Kopelman MD, Wilson BA.

van den Heuvel, M. P., Mandl, R. C. W., Stam, C. J., Kahn, R. S., and Hulshoff Pol, H. E. (2010). Aberrant Frontal and Temporal Complex Network Structure in Schizophrenia: A Graph Theoretical Analysis. *Journal of Neuroscience*, 30(47):15915–15926.

van den Heuvel, M. P. and Sporns, O. (2011). Rich-club organization of the human connectome. *Journal of Neuroscience*, 31(44):15775–15786.

Van Den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C., Cahn, W., Goni, J., Pol, H. E., and Kahn, R. S. (2013). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry*, 70(8):783–792.

Varela, F., Lachaux, J.-p., Rodriguez, E., and Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4):229–239.

Washburn, D. A. and Astur, R. S. (1998). Nonverbal working memory of humans and monkeys: Rehearsal in the sketchpad? *Memory & Cognition*, 26(2):277–286.

Weiskopf, N., Mathiak, K., Bock, S. W., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., and Birbaumer, N. (2004). Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Biomedical Engineering*, 51(6):966–970.

Yaple, Z. A., Stevens, W. D., and Arsalidou, M. (2019). Meta-analyses of the n-back working memory task: fMRI evidence of age-related changes in prefrontal cortex involvement across the adult lifespan. *NeuroImage*, 196(March):16–31.

Chapter 3

Simultaneous EEG-fMRI: Challenges & Solutions

3.1 Introduction

One solution to overcome the respective shortcomings of EEG and fMRI (discussed in the previous chapter) is to concurrently record both EEG and fMRI data. This poses several technical challenges stemming from the interaction of EEG electrodes and the rapidly changing magnetic fields used during fMRI scanning.

This is similar to the home experiment performed by repeatedly moving a permanent magnet in and around a loop of wire. This small-scale experiment shows a small current induced in the wire loop due to the magnetic field. The induced current is proportional to the change in magnetic flux (number of magnetic field lines per unit area) passing through the wire coil, and can therefore also be induced by a rapidly changing external magnetic field. The impact of the magnet on the current in the wire loop in this experiment is identical to fluctuating currents in the EEG electrodes

induced currents by the rapid changes in magnetic field required during fMRI imaging. Unfortunately, the magnitude of this induced current is orders of magnitude larger (on a scale of volts) than the EEG signal of interest (on a scale of microvolts), requiring additional filtering before the EEG signal can be used for further analysis. Since this artifact is derived from the MRI gradient fields, this artifact is called the gradient artifact (GA).

The paper discussed in this chapter describes the simultaneous EEG-fMRI setup used for the EEG-fMRI data collected in this thesis, and provides an overview of popular methods used to filter the GA. The paper additionally introduces two alternate GA filtering methods -

1. A parallelized implementation of FASTR, a popular GA filtering method, that leverages the multiple cores available on modern computers to greatly accelerate the GA filtering process.

2. A novel frequency-domain based GA filtering method that further reduces the computation time required to filter the GA.

The results of this paper informed the development of our EEG-fMRI analysis pipeline and validated the integrity of the filtered EEG data. Method 1 was ultimately incorporated into our analysis pipeline, greatly reducing the computation time required for the subsequent analysis steps in this study.

The primary author of this study (SBS) conceptualized and designed this study, collected the data, performed the analysis and wrote the manuscript.

3.2 Real-Time Filtering of Gradient Artifacts from Simultaneous EEG-fMRI Data

Shaw, S. B. (2017, June). Real-time filtering of gradient artifacts from simultaneous EEG-fMRI data. In 2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI) (pp. 1-4). IEEE.

Real-Time Filtering of Gradient Artifacts from Simultaneous EEG-fMRI Data

Saurabh Bhaskar Shaw, *PNB, McMaster University*

***Abstract*—EEG and fMRI are extremely popular tools to study patterns of functional brain activity. Their utility can be further enhanced when used together in simultaneous EEG-fMRI recordings. However, such recordings are ridden with artifacts due to the gradients switching within an MRI machine. These artifacts need to be filtered before the data can be further processed. Numerous tools exist for filtering such data. However, if one needed to use the data for real-time feedback (such as neurofeedback), the current methods would be too slow. This paper discusses parallel versions of the current methods and a novel FFT based method that reduces the computation time of current methods by a factor of 3 and 23 respectively. This facilitates the use of an EEG-fMRI dataset in real-time neurofeedback studies.**

***Index Terms*—Electroencephalography, EEG, functional Magnetic Resonance Imaging, fMRI, Gradient artifacts, real-time.**

I. Introduction

BOTH Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) are popular tools to probe the functional activity of the human brain. However, both have certain drawbacks owing to the nature of signals they are sensitive to. EEG has extremely high temporal resolution, however, it suffers from poor spatial resolution. This is due to the limited depth of neural tissue capable of influencing EEG voltages, and volume conduction effects that spatially blur the voltage recorded at each electrode[1]. On the other hand, fMRI has comparatively excellent spatial resolution, but poor temporal resolution due to the long acquisition times of fMRI scans and the lag introduced by the haemodynamic response driving the blood oxygen level dependent (BOLD) signal[2]. It is therefore beneficial to combine the two modalities together to obtain EEG and fMRI signals concurrently, providing high temporal and spatial resolutions respectively. Such data has been used to identify BOLD correlates of EEG evoked potentials and shifts in EEG band power[3], localize foci of intractable seizures for surgical planning purposes[4], and better constrain the EEG source localization problem for a more robust solution[5]. The equipment setup to acquire such simultaneous EEG-fMRI data is shown in figure 1. However, such a setup introduces a host of artifacts linked to the switching MRI gradients into the EEG signal and requires extensive filtering before it can be used. This paper explores current methods of filtering such gradient artifacts (GA) and proposes a novel method that allows for quick computation time.

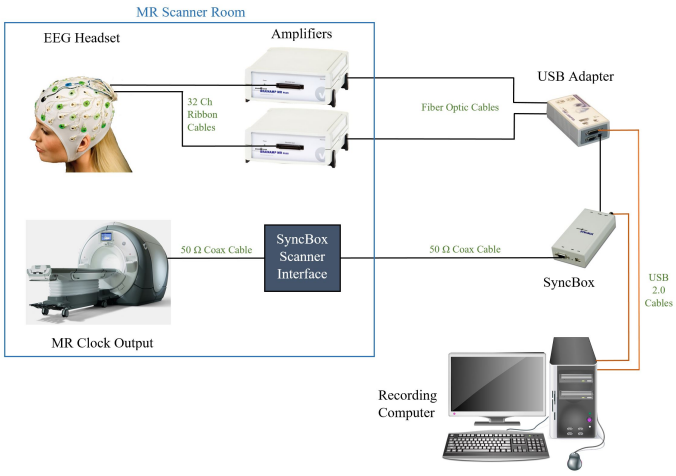

Fig. 1. The equipment setup for an EEG-fMRI system showing the components in the MRI scanner room (blue box), and the recording equipment outside the scanner room. The recording computer is placed outside the magnet room, with the amplified EEG data being transmitted via optic fiber cables going through the MRI room wave guide, ensuring integrity of the MRI magnet room RF shielding. The SyncBox is required to sync the clocks of the scanner and the acquisition computer. It also records the exact moment of the MRI Gradients switching on/off as markers in the EEG recording. This is essential for the GA filtering process.

II. Gradient Artifacts (GA)

During BOLD fMRI scanning, the MRI gradients switch back and forth rapidly to acquire multiple full brain images during the course of the scan. These rapidly switching magnetic fields induce a voltage in the electrodes of the EEG cap (according to Faraday's law). These result in massive spikes in the EEG signal trace and completely mask the underlying EEG data[6], as shown in figure 2.

Numerous approaches have been developed to remove such GAs[7] - Adaptive filtering, optimal basis set removal (OBS), alignment and subtraction (ITAS), and a realtime carbon-wire loop based method[8]. Combining multiple methods into a composite algorithm is known to increase the overall filtering efficacy of the system. One of the most popular combinations is the FASTR algorithm[9] implemented in the EEGLAB toolbox[10].

Another well packaged set of GA filtering tools is the FACET toolbox[11] for MATLAB, which implements sections

Saurabh Bhaskar Shaw was with the School of Biomedical Engineering, McMaster University before joining the Department of Psychology, Neuroscience and Behaviour (PNB), McMaster University.
E-mail: shaws5@mcmaster.ca

978-1-5386-3159-1/17/$31.00 ©2017 IEEE

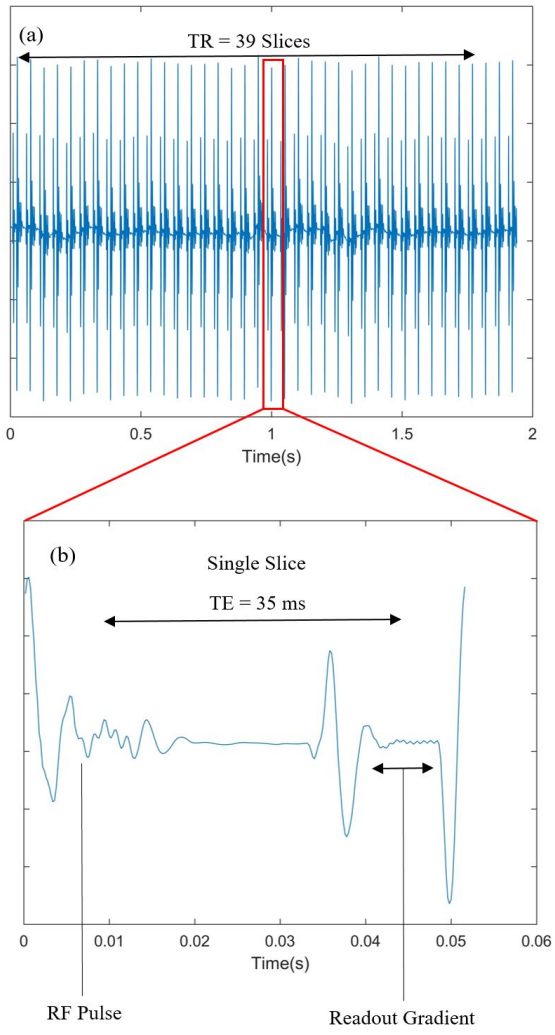

Fig. 2. Morphology of a typical gradient artifact observed over the course of 1 TR (2 seconds) is shown in panel (a). This is composed of 39 slices. One such slice is highlighted in the red box and presented in more detail in panel (b). The features seen in the artifact are a direct consequence of the EPI MR sequence used to acquire fMRI signals.

of FASTR along with another algorithm FARM, useful for electromygraphic (EMG) acquisitions in the scanner.

All algorithms work on a basic framework set by Allen, *et. al.*[12], where a GA model is created by averaging segments of gradient artifacts. The deterministic nature of these artifacts allows for the use of such a model to describe the GA throughout the signal block. The input signal (x) is segmented into recurring segments of the GA using gradient markers representative of the gradients switching on and off. The segmented data is averaged to obtain the GA model (y).

A. The FASTR Algorithm

The segmented input signal and the GA model are interpolated to a higher sampling rate, aligned and then subtracted (as shown in equation 1).

$$z_k = x_k - y \tag{1}$$

where, x_k and z_k are the raw version and filtered version of the k^{th} epoch, respectively. The residual artifact in the filtered signal is further minimized by using an optimal basis set (OBS) filtering algorithm[9]. This involves fitting the principal components from a PCA decomposition of the filtered signal (z_k) to that of the GA model (y), and removing the components with significant fits. Any further residual artifact is removed by using a recursive least squares (RLS) adaptive filter.

B. The ITAS Algorithm

This is a variant of the FASTR algorithm, where the template used for subtraction is generated by giving a greater weighting to artifact epochs closer to the current epoch while averaging. Such an algorithm helps in compensating for any drift in the GA template structure[7]. An example of such a weighting is shown in equation 2

$$y_k = \frac{\sum_{i=1}^{n} w^{|k-i|} \cdot x_k}{\sum_{i=1}^{n} w^{|k-i|}} \tag{2}$$

where w is the weight factor (0.9) and n is the total number of epochs available. This template is then subtracted from the raw signal for the k^{th} epoch as follows

$$z_k = x_k - y_k \tag{3}$$

C. Custom FFT Algorithm

This algorithm relies on the assumption that the noisy input signal (x) is a convolution of the EEG signal (x_{EEG}) and the gradient noise (n_{GA}).

$$x = x_{EEG} \star n_{GA} \tag{4}$$

The clean EEG signal (x_{EEG}) is recovered by de-convolving the EEG signal and the gradient noise, estimated using the GA model. De-convolution is difficult in the time domain, however, can be accomplished by division in the frequency domain. Applying a Fourier transform ($\mathcal{F}$) to the equation 4,

$$\mathcal{F}[x] = \mathcal{F}[x_{EEG} \star n_{GA}] \tag{5}$$

$$X = X_{EEG} \cdot N_{GA} \tag{6}$$

The estimated clean EEG signal can then be expressed as shown in equation 7, given that N_{GA} is non-zero.

$$x_{EEG} = \mathcal{F}^{-1}\left[\frac{X}{N_{GA}}\right] \tag{7}$$

III. Methods

Simultaneous resting EEG-fMRI data (eyes-closed) was acquired from a 30 year old, healthy male subject using a GE MR750 3T MRI and 32 channel head, neck and spine (HNS) RF coil (General Electric Healthcare, Milwaukee, WI). The fMRI scans were acquired using a 2D GRE EPI sequence with a TE of 35 ms and TR of 2000 ms. Each volume consisted of 39 slices. A BrainProducts (Brain Products GmbH, Gilching, Germany) 64 channel MR compatible EEG cap was used (BC-MR 64) to acquire the EEG data at a sampling rate of 5000 Hz.

TABLE I. The computational time taken by various versions of GA filtering algorithms on two computers with different computational power

Run Type	Computer 1	Computer 2
FASTR Stock code	1.876s ± 0.047s	0.819s ± 0.041s
FASTR CPU Parallelized code	0.675s ± 0.034s	0.338s ± 0.036s
Custom FFT	0.080s ± 0.041s	0.043 ± 0.036s
ITAS (FACET)	2.493s ± 0.071s	2.355s ± 0.232s

The setup shown in figure 1 was used to acquire the markers for the MR gradient onsets each fMRI slice acquisition.

The EEG data was segmented into 2 second long (1 TR length) epochs and each epoch was subject to 4 filtering algorithms:

1) Stock FASTR: The unchanged algorithm from the FMRIB plugin in the EEGLAB toolbox (http://fsl.fmrib.ox.ac.uk/eeglab/fmribplugin/).
2) CPU Parallelized FASTR: The stock FASTR code was modified to make use of multiple cpu cores.
3) Custom FFT algorithm: The custom algorithm based on deconvolution as described in section II-C.
4) ITAS algorithm: The ITAS algorithm implemented by the FACET toolbox [11](https://github.com/hansiglaser/facet).

The computation was performed on two computers of differing capabilities. Computer 1 was a Dell XPS 15 laptop with a 6th generation Intel i7 quad core processor (2.5 GHz), 16 GB RAM, and an Nvidia GeForce 750M graphics card. Computer 2 was a Dell desktop with a 4th generation Intel i7 quad core processor (3.5 GHz), 32 GB RAM and an Nvidia GeForce 760 graphics card.

IV. Results

The GA artifacts seen in figure 2 are greatly reduced and not noticeable in the filtered EEG traces shown in figure 3. Using the output from the stock FASTR algorithm as the standard, the custom FFT algorithm and ITAS algorithm achieved temporal correlations of 65% and 91%, respectively. Spectral correlations were much much higher and peaked at 85% and 98%, respectively. The custom FFT algorithm outperformed the other algorithms in computation time (shown in table I).

V. Discussion

The use of real-time processing in the study of neural patterns has steadily risen due to applications such as neurofeedback in EEG[13], fMRI[14] and even simultaneous EEG-fMRI[3]. However, real-time processing of EEG signal from a simultaneous EEG-fMRI recording requires the removal of gradient artifacts as discussed earlier. Apart from commercially available software (BrainVision RecView[3]), this is not feasible using currently available open-source tools, due to long computation times (as seen in table I). The use of commercial software is non-ideal for research as transparency and control over the modifications applied to the signal is lost. Hence there is a need to create open-source tools capable of performing GA filtering in real-time.

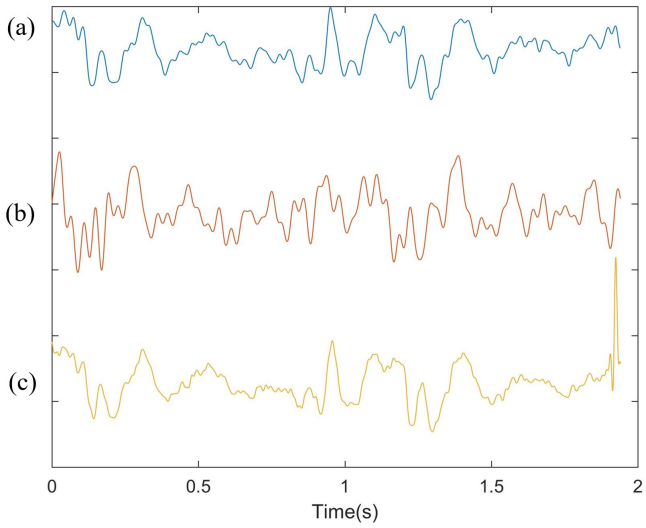

Fig. 3. The clean EEG signal from channel Fp1, filtered by (a) FASTR algorithm; (b) custom FFT algorithm; (c) ITAS algorithm. Both versions of the FASTR algorithm (stock and CPU accelerated) produced identical results.

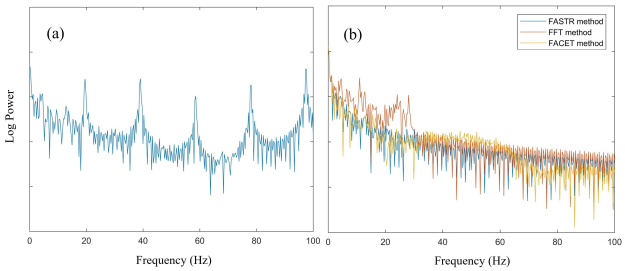

Fig. 4. The spectral content of the raw EEG signal from channel Fp1 (with the GA) is shown in panel (a). Note the high frequency spikes characteristic of the harmonics associated with the GA. The spectral content of the Fp1 EEG signal filtered using the FASTR algorithm, custom FFT algorithm, and the ITAS (FACET) algorithm, is shown in panel (b). The GA spectral spikes are missing in (b), and there is very little difference in the spectral content of the filtered signals from all three algorithms.

This study created two such tools - a modified version of the existing FASTR algorithm, and a novel FFT based algorithm. The use of parallelization to accelerate execution of the FASTR algorithm resulted in a x3 reduction in computational time. Since this relies on CPU parallelization, a faster CPU (Computer 2), outperformed a slower CPU (Computer 1) where the number of CPU cores was the same. The parallelization resulted in each CPU core filtering GAs on an independent EEG channel, accelerating the computation by roughly $(n-1)$ times, where n is the number of CPU cores. This reduced the latency of the FASTR algorithm from 1.9 seconds to 675 ms

The novel FFT based algorithm led to a greater boost in performance (x23) due to the extremely efficient operations involved in computing the fourier transform and the inverse fourier transform of the signals. Although this algorithm is effective in quickly filtering the GA (latency of 80 ms), the resultant signal carries some temporal distortions as compared to FASTR and ITAS (shown by lower temporal correlation

values in figure 3). However, the higher spectral correlation values indicate that all algorithms preserve the spectral content of the clean EEG data (shown in figure 4). Hence, the novel FFT algorithm will be ideal for real-time applications that primarily rely on spectral features. For applications that utilize temporal EEG features, the cpu accelerated FASTR is a better option.

VI. Future Directions

The real-time GA filtering algorithms described in this paper need to be further verified in tasks with known expected EEG patterns (such as ERPs). Modifications to the current algorithm need to be explored to reduce the temporal distortions. This will ensure that downstream processing of EEG temporal features remains unaffected. Furthermore, real-time algorithms need to be developed to filter ballistocardiogram (BCG) artifacts [15,16] present in EEG-fMRI recordings. This can be accelerated by using the GPU in addition to the CPU for parallelization.

References

[1] J. K. Rice, C. Rorden, J. S. Little, and L. C. Parra, "Subject position affects EEG magnitudes," *NeuroImage*, vol. 64, no. 1, pp. 476–484, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2012.09.041

[2] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer, "Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 966–970, 2004.

[3] V. Zotev, R. Phillips, H. Yuan, M. Misaki, and J. Bodurka, "Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback." *NeuroImage*, vol. 85 Pt 3, pp. 985–95, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811913005041

[4] E. L. So, "Integration of EEG, MRI, and SPECT in Localizing the Seizure Focus for Epilepsy Surgery," *Epilepsia*, vol. 41, no. s3, pp. S48–S54, 3 2000. [Online]. Available: http://doi.wiley.com/10.1111/j.1528-1157.2000.tb01534.x

[5] Y. Meir-Hasson, S. Kinreich, I. Podlipsky, T. Hendler, and N. Intrator, "An EEG Finger-Print of fMRI deep regional activation," *NeuroImage*, vol. 102, pp. 128–141, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2013.11.004

[6] A. Salek-Haddadi, K. J. Friston, L. Lemieux, and D. R. Fish, "Studying spontaneous EEG activity with fMRI," *Brain Research Reviews*, vol. 43, no. 1, pp. 110–133, 2003.

[7] P. Ritter, R. Becker, C. Graefe, and A. Villringer, "Evaluating gradient artifact correction of EEG data acquired simultaneously with fMRI," *Magnetic Resonance Imaging*, vol. 25, no. 6, pp. 923–932, 2007.

[8] J. N. van der Meer, A. Pampel, E. J. W. Van Someren, J. R. Ramautar, Y. D. van der Werf, G. Gomez-Herrero, J. Lepsien, L. Hellrung, H. Hinrichs, H. E. Möller, and M. Walter, "Carbon-wire loop based artefact correction outperforms post-processing EEG/fMRI corrections-A validation of a real-time simultaneous EEG/fMRI correction method." *NeuroImage*, vol. 125, pp. 880–894, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26505301http://dx.doi.org/10.1016/j.neuroimage.2015.10.064

[9] R. Niazy, C. Beckmann, G. Iannetti, J. Brady, and S. Smith, "Removal of FMRI environment artifacts from EEG data using optimal basis sets," *NeuroImage*, vol. 28, no. 3, pp. 720–737, 2005. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1053811905004726

[10] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[11] J. Glaser, R. Beisteiner, H. Bauer, and F. P. S. Fischmeister, "FACET - a "Flexible Artifact Correction and Evaluation Toolbox" for concurrently recorded EEG/fMRI data," *BMC neuroscience*, vol. 14, no. 1, p. 138, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24206927

[12] P. J. Allen, O. Josephs, and R. Turner, "A method for removing imaging artifact from continuous EEG recorded during functional MRI." *NeuroImage*, vol. 12, no. 2, pp. 230–9, 8 2000. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10913328

[13] R. C. Kluetsch, T. Ros, J. Théberge, P. A. Frewen, V. D. Calhoun, C. Schmahl, R. Jetly, and R. A. Lanius, "Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback," *Acta Psychiatrica Scandinavica*, vol. 130, no. 2, pp. 123–136, 2014.

[14] A. A. Nicholson, D. Rabellino, M. Densmore, P. A. Frewen, C. Paret, R. Kluetsch, C. Schmahl, J. Théberge, R. W. Neufeld, M. C. McKinnon, J. Reiss, R. Jetly, and R. A. Lanius, "The neurobiology of emotion regulation in posttraumatic stress disorder: Amygdala downregulation via real-time fMRI neurofeedback," *Human Brain Mapping*, vol. 00, no. August, 2016. [Online]. Available: http://doi.wiley.com/10.1002/hbm.23402

[15] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux, "Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction." *NeuroImage*, vol. 8, no. 3, pp. 229–239, 1998.

[16] J. C. de Munck, P. J. van Houdt, S. I. Gon??alves, E. van Wegen, and P. P. W. Ossenblok, "Novel artefact removal algorithms for co-registered EEG/fMRI based on selective averaging and subtraction," *NeuroImage*, vol. 64, no. 1, pp. 407–415, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2012.09.022

3.3 Discussion

The two methods described in this study for filtering GA artifacts in EEG data collected concurrently with fMRI, each reduced the computation time of GA filtering by leveraging parallel computing and frequency-domain based deconvolution, respectively. While the latter method showed the greatest reduction in computation time, it was susceptible to distortions of EEG data in the time-domain.

Consequently, all EEG-fMRI analyses reported in the remainder of this thesis were performed using the parallelized GA filtering method, which provided an acceleration factor that scaled with the number of processor cores available in the computer; these analyses were performed using Compute Canada's "Graham" super-computing cluster, which allows for processing jobs with up to 32 cores, and resulted in an approximately 25-fold reduction in GA filtering time compared to the traditional, non-parallel implementation of GA filtering.

The novel methods described in this chapter address gradient artifacts that the magnetic field induces on the recorded EEG signal. In addition to gradient artifacts, the subtle movements caused by the cardiac pulse induce EEG artifacts due to the motion of the EEG electrodes within the large static magnetic field of the MRI scanner. While this artifact is also present in regular EEG recordings outside the MRI scanner, its amplitude is negligible due to Earth's extremely weak magnetic field. However, a 3 Tesla MRI scanner creates a magnetic field that is 60,000 times stronger than the Earth's magnetic field, resulting in the amplification of this ballistocardiogram (BCG) artifact to levels comparable to EEG signals. Hence, the BCG artifact also needs to filtered before further EEG analysis is possible.

Since the source of the BCG data is rhythmic, the creation of a template BCG

artifact followed by its' aligned subtraction from the signal removes majority of the BCG artifact (Allen et al., 1998; de Munck et al., 2013; Kruggel et al., 2000). Another strategy to remove the BCG artifact utilizes ICA to remove the BCG component in the EEG data (Srivastava et al., 2005). Any residual BCG artifacts can be further removed via optimal basis set (OBS) estimation and removal. This method assumes that each artifact instance is a linear combination of a set of orthogonal basis functions, which can be used to create an accurate model of each artifact instance (Niazy et al., 2005). Lastly, data driven methods (Gonçalves et al., 2007) and wavelet transforms (Wan et al., 2006) can also be used to generate the templates for subtraction.

Due to the superior performance of OBS-based BCG filtering, it was used as the preferred algorithm to remove the BCG artifact from all EEG-fMRI data in this thesis.

Bibliography

Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., and Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *NeuroImage*, 8(3):229–239.

de Munck, J. C., van Houdt, P. J., Gonçalves, S. I., van Wegen, E., and Ossenblok, P. P. W. (2013). Novel artefact removal algorithms for co-registered EEG/fMRI based on selective averaging and subtraction. *NeuroImage*, 64(1):407–15.

Gonçalves, S., Pouwels, P., and Kuijer, J. (2007). Artifact removal in co-registered EEG/fMRI by selective average subtraction. *Clinical*.

Kruggel, F., Wiggins, C. J., Herrmann, C. S., and Von Cramon, D. Y. (2000). Recording of the event-related potentials during functional MRI at 3. 0 tesla field strength. *Magnetic Resonance in Medicine*, 44(2):277–282.

Niazy, R., Beckmann, C., Iannetti, G., Brady, J., and Smith, S. (2005). Removal of FMRI environment artifacts from EEG data using optimal basis sets. *NeuroImage*, 28(3):720–737.

Srivastava, G., Crottaz-Herbette, S., Lau, K. M., Glover, G. H., and Menon, V. (2005). ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. *NeuroImage*, 24(1):50–60.

Wan, X., Iwata, K., Riera, J., Ozaki, T., Kitamura, M., and Kawashima, R. (2006). Artifact reduction for EEG/fMRI recording: Nonlinear reduction of ballistocardiogram artifacts. *Clinical Neurophysiology*, 117(3):668–680.

Chapter 4

EEG Microstates

4.1 Introduction

The benefits of using EEG to detect the activity and dynamics of large-scale brain networks have motivated the development of numerous EEG-based measures that correlate with ICN activity. EEG-based microstates is one such method that has gained considerable popularity in recent years (see Figure 4.1), especially in characterizing psychopathologies such as schizophrenia (Nishida et al., 2013; Rieger et al., 2016; Strelets et al., 2003; Stevens et al., 1997), Alzheimer's (Strik et al., 1997; Khanna et al., 2015), Tourette's syndrome (Khanna et al., 2015), Panic Disorder (Kikuchi et al., 2011), multiple sclerosis (Gschwind et al., 2015), and concussion (Corradini and Persinger, 2014). This measure derives clusters of EEG scalp topologies that are assumed to be discontinuous quasi-stable states that the brain switches between while performing everyday tasks. The pathological changes in these quasi-stable states corresponding to such mental health disorders are given in table 4.1.

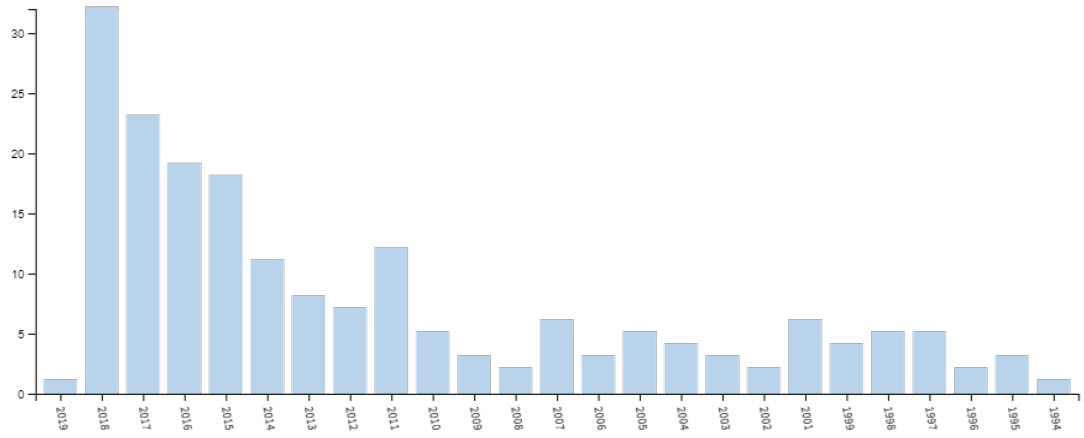

Figure 4.1: Number of publications investigating EEG microstates from 1994 to 2019. Note the recent increase in publications. Data extracted from Web Of Science.

However, despite such widespread adoption, a principled assessment of the core assumptions made during EEG Microstate analysis was missing at the time of performing this study. This motivated the study presented in this chapter that scrutinized each of microstates's three major assumptions, assessing its adequacy for use as an EEG-based biomarker of ICN activity.

The primary author of the paper was involved in the conception and development of the idea, with the help of the co-authors. The primary author also performed all the analyses and primarily wrote the paper.

Condition	Studies	Microstate changes											
		DUR				COV				OCC			
		A	B	C	D	A	B	C	D	A	B	C	D
Healthy Resting EC	Koenig et al. (2002); Milz et al. (2017)	0	0	0	0	0	0	0	0	0	0	0	0
Healthy Resting EO	Seitzman et al. (2017)	-	0	0	-	0	+	0	-	0	+	+	-
Healthy Task†	Seitzman et al. (2017)	0	0	-	0	0	0	-	+	0	0	-	+
Meditation	Faber et al. (2017)	0	0	0	-	+	0	0	-	+	0	0	-
Fluid Intelligence	Santarnecchi et al. (2017)	0	0	0	0	0	0	0	0	0	-	-	0
Schizophrenia	Strelets et al. (2003); Stevens et al. (1997); Nishida et al. (2013); Rieger et al. (2016)	0	-	0	-	0	0	0	0	+	0	+	0
FTD	Nishida et al. (2013)	0	0	-	0	0	0	0	0	0	0	0	0
Alzheimer††	Strik et al. (1997); Khanna et al. (2015)	-	-	-	-	-	0	0	0	0	0	0	0
Depression††	Khanna et al. (2015)	-	-	-	-	-	0	0	0	0	0	0	0
Tourette's	Khanna et al. (2015)	0	0	0	0	0	0	0	0	+	0	0	0
Panic Disorder	Kikuchi et al. (2011)	+	0	0	0	0	0	0	0	0	0	-	0
Multiple Sclerosis	Gschwind et al. (2016)	+	+	0	0	+	+	0	0	0	0	0	0
Concussion	Corradini and Persinger (2014)	-	0	-	-	0	0	0	0	0	0	0	0

Table 4.1: Changes in microstate metrics due to various cognitive states and pathophysiologies. Abbreviations are EC - Eyes Closed, EO - Eyes Open and FTD - Fronto-temporal Dementia. The symbol "-" corresponds to a significant decrease, "+" corresponds to a significant increase; and "0" corresponds to no significant change in the microstate metric during the given condition, as compared to the healthy resting eyes closed condition. †Milz et al. (2016) find conflicting evidence where COV and OCC of microstate D are decreased (-). †† These studies also report a decrease in overall average microstate duration, albeit with different microstates than the 4 mentioned here.

4.2 Capturing the Forest but Missing the Trees: Microstates Inadequate for Characterizing Shorter-Scale EEG Dynamics

Shaw, S. B., Dhindsa, K., Reilly, J. P., & Becker, S. (2019). Capturing the Forest but Missing the Trees: Microstates Inadequate for Characterizing Shorter-Scale EEG Dynamics. Neural computation, 31(11), 2177-2211.

LETTER Communicated by Lawrence Ward

Capturing the Forest but Missing the Trees: Microstates Inadequate for Characterizing Shorter-Scale EEG Dynamics

Saurabh Bhaskar Shaw
shaws5@mcmaster.ca
Neuroscience Graduate Program, McMaster University, Hamilton, ON L8S 4L8, Canada

Kiret Dhindsa
dhindsj@mcmaster.ca
Research and High Performance Computing, McMaster University, Hamilton, ON L8S 4L8, Canada, and Vector Institute for Artificial Intelligence, Toronto, ON M5G 1M1, Canada

James P. Reilly
reilly@mail.ece.mcmaster.ca
Vector Institute for Artificial Intelligence, Toronto, ON M5G 1M1, Canada, and Department of Electrical and Computer Engineering and McMaster School of Biomedical Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada

Suzanna Becker
becker@mcmaster.ca
Department of Psychology Neuroscience and Behaviour, McMaster University, Hamilton, ON L8S 4L8, Canada, and Vector Institute for Artificial Intelligence, Toronto, ON M5G 1M1, Canada

The brain is known to be active even when not performing any overt cognitive tasks, and often it engages in involuntary mind wandering. This resting state has been extensively characterized in terms of fMRI-derived brain networks. However, an alternate method has recently gained popularity: EEG microstate analysis. Proponents of microstates postulate that the brain discontinuously switches between four quasi-stable states defined by specific EEG scalp topologies at peaks in the global field potential (GFP). These microstates are thought to be "atoms of thought," involved with visual, auditory, salience, and attention processing. However, this method makes some major assumptions by excluding EEG data outside the GFP peaks and then clustering the EEG scalp topologies at the GFP peaks, assuming that only one microstate is active at any given time. This study explores the evidence surrounding these assumptions by studying the temporal dynamics of microstates and its clustering space using tools from dynamical systems analysis, fractal, and chaos theory

Neural Computation **31, 2177–2211** (2019)
https://doi.org/10.1162/neco_a_01229

to highlight the shortcomings in microstate analysis. The results show evidence of complex and chaotic EEG dynamics outside the GFP peaks, which is being missed by microstate analysis. Furthermore, the winner-takes-all approach of only one microstate being active at a time is found to be inadequate since the dynamic EEG scalp topology does not always resemble that of the assigned microstate, and there is competition among the different microstate classes. Finally, clustering space analysis shows that the four microstates do not cluster into four distinct and separable clusters. Taken collectively, these results show that the discontinuous description of EEG microstates is inadequate when looking at nonstationary short-scale EEG dynamics.

1 Introduction

The brain is an extremely complex system driven by electrical and metabolic activity that enables us to assess our environment, adequately respond to stimuli, and formulate complex thoughts that ultimately determine our personality and behavior. This is accomplished by an intricate network of billions of neurons, glia, and other cell types that actively work together to perform complex cognitive tasks. The brain is also active when one is not overtly performing a cognitive task (Sadaghiani, Hesselmann, Friston, & Kleinschmidt, 2010), whether the mind is wandering involuntarily or engaged voluntarily in internally driven, purposeful thought. These states of the brain are often referred to collectively as the "wakeful resting state," or simply the "resting state" (Meehan & Bressler, 2012). Characterizing the nature of this resting state activity has been an area of active research over the past few decades.

A very well-documented approach to study resting state activity is to search for networks of subregions that fire in a functionally correlated and synchronous manner. This reveals numerous functional networks (Wang, Kang, Kemmer, & Guo, 2016; Hutchison et al., 2013) that have been associated with various cognitive processes, including vision, audition, working memory, attention, and salience detection. The coordinated activity of these functional networks is thought to be essential for normal behavior and cognition and is dysregulated in neuropsychological pathologies like concussion (Mayer & Bellgowan, 2014), posttraumatic stress disorder (Parlar et al., 2017), depression (Greicius et al., 2007), schizophrenia, Alzheimer's (Sun, Yin et al., 2014), mental fatigue (Sun, Lim, Kwok, & Bezerianos, 2014), and alcohol and drug impairment (Mayhugh et al., 2016). Studying the spatial characteristics and temporal dynamics of these networks can help improve our understanding of neuropathologies and potentially create tools for clinical treatment.

Functional magnetic resonance imaging (fMRI) has predominantly been used to study such networks, owing to the excellent spatial resolution of

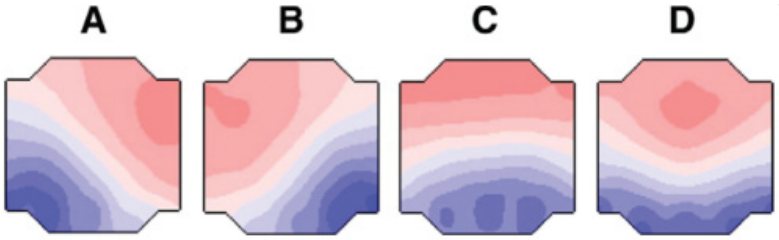

Figure 1: The four widely used microstates as defined by Milz et al. (2015). Microstates A, B, C, and D are are thought to be responsible for, respectively, auditory processing, visual processing, saliency processing, and attention reorientation (Van de Ville et al., 2010; Britz et al., 2010). This represents the scalp from a top-down view with the nose at the top of the map.

the modality. However, the high spatial resolution of fMRI acquired BOLD signal comes at the cost of poor temporal resolution (Menon & Kim, 1999; Shaw, 2017), as events can only be resolved in fMRI on a timescale of several seconds (Kim, Richter, & Ugurbil, 1997), whereas neural activity evolves on a millisecond timescale. Hence, alternate modalities with finer temporal resolution are used to study the temporal dynamics of resting state brain activity. One such modality is electroencephalography (EEG), which records electrical voltages on the scalp, generated by the summed activity of pyramidal neurons in the cortex (Mayer & Bellgowan, 2014).

While most research on resting-state EEG (rsEEG) has involved the use of methods from dynamical systems, including embedded state-space models (Wackermann, 1999) and chaotic time series modeling (Natarajan et al., 2004), an alternate method based on the identification of EEG microstates (Michel & Koenig, 2018) has gained popularity in recent years. Lehmann (1971) introduced the concept of microstates as quasi-stable (relatively stable over short 60–120 ms time periods; Michel & Koenig, 2018), periodically recurring patterns in the spatial distribution of EEG voltage topologies. Identified by clustering EEG activity at the peaks of the global field potential, these recurring patterns have come to be referred to as microstates and are postulated to be the basic "atoms of thought" (Lehmann & Koenig, 1997) making up complex conscious thought processes. These recurring patterns were originally used by Lehmann (1971) and Lehmann and Koenig (1997) to better describe the structure of the neural populations creating the patterns of inverting spatial polarity observed in EEG alpha band activity. However, it has since become a method to more generally describe EEG dynamics in terms of the scalp topologies (Khanna, Pascual-Leone, Michel, & Farzan, 2015). These microstates are clustered into a discrete set of characteristic scalp topologies, defining distinct microstate classes, as shown in Figure 1. The evolution of these microstates is described as a discontinuous process that periodically switches between the different microstate classes (Lehmann & Koenig, 1997; Lehmann et al., 2005; Milz,

Pascual-Marqui, Achermann, Kochi, & Faber, 2017; Van de Ville, Britz, & Michel, 2010). The occurrence of each microstate class is found to shortly precede perceived spontaneous thought processes such as visual imagery, verbal imagery, and abstract imagery and is consequently postulated to represent atoms of thought that need to be concatenated together to form complete thoughts (Lehmann, 1990; Lehmann & Koenig, 1997; Lehmann, Strik, Henggeler, Koenig, & Koukkou, 1998; Koenig, Kochi, & Lehmann, 1998).

Considerable work has gone into identifying the number of unique microstate classes required to adequately describe the variability in EEG scalp topologies across individuals. Despite the possibility of clustering into many more classes (Yuan, Zotev, Phillips, Drevets, & Bodurka, 2012), a cross-validation-based optimization of residual clustering variance finds that four or five classes (Pascual-Marqui, Michel, & Lehmann, 1995; Brodbeck et al., 2012) capture a large portion of the variance in the microstate data across multiple participants. Consequently, the most widely used number of clusters is four, which explains around 60% to 80% of total variance in the EEG microstates across different individuals (Britz, Van De Ville, & Michel, 2010; Koenig et al., 2002). The topologies of these four microstates are illustrated in Figure 1. Some studies have tried to identify the functional significance of these microstates by identifying the active brain regions seen in simultaneously acquired fMRI BOLD signal (Britz et al., 2010) or by using EEG source localization to identify the sources generating the scalp topology that constitutes each microstate (Custo, van der Ville, Wells, Tomescu, & Michel, 2017; Milz et al., 2016; Pascual-Marqui et al., 2014). According to these studies, the activity of the four microstates can be interpreted as visual, auditory, salience, and attention network activity. However, different studies lead to different conclusions about the significance of each microstate, as summarized in Table 1. For example, Britz et al. (2010) find microstates A and B linked with activation in predominantly auditory and visual areas, respectively, while Milz et al. (2016) find the reverse. Furthermore, the morphology of each microstate class varies considerably among different individuals (see Figure 1 of Britz et al., 2010) when studied using a dense electrode array (30 or more electrodes) and varies much less when the signal is sampled by a smaller number of electrodes (8 electrodes) (Khanna, Pascual-Leone, & Farzan, 2014). This suggests that the microstate class topologies only accurately capture global structure in the scalp topology, and that they miss the local variability in the global structure, which starts to show in finer-scale spatial topologies made possible by using arrays with a higher number of electrodes.

Given that current microstate analysis assumes that brain states transition between discrete microstates, a first-order Markov chain model can be used to estimate the transition probabilities. Assuming a conventional microstate analysis with four unique microstates, this is often summarized as a 4×4 matrix of transition probabilities for every possible transition. This could serve as a powerful descriptor of brain states and has been used as

Table 1: The Anatomical and Functional Significance of Each of the Four Major Microstate Classes.

Microstate	Regions	Significance
A	• Bilateral superior and middle temporal gyri (Britz et al., 2010) • Left middle frontal gyrus	• Auditory phonological processing (Britz et al., 2010)[a] • Visualization (Milz et al., 2016; Faber, Travis, Milz, & Parim, 2017)[a] • Sensorimotor processing (Yuan et al., 2012)
B	• Bilateral occipital areas (Britz et al., 2010) including bilateral inferior occipital gyri, bilateral cuneus, left lingual, and middle occipital gyrus	• Visual processing (Britz et al., 2010)[a] • Verbalization (Milz et al., 2016)
C	• Insular-cingulate network (Britz et al., 2010) and hubs of default mode network (DMN) (Pascual-Marqui et al., 2014), including anterior cingulate cortex bilateral inferior frontal gyri, right anterior insula, right amygdala • Salience network	• Salience and emotion regulation (Nishida et al., 2013) • Subjective interoceptive autonomic processing (Mantini, Perrucci, Del Gratta, Romani, & Corbetta, 2007) • Switching between DMN and central executive network (Menon, 2011) • Increased OCC linked to hallucinations (Khanna et al., 2015) • Task-negative microstate (Seitzman et al., 2017)
D	• Fronto-parietal regions (Britz et al., 2010), including right superior and middle frontal gyrus, right superior and inferior parietal lobes • Dorsal attentional control network (Seitzman et al., 2017)	• Executive control, working memory (Britz et al., 2010) • Focus-switching and attention reorientation (Milz et al., 2016) • Decreased DUR causes hallucinations (Nishida et al., 2013) • Task-positive microstate (Seitzman et al., 2017)

[a]There is contradictory evidence surrounding the significance of microstates A and B: Britz et al. (2010) find microstates A and B linked with activation in predominantly auditory and visual areas respectively, while Milz et al. (2016) finds the reverse.

such in numerous studies (Lehmann et al., 2005; Nishida et al., 2013; Brodbeck et al., 2012). However, recent studies have falsified the memoryless assumption of Markov processes (that the probability distribution of future transitions does not depend on past states) with respect to microstate transitions while also showing that there is considerable nonstationarity in short-range microstate transitions (von Wegner, Tagliazucchi, & Laufs, 2017; Gärtner, Brodbeck, Laufs, & Schneider, 2015), implicating longer

Table 2: Commonly Used Metrics to Describe Group Microstate Differences and Their Definitions.

Microstate Metric	Abbreviation	Definition
Duration/life span	DUR	The length of continuous time (in ms) for which a microstate class persists without switching to other microstate classes.
Coverage/percentage total time	COV	The percentage of epoch time occupied by each microstate class.
Occurrence/microstates per second/frequency	OCC	The number of times a microstate class recurred per second.
Global explained variance	GEV	The percentage of global variance explained by each microstate class.
Transition probability	TP	The probability of transition from one microstate class to another.

Note: Refer to Koenig et al. (2002) for normative values of duration, percentage total time, and microstates per second across different age groups.

range dependencies (LRD) in microstate transitions (Gschwind, Michel, & Van De Ville, 2015). Simultaneous EEG-fMRI studies have also found long-range dependencies in microstate transitions that correlate with fMRI BOLD signal along longer timescales (Van de Ville et al., 2010). These findings imply that analyses of microstate dynamics over short timescales might yield inaccurate results and suggest reserving microstate transition analysis for studying longer timescales.

Despite these shortcomings, microstate computation procedures are included in several widely used EEG software packages and have come into widespread use as biomarkers for characterizing both normal and clinical populations (for a review, see Michel & Koenig, 2018). For example, individuals with schizophrenia show increased occurrence of microstates A (Nishida et al., 2013) and C (Nishida et al., 2013; Rieger, Hernandez, Baenninger, & Koenig, 2016), and decreased average duration of microstates B and D (Nishida et al., 2013). A decrease in the average duration of each microstate is also found in depression (Khanna et al., 2015) and concussion (Corradini & Persinger, 2014). Individuals with multiple sclerosis exhibit an increase in the average duration of microstates A and B (Gschwind et al., 2016). Trends observed in such metrics computed from the sequence of microstate transitions (refer to Table 2 for definitions of the metrics) have been used as biomarkers for identifying certain neuropathologies, understanding their neurological underpinnings, and even as a neurofeedback target in a neurofeedback protocol aimed to increase the average duration spent in one microstate, in order to move the microstate metrics toward normalcy (Diaz Hernandez, Rieger, Baenninger, Brandeis, & Koenig, 2016). Such

differences in microstate metrics have been observed not only in various neuropathologies but also in healthy participants during various stages of consciousness and sleep (Brodbeck et al., 2012).

Despite the increasing acceptance of microstate analysis as a valid measure of brain dynamics, the approach has some serious methodological limitations:

1. **GFP peaks:** Most studies consider EEG data only at the GFP peaks for microstate clustering, thereby ignoring the majority (more than 90%) of the data. This removes most of the time periods from the EEG data, effectively downsampling to an extremely low sampling rate of approximately 20 Hz (based on the average interval between consecutive GFP peaks). Based on the Nyquist criterion, this signal cannot adequately represent dynamics occurring at rates above 10 Hz, which is the purported frequency of microstate transitions occurring every 60 ms to 120 ms (Michel & Koenig, 2018). One might then wonder if the observed dynamics of microstate transitions might be an artifact of such a subsampling of data. A caveat is that it assumes uniform sampling, which might not be the case when sampling at GFP peaks occurring at nonuniform intervals, resulting in nonuniform data sampling. Such nonuniform sampling could represent signal dynamics of rates up to a magnitude higher than the Nyquist limit, given that the spectrum of the signal is sparse (Wakin et al., 2012). However, the spectra of EEG signals do not satisfy this sparsity constraint due to relevant neural information being represented in a contiguous chunk of frequencies between 0 and 50 Hz and therefore cannot exploit the higher Nyquist limit afforded by nonuniform sampling. Consequently, the true Nyquist limit of the EEG signal sampled at GFP peaks is closer to that assuming uniform sampling.

Gärtner et al. (2015) and Custo et al. (2017) rationalize subsampling at the GFP peaks by viewing the peaks as time points of maximal signal-to-noise ratio (SNR), arguing that the nonstationarity in the data between GFP peaks is due to noise. However, this is not necessarily true, as nonstationarity can arise from meaningful changes in a dynamical system rather than noise. No evidence is provided to decisively rule out such dynamical changes as the contributing factor for the nonstationarity. Furthermore, GFP peaks could refer to time periods where a single neural source might be driving most of the EEG topology. Limiting analysis to just the GFP peaks would consequently pick time periods of single-source activity a priori, and it is therefore not surprising that the subsequent microstate analysis identifies microstate classes representing separate sources a posteriori (Custo et al., 2017). This is in line with the original intent of microstates as a method of describing distinct neural populations generating different EEG scalp topologies (Lehmann, 1971; Lehmann & Koenig, 1997). This leaves open the possibility that multiple neural sources might be active between GFP peaks, which could correspond to active information transfer between different brain regions. Consequently, there may be complex brain dynamics

reflected in the EEG data between GFP peaks that is being missed by ignoring time segments between GFP peaks on the premise of maximizing SNR.

2. **One microstate at a time:** Microstate analysis is based on the assumption that no more than one microstate is active at any given time. Justification for this assumption relies on the discontinuous description of brain state dynamics adopted by this analysis pipeline. Michel and Koenig (2018) argue that sufficient changes in the substates constituting a global state can be viewed as a new global state with distinct functional significance that does not overlap with the other global states. However, this assumption does not hold true, given that the different microstate classes are shown to correspond to functional networks in the brain (Custo et al., 2017) that are known to continuously vary and might also be simultaneously active (Menon, 2011). Hence microstate analysis might be imposing a discontinuous model on activity that is inherently continuous and dynamically changing.

3. **Clustering:** The analysis pipeline uses clustering for identification of the microstate class topologies, most often using k-means clustering (as well as a few other clustering algorithms, described in the next section: Brunet, Murray, & Michel, 2011; Custo et al., 2017; Pascual-Marqui et al., 1995). K-means clustering assumes that the prior probabilities of the k clusters are equal, implying that the occurrences of the microstates should be equal. However, this is clearly not the case, since microstate C is the most frequently observed microstate, while A and B are the least frequent in adults (Koenig et al., 2002). Also, most clustering is performed on EEG channel voltages with 32 to 256 channels, creating a high-dimensional space that might lead to the artificial separation of data into clusters. Furthermore, k-means clustering often finds the local minima, which may be far from the globally optimal class structure of the data due to the tendency to arrive at suboptimal solutions for nonconvex optimization problems and thus may not be reliable for microstate clustering. For example, a single cluster that has a much larger variance in some dimensions than others (e.g., an ellipsoid) might be mistakenly divided into two clusters that seem orthogonal. This is of particular concern for microstates, since classes A and B are orthogonal and might inherently be better described as a larger single cluster. This may be why different studies find opposing results for the functional significance of microstates A and B (Britz et al., 2010; Milz et al., 2016). Hence, from points 2 and 3, little to no evidence supports the view that clustering, which assumes a discrete state space, is an appropriate model to describe continuous EEG data.

The evidence presented above suggests that clusters of activation identified by microstate analyses capture longer-scale brain dynamics, averaged over a large spatial domain and fail to capture finely grained spatio temporal EEG dynamics. In this letter, we apply several methods from complex

dynamical systems to study EEG dynamics. Our results call into question the major assumptions embedded in standard microstate analysis. We first consider whether EEG scalp topologies evolve discontinuously over time, as assumed, while computing microstates from the GFP peaks. Next, the validity of the winner-takes-all model is assessed, followed by an analysis of the clustering space to ask if clustering into four classes is an appropriate model to use.

2 Methods

We conducted a series of analyses of resting-state EEG data to assess the adequacy of the microstate description for characterizing the spatiotemporal dynamics of brain activity. Specifically, we sought to assess the adequacy of the following assumptions:

1. *GFP peaks*. Are there complex dynamics between GFP peaks that are being missed by the microstate analysis since it only considers EEG scalp topologies at GFP peaks?
2. *One microstate at a time*. The winner-takes-all model of a single microstate being active at a time assumes trivial activity of the nonwinning microstate classes. Is there evidence that supports this assumption?
3. *Clustering*. Clustering as a required step in identifying microstates assumes that microstates constitute a discrete set of patterns of brain activity. Is that assumption justified, or is there really a continuous space of microstates?

To address these questions, we employed a range of methods, including a study of the clustering space generated by microstate analysis, dynamical systems analysis, and analysis of chaotic and fractal structure of the EEG scalp topologies.

A publicly available data set containing eyes-closed, resting-state EEG data from 12 right-handed healthy participants (26.6 ± 2.1 years) was used for the analyses reported here (Sockeel, Schwartz, Pélégrini-Issac, & Benali, 2016). The data were acquired at a sampling rate of 5 kHz using a 64-channel BrainProducts BrainAmp (Brain Products GmbH, Gilching, Germany) system, featuring 62 EEG channels (based on the international 10-20 system), 1 ECG channel, and 1 EOG channel. The data were preprocessed by bandpass filtering between 0.1 Hz and 50 Hz, followed by manual ICA-based removal of eye and muscle artifacts, after which the data were average-referenced and downsampled to 500 Hz. An artifact-free duration of 300 seconds (5 minutes) was used for all analyses.

2.1 Microstate Computation. EEG microstates are identified from the electrode voltages at peaks of the global field potential (GFP), computed from the preprocessed data for each participant. The GFP for an n channel

EEG system at time t is the spatial standard deviation (square root of the spatial variance) in the EEG signal across all electrodes, given by

$$GFP(t) = \sqrt{\frac{\sum_{i=1}^{n}(v_i(t) - \bar{v}(t))^2}{n}}, \tag{2.1}$$

where $v_i(t)$ is the voltage at channel i and $\bar{v}(t)$ is the mean voltage across all channels. The local maxima in GFP are peaks in EEG field strength (Milz et al., 2016). The EEG activity coinciding with all such peaks, across all electrodes, is then considered when computing the EEG microstates. Importantly, the EEG data at all times between these GFP peaks, which account for most of the data, are excluded from traditional microstate analyses. Thus, the GFP peaks are inherently assumed to represent discontinuities in the EEG topology that form quasi-stable attractor states, where the EEG scalp topology remains stable over a short time duration.

The EEG patterns across the entire set of electrodes at each GFP peak are aggregated across all participants and analyzed by a clustering algorithm (k-means or topographic atomize and agglomerate hierarchical clustering (TAAHC; Khanna et al., 2014) to identify classes of microstate topologies. This clustering process ignores the polarity of the EEG topologies, clustering opposite polarities into the same microstate class. Once the classes are identified, the scalp topology corresponding to each GFP peak is classified into one of the identified microstate classes based on a spatial similarity metric. The microstate with the highest spatial similarity is the assigned microstate class for that peak duration. This generates a time series of the similarity of the microstate classes to the spatial topology of rsEEG at each time point. The sequence of the assigned microstate classes creates a time series of microstate transitions. Microstate metrics, defined in Table 2, are extracted from this time series. This procedure is summarized in Figure 2. Numerous toolboxes have implemented this pipeline. Keypy (Milz, 2016) is one such package (implemented in Python) that uses precomputed maps from Milz et al. (2016) to sort the identified microstate classes into previously identified microstates A,B,C, and D (as shown in Figure 1). CARTOOL (Brunet et al., 2011) is another such package that uses a stand-alone program to compute these parameters.

This study used the Python Keypy package to compute the microstate class topologies and transitions, followed by temporal dynamics and clustering analyses of the identified microstates using custom scripts in Matlab.

2.2 Temporal Dynamic Analysis. The traditional view of microstates is that they represent a series of topographically discontinuous quasi-stable states in EEG that remain stable over 60 ms to 120 ms (Michel & Koenig, 2018). By this view, we should see a persistent scalp topology corresponding to the identified microstate for a period of close to 60 ms to 120 ms, followed

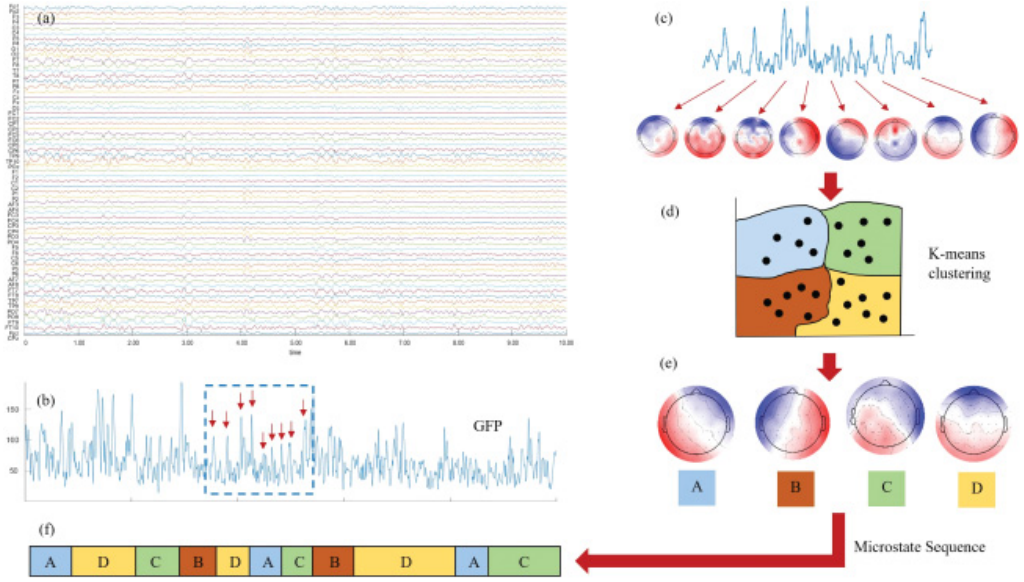

Figure 2: The procedure to identify microstate sequences. The EEG channel data (a) are used to compute GFP time series (b) using equation 2.1. The peaks in the GFP are identified (red arrows) and the EEG topologies at these peaks are recorded as the microstates. This is shown in panel c, where the boxed segment of the GFP trace is magnified. The identified microstates undergo k-means clustering (illustrated in panel d using two electrodes, which are the two axes shown) to identify dominant microstate topologies (e), while ignoring polarity. These dominant microstate classes are backprojected onto the microstate time series to generate the microstate sequence shown in panel f.

by an abrupt transition from one microstate to another at the next GFP peak. The EEG may exhibit much more complex dynamics between GFP peaks (assumption 1 at the start of section 2), that may or may not support the winner-takes-all model imposed on microstates (assumption 2 at the start of section 2). To assess this, we studied the temporal dynamics and the chaotic and fractal behavior of EEG data between the GFP peaks.

2.2.1 Descriptors of Microstate Transitions. To determine the behavior of EEG scalp topologies between the GFP peaks, we extracted:

- A distance measure—the Euclidean distances (L2-norm version of the Minkowski distance) between the four microstate cluster centroids and the EEG scalp topology, at each time point (including the time between GFP peaks)
- The path-step length—the Euclidean distances between each pair of consecutive time points along the trajectory of the EEG scalp topologies over time

To characterize the complex dynamics in microstate transitions, the fractal dimension, phase space, and Lyapunov exponents of the distance measure and path-step length (defined above) were computed using a sliding window of 256 ms and a step size of 2 ms.

2.2.2 Fractal Dimension. Fractal distance (FD) is a method to quantify the level of complexity in time series data. Patterns that have a fractal structure maintain similar structure across different scales. Fractal structure is abundant in physiological signals, and fractal dimension characterizes the degree of complexity in the data (Eke, Hermán, Bassingthwaighte et al., 2000).

Numerous algorithms exist to estimate the level of fractality in a data set, one of which is the box-counting method (Eke, Hermán, Kocsis, & Kozak, 2002). This method tries to cover the data set (normalized to unit length) using nonoverlapping boxes of progressively smaller edge lengths (ϵ) (starting with the trivial case of one box for the whole data set), and counting the minimum number of boxes required (N_ϵ) to cover the entire data set for each edge length. If the complexity of the data set is low (i.e., a low topological dimension), the number of boxes required to cover the data set is on the order $1/\epsilon$ for each ϵ. However, if the data set has a high topological dimension, the required number of boxes will be $1/\epsilon^d$, for each ϵ, where $d > 1$ is the topological dimension:

$$N_\epsilon = 1/\epsilon^d. \tag{2.2}$$

This is computationally estimated by embedding the data set in the unit box (box of unit length), normalizing the time series, and then computing N_ϵ for progressively smaller boxes ϵ. The fractal dimension d is then given as the slope of the log of N_ϵ as a function of the log of $1/\epsilon$:

$$d = \frac{\log N_\epsilon}{\log 1/\epsilon}. \tag{2.3}$$

The fractal dimension is also referred to as the capacity dimension or counting dimension and is a practical method to estimate the more general Hausdorff dimension of the data set (Eke, Hermán, Kocsis et al., 2002). Intuitively, the fractal dimension captures the degree to which data exhibit complexity at many different scales.

Compared to other measures of quantifying complexity in biomedical time series data, measures such as FD and sample entropy (Richman & Moorman, 2000) are relatively invariant to the number of data used (window length) for estimating the measure (Ferenets et al., 2006). This is of particular importance in the context of studying microstates due to the short timescale of its dynamics (on the order of a few hundred milliseconds).

2.2.3 Phase Space. Any system might exist in a range of different states and could have different paths of transitioning from one state to another. The phase space of a system represents this state space that the system could potentially occupy and shows the various trajectories the system takes over time (called orbits). The pattern of the phase-space orbits shows the dynamical characteristics of the system (Suzuki, Lu, Ben-Jacob, & Onuchic, 2016). Visualizing the phase space of an oscillatory time series is a useful way to qualitatively understand its temporal dynamics. One way to quantify this is by using Lyapunov exponents, described next.

2.2.4 Lyapunov Exponent. Lyapunov exponent (LE) is a quantitative descriptor of the phase-space trajectory dynamics that estimates the rate of divergence of infinitesimally close phase space trajectories. Since different initial conditions might have different divergence rates, a spectrum of Lyapunov exponents is computed, and the maximum Lyapunov exponent is taken as an upper bound on the divergence rate (Wolf, 1986). Henceforth, when this letter refers to the Lyapunov exponent, we are referring to the maximum Lyapunov exponent. Chaotic phase-space trajectories eventually diverge and, hence, have a positive LE, compared to a negative LE for nonchaotic phase space trajectories (periodic/oscillatory).

Neural circuits are known to produce chaotic dynamics (Skarda & Freeman, 1987; Stam, 2005) and can even switch between periodic and chaotic dynamics (Alonso, 2017). Studying descriptors of chaotic dynamics (such as LE) alongside microstate transitions can help identify the dominant dynamical operating paradigm.

The Jacobian method of computing LE (BenSaïda, 2015) from a scalar time series ($\{x_t\}$) was used in the analyses reported here and is briefly described. Any time series ($\{x_t\}$) can be expressed as a noisy chaotic system in terms of a time delay (L), embedding dimension (m), and added noise (ϵ_t):

$$x_t = f(x_{t-L}, x_{t-2L}, \ldots, x_{t-mL}) + \epsilon_t. \tag{2.4}$$

The function f can be estimated using nonlinear least squares (see equation 2.5) implemented using a neural network with q hidden layers and tanh activation function:

$$x_t \approx \alpha_0 + \sum_{j=1}^{q} \alpha_j \tanh\left(\beta_{0,j} + \sum_{i=1}^{m} \beta_{i,j} x_{t-iL}\right) + \epsilon_t \tag{2.5}$$

This estimate of the function f is used to compute $T_M = \prod_{t=1}^{M-1} J_{M-t}$, a product of Jacobian matrices (J_t):

$$J_t = \begin{bmatrix} \frac{\partial f}{\partial x_{t-L}} & \frac{\partial f}{\partial x_{t-2L}} & \cdots & \frac{\partial f}{\partial x_{t-mL+L}} & \frac{\partial f}{\partial x_{t-mL}} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{2.6}$$

The LE is then estimated as $LE = \frac{1}{2M} \ln(\nu_1)$, where M is an integer that is picked to be $M \approx T^{2/3}$ and ν_1 is the largest eigenvalue of $(T_M U_0)'(T_M U_0)$. This is repeated of each triplet of (L, M, q), generating a spectrum of Lyapunov exponents, from which the largest LE is picked.

To appropriately verify the existence of nonlinear structure in the data, surrogate tests were performed by comparing the LE computed from surrogate data to that derived from actual data using a one-tailed t-test. Only significant ($p > 0.05$) LE values were retained for this analysis.

2.3 Predictions for Microstate Analysis. The traditional view of microstates is that EEG scalp topologies move between discontinuous quasi-stable states with a single active state at any one time. If this view is correct, we predict that the time of the transitions between microstates will be less than the time spent in a given microstate. This would be reflected in lower FD (less complexity), negative LE (nonchaotic behavior), and higher path-step lengths (large jump in state-space) during the microstate transitions between two GFP peaks. However, if there is competition between the different microstates during the transitions, the FD should be higher (higher complexity), with positive LE (chaotic behavior) and shorter path-step lengths (smaller steps) in this time span as the scalp topologies traverse a few different states before finally resulting in one microstate.

Furthermore, if the phase space and LE indicate that the brain is operating under chaotic dynamics, representing it using a small number of discontinuous states might not be appropriate.

2.4 Clustering Analysis. As stated in assumption 3, there is little evidence supporting that microstates are indeed a discrete set of states that can be adequately described by clustering. To investigate this, the k-means clustering algorithm was rerun, varying the number of clusters from two to eight, computing the gap statistic and Davies-Bouldin index for each number of clusters. The high-dimensional clustering space was also embedded into three dimensions using t-SNE (Van Der Maaten & Hinton, 2008) for visualization of the cluster separation and spread. The time course of the data was also plotted within the clustering space to visualize the trajectory of EEG scalp topologies.

2.4.1 Gap Statistic. The gap statistic compares the intracluster dispersion of a clustering solution with that derived from a null distribution (random uniform distribution), defined as

$$Gap_n(k) = E_n^* \left\{ \log W_k \right\} - \log W_k, \tag{2.7}$$

where n is the sample size and k is the number of clusters being evaluated. W_k is a pooled within-cluster dispersion measurement given by

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r, \tag{2.8}$$

where, $D_r = \sum_{i,j=1;i\neq j}^{n_r} \|\mathbf{x_j} - \mathbf{x_i}\|_2$ is the summed Euclidean distance between each pair of points in cluster r containing a total of n_r data points. The optimal number of clusters is identified as the smallest cluster number at which the intracluster dispersion falls farthest below the null distribution, as identified by a formalization of the "elbow method" (Tibshirani, Walther, & Hastie, 2001). Mathematically, this is expressed as identifying the smallest k, such that $Gap(k) \geq Gap(k+1) - std(k+1)$ (Tibshirani et al., 2001). Hence, the appropriate number of clusters occurs at the maximum of the gap statistic, where the distributions found in the data are farthest from that of a random uniform distribution.

2.4.2 Davies-Bouldin Index. The Davies-Bouldin (DB) index is based on the ratio of the within-cluster distances to the between-cluster distances, indicating the spread of the identified clusters. For k clusters, it is defined as

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{i\neq j} \left\{ \frac{(\bar{d}_i + \bar{d}_j)}{d_{i,j}} \right\}, \tag{2.9}$$

where $\bar{d}_i$ and $\bar{d}_j$ represent the average distance between each point in the cluster and the cluster centroid for the ith and jth clusters, and $\bar{d_{i,j}}$ is the Euclidean distance between the centroids of the ith and jth clusters.

Minimizing this index gives the optimal number of clusters with small within-cluster spread and large between-cluster distances.

2.4.3 t-Distributed Stochastic Neighbor Embedding. t-SNE is a dimension-reducing algorithm that embeds a higher-dimensional space into a lower number of dimensions and is particularly useful for visualizing high-dimensional clustering spaces in two or three dimensions (Van Der Maaten & Hinton, 2008). This is an iterative procedure that attempts to maintain the local structure of the high-dimensional space, such that points that are

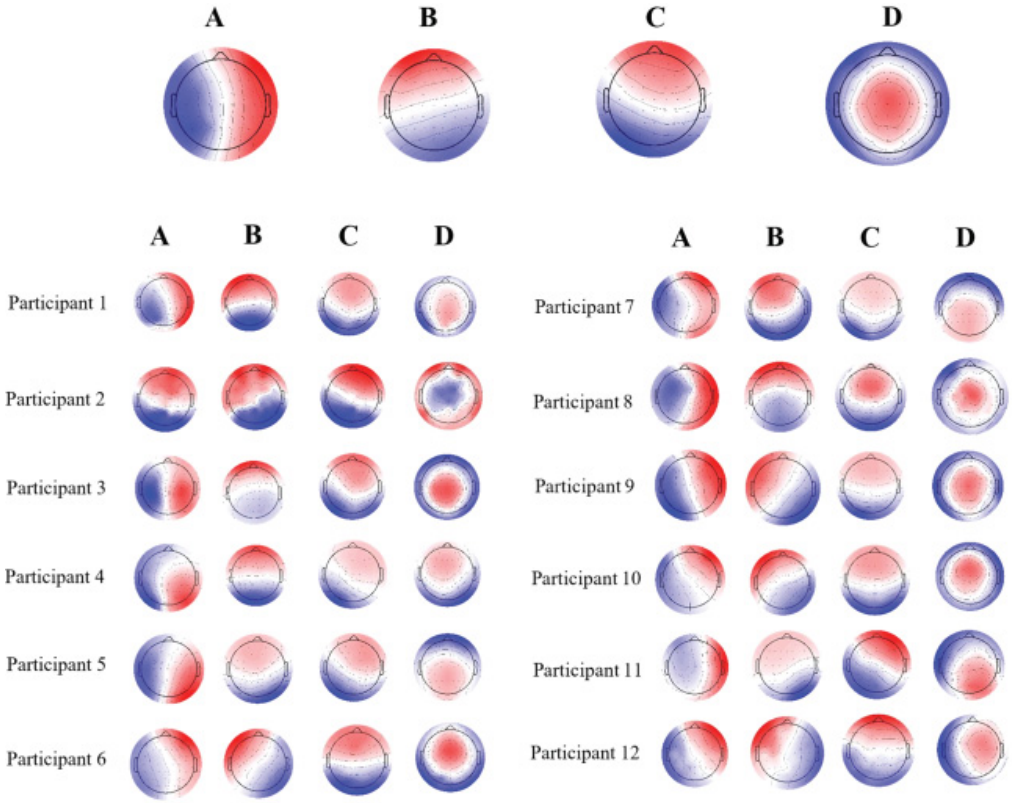

Figure 3: The individual differences in microstate maps identified from the data set, along with the participant level maps. Note the variability between the maps of different participants. The microstate maps are sorted according to the maps shown in Figure 1 (Milz et al., 2015). Note that the microstate maps ignore the polarity of the EEG scalp topology. This is seen in map D of participant 2, where the map polarity is opposite that of other participants. Also, note the similarity between maps A and D of participant 12. This is why they cluster together in the clustering space shown in Figure 12a.

closer together in the original high-dimensional space remain close in the lower-dimensional space. This allows for visualization of the local (and potentially global) structure of the higher-dimensional space. This algorithm was used to visualize the clustering space generated by microstate analysis.

3 Results

3.1 Individual Variability in Identified Microstates. The four microstate class topologies computed from the aggregate rsEEG data of the participants are shown at the top in Figure 3. The topologies identified on a per participant basis are also shown. The large variance in interparticipant microstate class topologies is evident from Figure 3 and is in line with the current literature (Britz et al., 2010).

3.2 Temporal Dynamic Analysis

3.2.1 Distance from Microstate Classes. Visualizing the distance of the GFP scalp topology to each microstate cluster centroid over time and its FD reveals temporal dynamics between GFP peaks and microstate transitions that are inconsistent with the view that microstates are quasi-stable discrete states.

The time series of the distance between the topology at a given time point and the four microstate classes is shown to vary with time, leading to a lower distance from one of the microstate centroids at each GFP peak. However, the other microstates with larger distances still show nontrivial activity at the peaks. Furthermore, the assigned microstate label did not always correspond to the cluster with the lowest distance, as seen in Table 2. Microstates A, B, C, and D were closest to the respective cluster centroids only 50%, 46%, 44%, and 38% of the time, respectively. Time points labeled as microstate A were often closest to the centroids of microstates B (19%) and D (20%), while 25% of the time points labeled as B were closest to microstate A centroid. Time points labeled as microstate C were found to be closest to the centroids of microstate D and B 25% and 22% of the time, respectively, whereas 26% of the time points labeled as microstate D were closest to microstate A centroid. Also, the distances of the GFP peak topology from the microstate classes were sometimes equal, with no clear winner. These results might stem from competition between the different microstates, suggesting the inadequacy of a winner-takes-all model to describe the dynamics of rsEEG. This is shown in Figure 4 and summarized in Table 3.

Fractal dimension. The fractal dimension (FD) of the microstate distances is observed to be higher (1.38 ± 0.11) than the baseline FD (equal to 1, as described in section 2.2.2), not only at GFP peaks as suggested by assumption 1, but also between the GFP peaks (see Figure 5). This indicates dynamic events occurring between GFP peaks that could contain useful information regarding the ongoing dynamics of rsEEG. However, microstate analysis could be missing these dynamic events, given that it considers only GFP peaks. This might explain the poor reliability and nonstationarity of microstates at determining short-range dynamics.

3.2.2 Path-Step Length. The step lengths of the EEG scalp topology trajectory (see Figure 6) show a more continuous traversal of the EEG scalp topology space than is implied by the discontinuous microstate description.

The lengths were found to be mostly small (0.09 ± 0.04), with some occasional peaks. The small step path lengths indicate a smooth traversal of most of the EEG scalp topology space. The majority (71.8%) of the peaks in step path lengths were located close to the GFP peaks and associated with a polarity reversal in scalp topology rather than following microstate attractor-like behavior expected to cause peaks in step path lengths close to

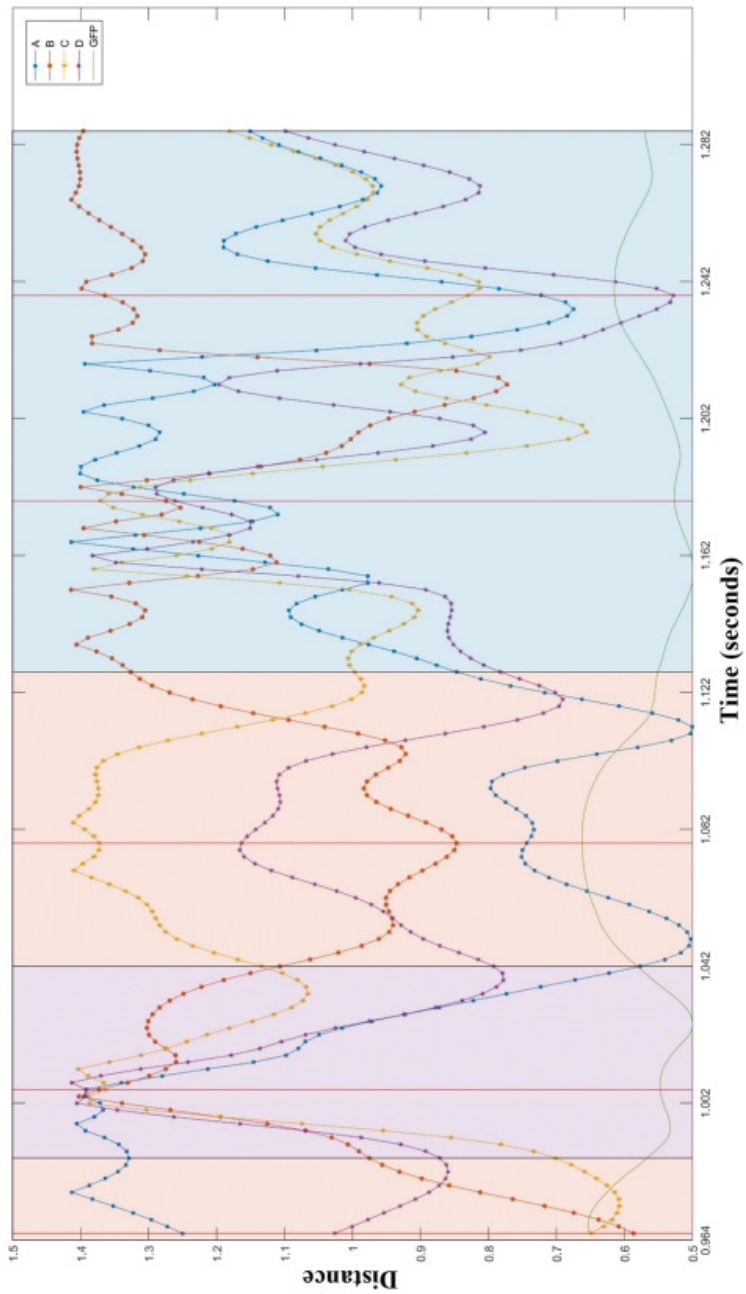

Figure 4: The distance between the peak topology and the four microstates is shown: microstate A (blue square), B (orange circle), C (yellow diamond), D (purple star). The GFP curve (green) used to identify microstates is also shown, along with its peaks (red vertical lines). The microstate label assigned by the microstate algorithm is indicated by the color of the background. The assigned microstate label does not always correspond to the microstate with the lowest distance from each time point. Furthermore, the distance of the scalp topology from the nonproximal (minimum distance) microstates is nontrivial and ranks differently from one GFP peak to another. This might be indicative of more complex dynamics of rsEEG, which is being missed by the winner-takes-all model.

Table 3: Distribution of Distances from the Time Points Labeled as a Particular Microstate to the Cluster Centroids of All Microstate Classes.

Closest	Labeled Microstate Class															
Microstate	A				B				C				D			
Class	**A**	B	C	D	A	**B**	C	D	A	B	**C**	D	A	B	C	**D**
1 (closest)	**50**	19	10	20	25	**46**	18	11	10	22	**44**	25	26	13	22	**38**
2	**23**	26	16	35	28	**25**	28	18	12	18	**34**	35	22	14	29	**34**
3	**14**	24	31	30	20	**14**	30	36	30	30	**14**	26	29	23	30	**18**
4 (Farthest)	**12**	31	42	15	26	**15**	24	35	48	30	**8**	14	23	49	18	**10**

Notes: This is expressed in terms of the percentage of time points that ranked closest (shortest distance) to a microstate cluster, while being labeled a particular microstate class. If data were well clustered, the closest microstate cluster should be the same as the labeled class (marked in bold). However, as evident from the table, a sizable proportion of time points show proximity to microstate clusters other than their labeled class. This can be due to the poor separation of the clusters, as shown in Figure 12a. The Euclidean distance (L2-norm version of the Minkowski distance) was used for this computation.

microstate transitions (seen only in 28.2% of step path length peaks, shown in Table 4).

Fractal dimension. The FD of the step length is found to be high (1.57 ± 0.13), indicating many small changes in step lengths throughout the time duration studied.

Phase space. The phase space of the step lengths of the EEG scalp topology trajectory (shown in Figure 7) resembles that of a chaotic system (Suzuki et al., 2016). This is confirmed by the positive Lyapunov exponents seen in Figure 8. Some regions of negative Lyapunov exponents further suggest periods of nonchaotic behavior interspersed between chaotic dynamics.

Lyapunov exponent. The LE of the step length (see Figure 8) is found to be mostly positive, with some periods of negative values, indicating chaotic dynamics interspersed with some periods of nonchaotic behavior.

3.3 EEG Scalp Topologies. Qualitatively, the EEG topology was found to vary smoothly between different scalp topologies (shown as a montage in Figure 9). Furthermore, the trajectory of EEG topology within the clustering space (see Figure 12b) seems to have smooth segments (short path-step length) as well as sharp jumps (large path-step length). These sharp jumps are found to be primarily located away from microstate transitions between different states (see Table 4) and correspond to within-microstate transitions, that is, transition from one polarity of the microstate class topology to the opposite polarity of the same microstate class. This is seen in the first two microstates shown in Figure 9, where the EEG topology flips in polarity and stays in microstate B. This in contrast to the discontinuous behavior suggested by the quasi-stable description of EEG microstates, which

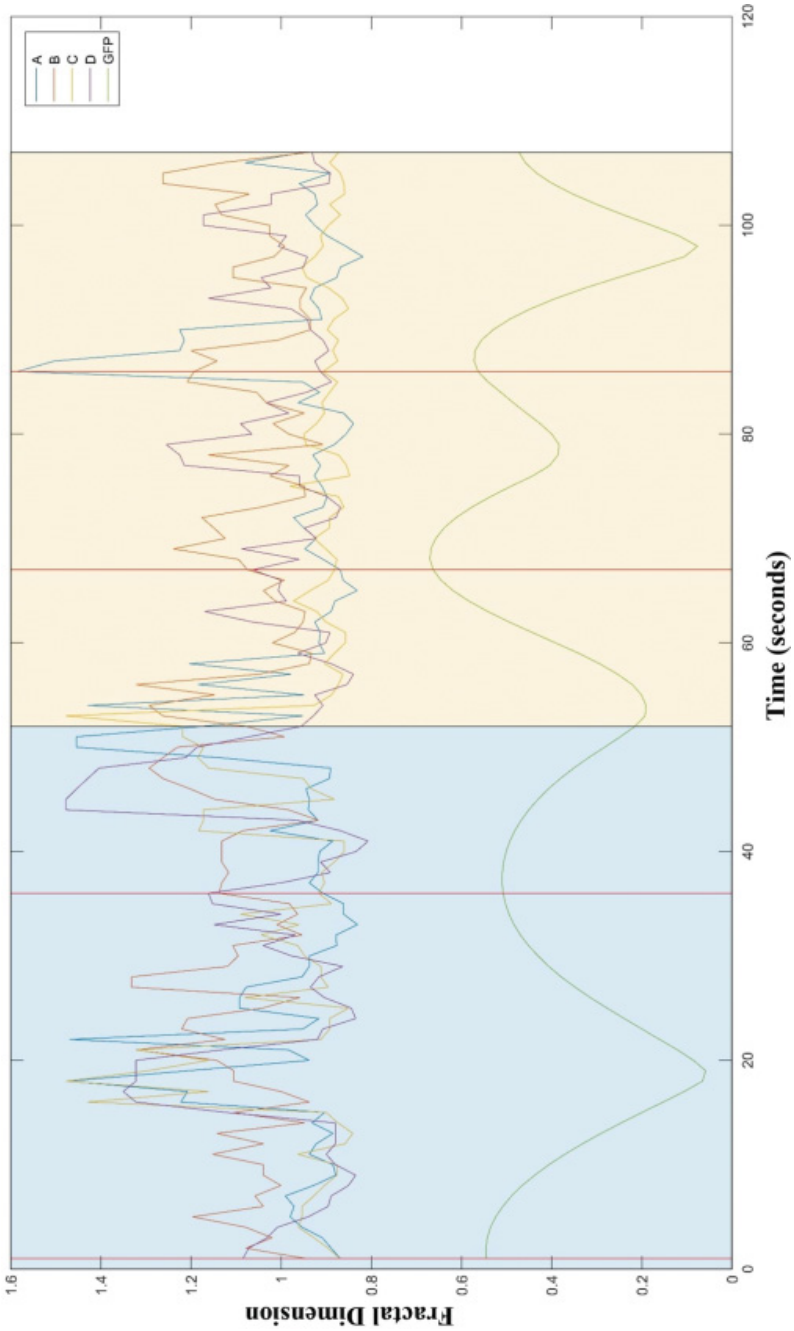

Figure 5: The fractal dimension (FD) of distance between the peak topology and the four microstates is shown: microstate A (blue), B (orange), C (yellow), D (purple). The GFP curve (green) used to identify microstates is also shown along with its peaks (red vertical lines). The microstate label assigned by the microstate algorithm is indicated by the color of the background. An increase in FD is noticed between GFP peaks, indicating an increase in complex dynamics during this timescale that might be missed by focusing only on GFP peaks.

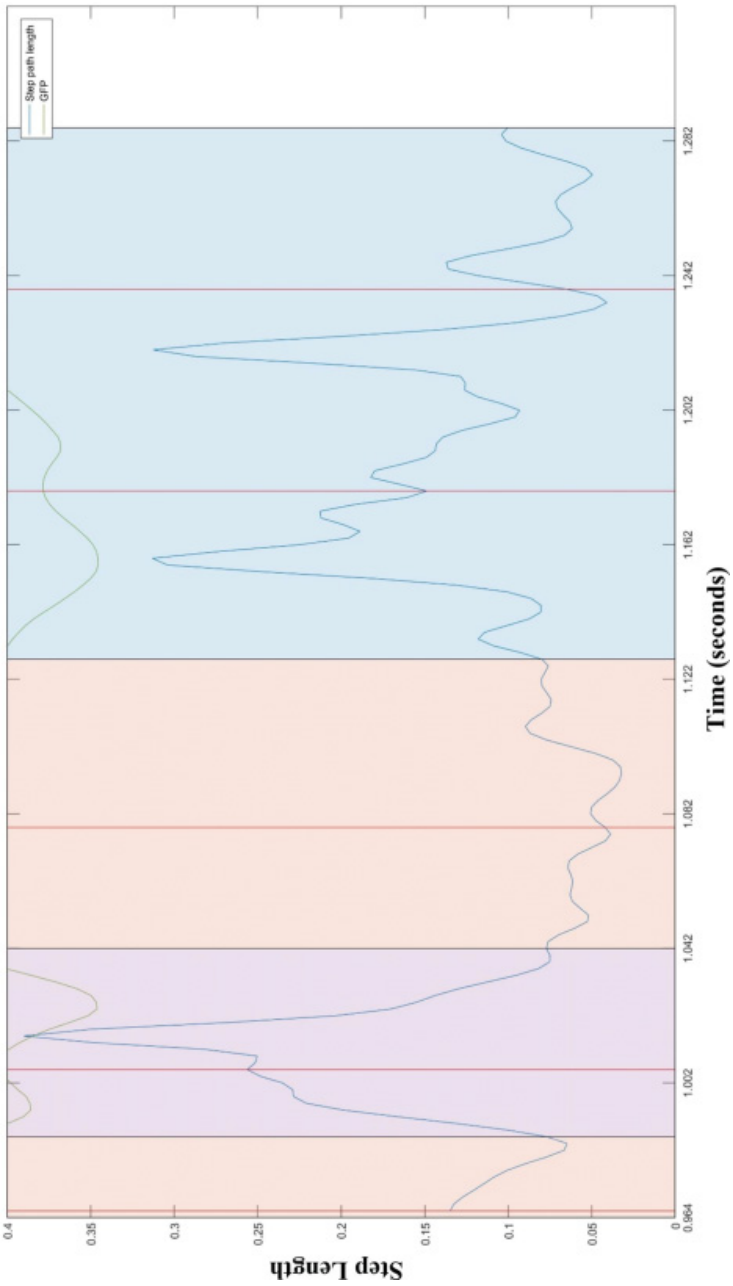

Figure 6: The step lengths of the EEG scalp topology trajectory plotted over a short time course. The GFP curve (green) used to identify microstates is also shown along with its peaks (red vertical lines). The microstate label assigned by the microstate algorithm is indicated by the color of the background: microstate A (blue), B (orange), C (yellow—not present in this example), D (purple). An increase in step length is noticed around GFP peaks, indicating a large change in the EEG scalp topology (seen in Figure 12b). Most of the peak lengths are small and correspond to smooth traversal of the clustering space. The peaks in step lengths occur close to GFP peaks and correspond to reversal of EEG polarity, meaning that, by definition, it remains in the same microstate (as seen in the last two GFP peaks in this figure).

Table 4: Distribution of Path-Step Lengths Plotted in Figure 6.

Metric	During Microstate Transition	Not During Microstate Transition
Latency (in ms) from step-path peaks	5.7 ± 5.7	—
Percentage of step-path peaks	28.2	71.8

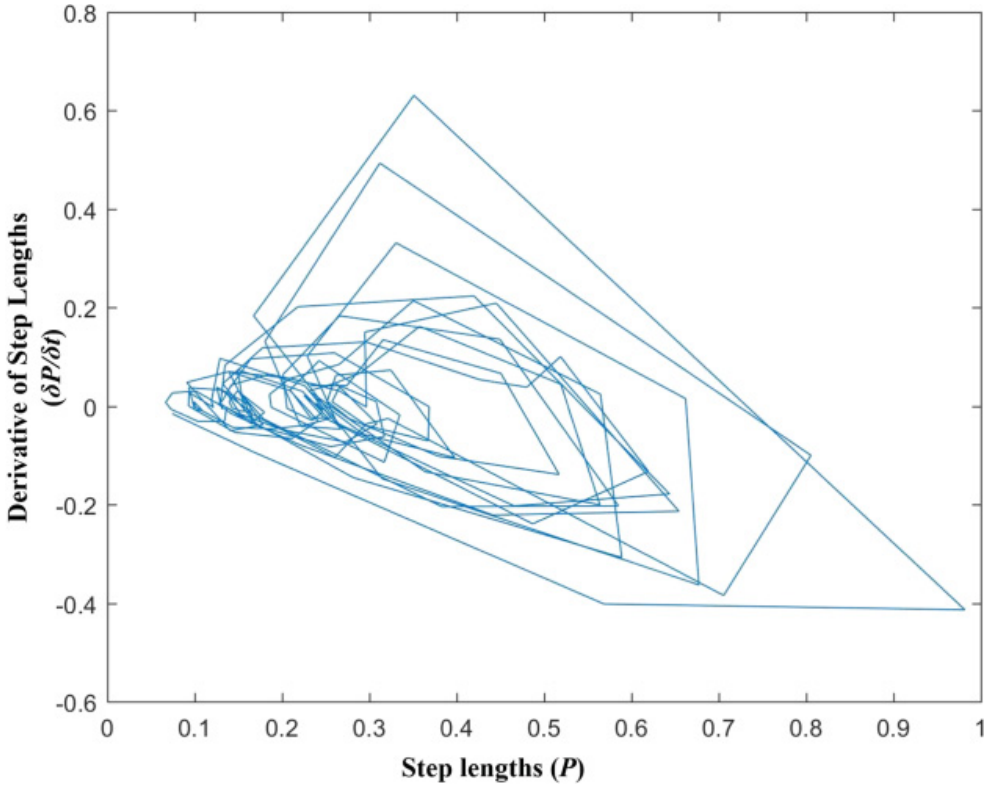

Figure 7: The phase space of the step lengths of the EEG scalp topology trajectory plotted over a short time course. The orbits of this phase space resemble that of a chaotic system (Suzuki et al., 2016).

would have shown abrupt and discontinuous variation in scalp topologies between different microstates rather than transitions within the same microstate.

3.4 Clustering Analysis. The gap statistic and Davies-Bouldin index for varying numbers of clusters are shown in Figures 10 and 11, respectively. The gap statistic values for the widely used cluster number of four show that it is close to the clustering structure derived from a random uniform distribution (since it is close to unity), indicating that representing the EEG sequence by four microstate clusters constitutes an inadequate description

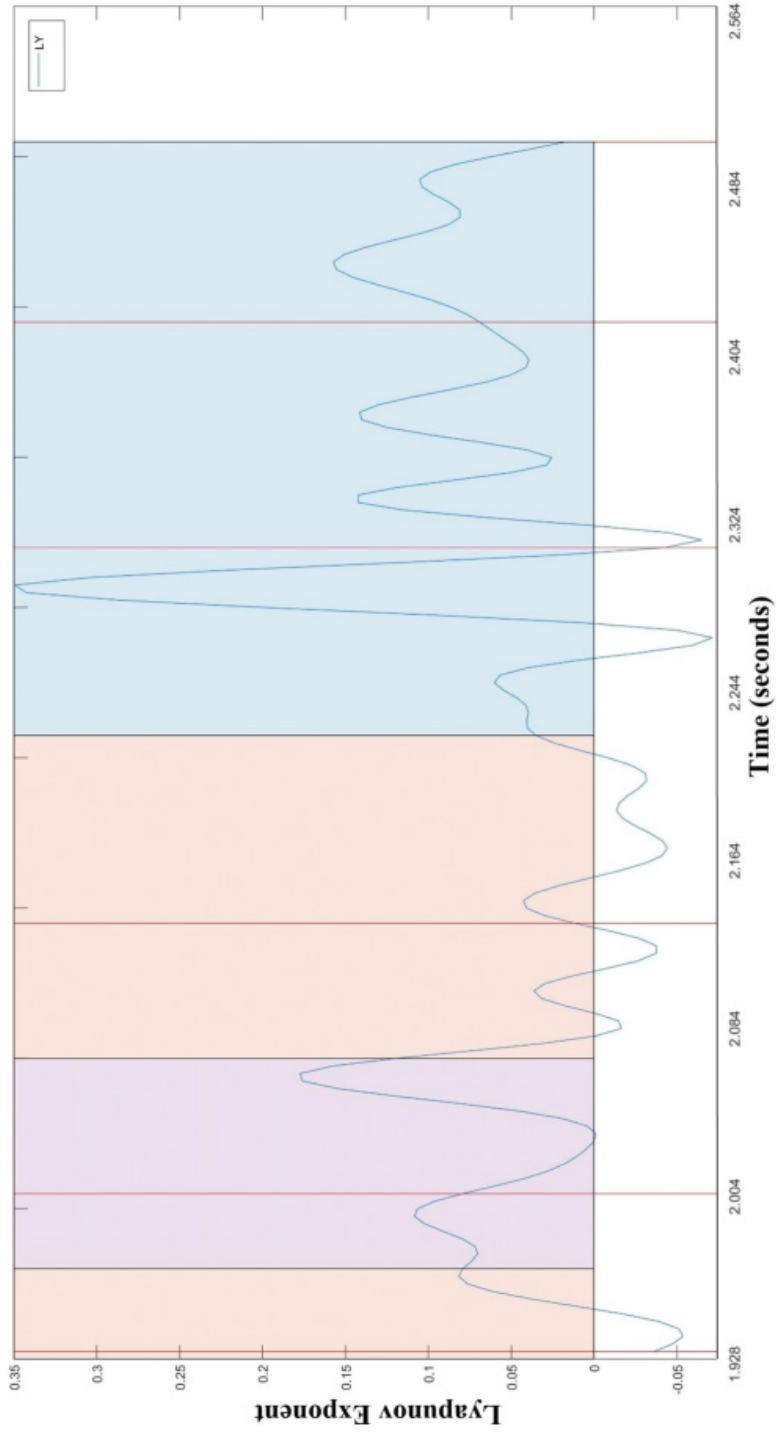

Figure 8: The Lyapunov exponent of the step lengths of the EEG scalp topology trajectory (blue) shown over a short time course. The GFP curve peaks (red vertical lines) used to identify microstates are also shown. The microstate label assigned by the microstate algorithm is indicated by the color of the background: microstate A (blue), B (orange), C (yellow), D (purple). The LE is found to be primarily positive, indicating chaotic dynamics. There also are some regions of negative LE, indicating switching between chaotic and nonchaotic dynamical paradigms.

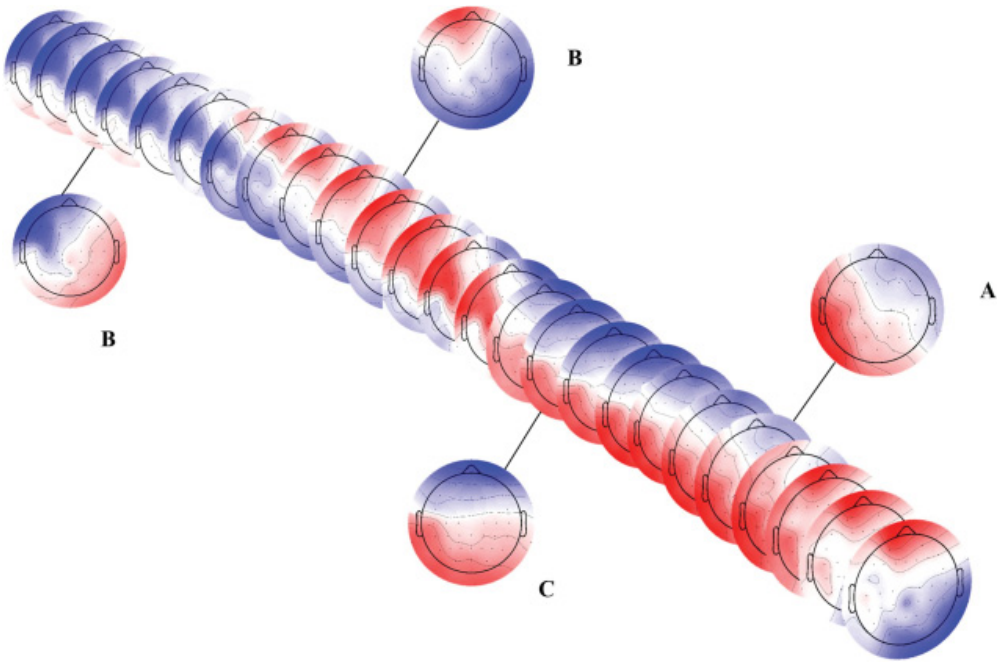

Figure 9: Chronological progression of the EEG topographic maps, plotted over the course of 300 ms. The EEG topologies at the GFP peaks during this time segment are shown separately, along with the classified microstate class. Note that the topologies smoothly vary over time, in contrast to the abrupt changes suggested by the discontinuous microstate model. Furthermore, note that the first two GFP peak topologies are opposite in polarity and are classified into the same microstate class (B), since microstate analysis ignores EEG polarity.

of the data at the GFP peaks. The Davies-Bouldin index also shows that four clusters is nonoptimal, when considering the within-cluster spread to between-cluster distance, indicating that there may be overlap between the clusters identified. Both methods identified two clusters as being the most appropriate clustering solution. Visualization of the clustering space using t-SNE (see Figure 12a) illustrates why four clusters is not a good description of the data. The EEG scalp topology data forms a ring structure, where opposite polarities of the same microstate classes are seen on opposite ends of the ring. The wide spread in the clusters and poor separation between different clusters are detected by the Davies-Bouldin index, showing poor clustering structure within the EEG scalp topology data.

4 Discussion

Studying the dynamics of brain activity at a finer time resolution (on the order of hundreds of milliseconds) can be extremely useful in uncovering the mechanisms of behavior and cognition in healthy and diseased populations. Microstate analysis is a relatively new method that attempts to

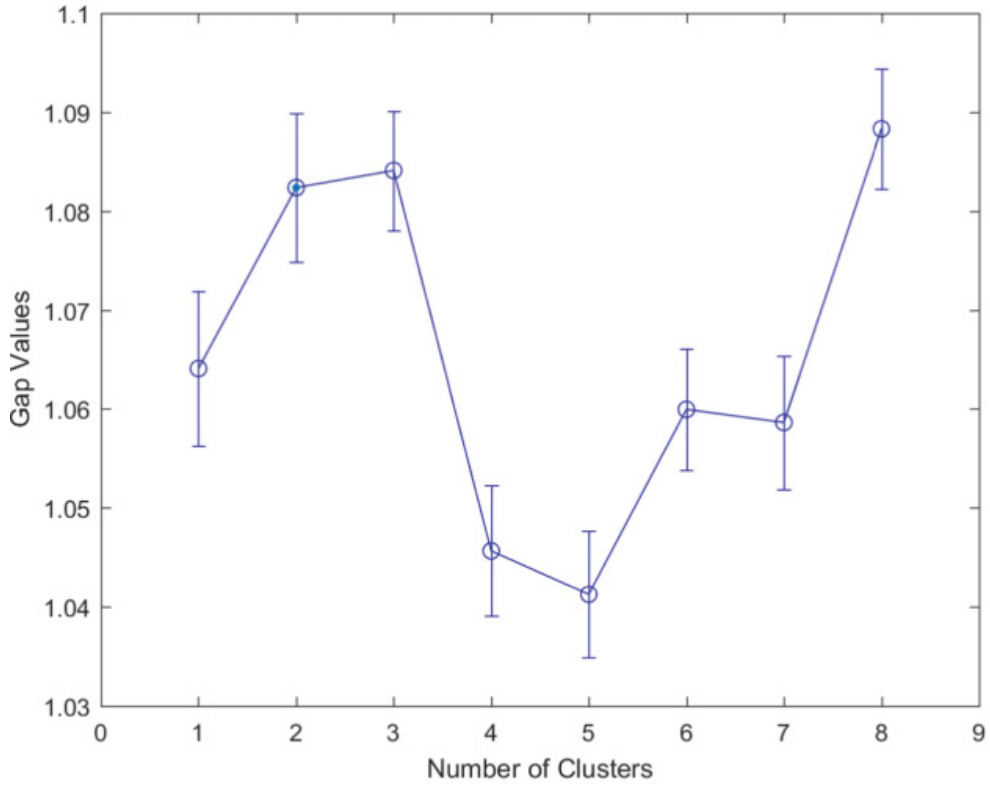

Figure 10: The gap statistic is shown for varying cluster numbers (2–8), with the optimal value identified as 2 to 3. The optimal cluster number is the one that maximizes the gap statistic, indicating maximal deviation from a random uniform clustering pattern. Note that the widely used cluster number of 4 is nonoptimal.

describe complex brain dynamics by sequences of discrete patterns in EEG scalp topology, termed *microstates*. In doing so, it makes some key assumptions about the spatiotemporal dynamics of EEG data, which have not been carefully scrutinized. Here, we report the results of a series of nonlinear dynamic analyses to test some of these basic assumptions and found that they do not hold up to scrutiny.

Assumption 1: GFP peaks. Complex dynamics does exist between GFP peaks and is being missed by microstate analysis One of the primary assumptions inherent in the computation of EEG microstates is that scalp topology at the GFP peaks defines approximately four quasi-stable states that the brain spends most of its time in, switching between the states every 60 to 120 ms (Michel & Koenig, 2018). If this assumption is correct, the EEG topology over time should not vary smoothly, spending most of its time in a particular microstate class topology and then rapidly accelerating toward another microstate class topology while spending minimal time in the transition. However, EEG scalp topologies are observed to vary smoothly, as

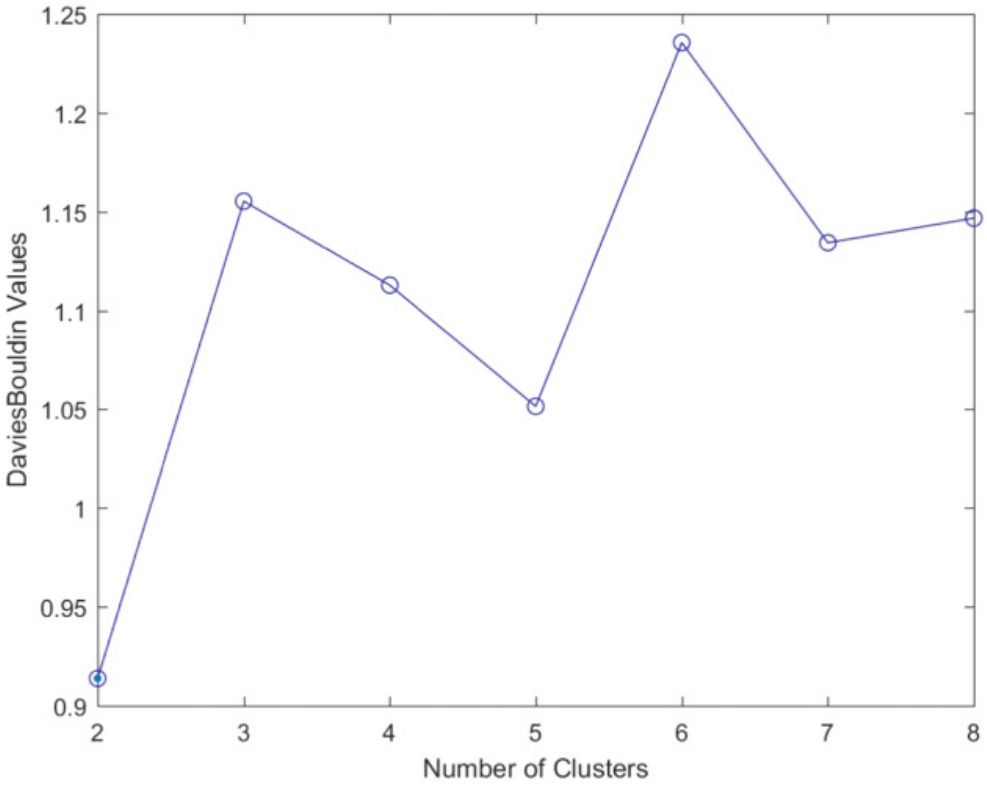

Figure 11: The Davies-Bouldin index is shown for varying cluster numbers (2–8), with the optimal value identified as 2. The optimal cluster number is the one that minimizes this index, indicating minimal within-cluster spread and maximal distance between the clusters. The widely used cluster number of 4 is found to be nonoptimal, with the closest local-minimum occurring at 5 clusters.

evident from the EEG scalp topologies over time (see Figure 9), the step-path length (see Figure 6), and the trajectory of EEG scalp topology in the clustering space (see Figure 12b). Most of the abrupt changes in scalp topology are found to represent polarity reversals within the same microstate (70%) rather than transitions between different microstate classes (see Table 4). Furthermore, despite the GFP peaks lasting only a short duration (full-width-half-maximum $= 22$ ms $\pm$ 20 ms), the microstate assigned at the GFP peak is assumed to be quasi-stable for a much longer duration before and after the GFP peak, thereby missing EEG dynamics between the peaks (shown in Figures 4 and 5).

Assumption 2: One microstate at a time. The winner-takes-all approach might not be appropriate due to observed competition between the microstate classes The clustering assumption implies a winner-takes-all approach. This is appropriate where there is good separation between the classes and physiologically appropriate when only one class is active at one time. However, previous evidence suggests significant overlap between

Figure 12: (a) Microstate clustering space visualized in three dimensions using t-SNE (Van Der Maaten & Hinton, 2008). The clustering space forms a ring structure that has opposite-polarity maps of the same microstate class on opposite ends of the ring structure. Poor separation of the four microstate classes is seen—A (blue square), B (orange circle), C (yellow diamond), D (purple star)—resulting in the poor performance of the four-cluster solution with the Davies-Bouldin index and the gap statistic. Also, note that maps A and D cluster together due to the similarity between the identified scalp topology maps for microstates A and D in participant 12 (see Figure 3). (b) The trajectory of EEG scalp topology maps (green trace) over a 300 ms time course, visualized within the microstate clustering space. The trajectory has smooth sections with small step sizes, showing the smooth traversal of EEG scalp topology space. However, the trajectory also has some large jumps that correspond to an abrupt change in scalp polarity. This is seen to primarily occur in the absence of microstate transitions (see Table 4).

the classes (Custo et al., 2017) and suggests no reason for the regions identified in Table 1 to remain mutually exclusive. This can also be seen in the distance from microstate centroids over time (see Figure 4), where the distances from cluster centroids are sometimes equal at the GFP peaks, showing competition between the different microstate classes. This is further supported by the data in Table 2, which show that the EEG scalp topology is not always closest to its labeled microstate class centroid. This can happen if the closer cluster is much more compact and is assigned a lower probability. For most clustering methods, distance to class centroids is not linearly correlated with the probability of belonging to that cluster, suggesting that if a pattern belongs to one microstate, it might not be the microstate it is most similar to. However, this goes against the winner-takes-all approach, assigning the microstate class on the basis of the cluster it is most similar to. Furthermore, visualizing the trajectory of the EEG scalp topologies (see Figure 12b) in the clustering space shows the competition between the different microstate classes, as the scalp topology smoothly traverses through the space, occasionally jumping to the opposite polarity. The Davies-Bouldin index (see Figure 11) and the clustering space visualization (see Figure 12a) also indicate that the four-class data structure is nonoptimal because of overlap between clusters, further suggesting that the winner-takes-all approach is inadequate for EEG scalp topologies.

Assumption 3: Clustering. The different clusters representing the four microstate classes have significant overlap and inhomogeneity Another assumption made is the existence of clusters in the EEG scalp topologies at GFP peaks. However, visualizing the clustering space in Figure 12a shows significant overlap between these clusters, as supported by the quantitative results of the gap statistic (see Figure 10) and Davies-Bouldin index (see Figure 11). This is also supported by the large variance observed in the clustered scalp topologies between participants (see Figure 3) and the nonstationarity of EEG data, suggesting that the microstate class topologies are not stable and might vary with time. This is supported by the gap statistic results (see Figure 10), showing that a four-class EEG clustering is most similar to clusters in a random uniform data set, most likely causing the large variance seen in the clustered scalp topologies and the nonstationary behavior. This suggests that a discontinuous description of four stable classes that the brain switches between might not be accurate.

EEG chaotic dynamics might contribute to the nonstationarity observed in microstate analyses over short timescales Stemming from the assumption of a discrete winner-takes-all approach resulting in one microstate active at a time (see assumption 2) is yet another assumption: the Markov property of microstates. This assumption has been previously examined and consequently is not explicitly studied in this letter. Numerous studies using microstate analysis have assumed that microstate transitions portray first-order Markovian behavior and have computed transition matrices accordingly. However, recent evidence suggests that this Markovian

assumption is appropriate only for longer timescales and falls apart over short time ranges. Von Wegner et al. (2017) find a lack of Markov property (of any order) and stationarity in short-range microstate transitions. This implicates the existence of either complex or chaotic microstate behavior that translates to long-range dependencies over longer timescales (Gschwind et al., 2015; Van de Ville et al., 2010). This is in line with our FD (see Figure 5) and phase space (see Figure 7) results, which suggest complex dynamics (peak in FD) while switching between microstate classes, in the periods between GFP peaks. Switching of the dominant dynamical paradigm (see Figure 8) from chaotic to nonchaotic and back to chaotic may contribute to the instability of microstate transitions over shorter timescales. This nonstationarity might also arise from the discontinuous parsing of EEG data into microstates when EEG data are not inherently discontinuous.

All of the evidence we have presented goes against the assumption of quasi-stable discontinuous behavior inherent in microstate analysis and is in agreement with a large body of work suggesting that EEG data have complex dynamics and can exhibit chaotic behavior, leading to nonstationarities in the data. Using techniques from dynamical systems analysis, some studies have embedded EEG data in state-space to study general EEG dynamics (Wackermann, 1999) and synchronization between hubs of EEG activity (Carmeli, Knyazeva, Innocenti, & De Feo, 2005). Other nonlinear chaotic descriptors of EEG data have also been used to characterize dynamics of numerous brain states (Stam, 2005). These include fractal dimension (Nan & Jinghua, 1988), chaotic Lyapunov exponents (Natarajan et al., 2004), and entropy of EEG data (Kannathal, Choo, Acharya, & Sadasivan, 2005; Mizuno et al., 2010). Entropy has also proven to be a useful predictor of a participant's ability to control a brain-computer interface (BCI)—the BCI inefficiency effect (Zhang et al., 2015).

In addition to the questionable assumptions inherent in microstate analyses, another issue is the lack of correspondence between the four commonly identified microstates and the core functional brain networks well characterized in resting-state fMRI. In particular, given the predominence of default mode network (DMN) activity in rsfMRI, one would expect to see DMN activity in rsEEG. However, there is conflicting evidence regarding the correlation of the DMN with EEG microstates. Some studies find no correlation between any of the four traditional EEG microstates and DMN activity (Britz et al., 2010), which seems to generate its own unique EEG scalp topologies (Panda et al., 2016). Other studies have found that microstate class C has some link with the DMN (Custo et al., 2017) or that DMN activity is distributed over all four microstate classes and is not specific to any one class (Pascual-Marqui et al., 2014). This might be due to the assumption of only one microstate being active at any one time, whereas the DMN is shown to be the predominantly active network during rsEEG. One other reason for this might be the inadequacy of microstate analysis to study short

timescale variations (Gschwind et al., 2015), since the DMN shows multiple temporal signatures when studied at high temporal resolution (Gschwind et al., 2016; Panda et al., 2016). This could be missed by microstate analysis, which is better able to capture longer-range dependencies than fine timescale spatiotemporal dynamics. Yet another reason for microstates inadequately describing the DMN could be its inability to capture common or latent inputs that modulate the behavior between two brain regions, as seen in the DMN, where the posterior cingulate cortex (PCC) and ventral anterior cingulate cortex (vACC) nodes behave differently when coupled with other networks in the brain (Uddin et al., 2009; Das et al., 2017).

5 Conclusion

We have found that several assumptions underlying the extraction and interpretation of EEG microstates do not hold under empirical investigation, and as a result, microstates may not be an accurate description of the temporal dynamics of the EEG. Although the discontinuous behavior suggested by the microstate model might capture some information relating to the global scalp topologies over longer timescales, it seems to be inadequate to describe the nonstationary and chaotic nature of EEG data over shorter timescales. Consequently, those using this method should be cautioned that it may severely underrepresent, or miss entirely, the detailed complex spatiotemporal dynamics that are crucial to many research questions and applications of EEG. Future work should focus on developing better dynamical methods that can capture such complex behavior from continuous EEG data.

Acknowledgments

This work was supported by an NSERC Discovery grant to S.B., an NSERC discovery grant to J.R., and an NSERC CGS-D scholarship to S.B.S.

References

Alonso, L. M. (2017). Nonlinear resonances and multi-stability in simple neural circuits. *Chaos*, *27*(1), 013118. doi:10.1063/1.4974028

BenSaïda, A. (2015). A practical test for noisy chaotic dynamics. *SoftwareX*, *3-4*, 1–5. doi:10.1016/j.softx.2015.08.002

Britz, J., Van De Ville, D., & Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, *52*(4), 1162–1170. doi:10.1016/j.neuroimage.2010.02.052

Brodbeck, V., Kuhn, A., von Wegner, F., Morzelewski, A., Tagliazucchi, E., Borisov, S., . . . Laufs, H. (2012). EEG microstates of wakefulness and NREM sleep. *NeuroImage*, *62*(3), 2129–2139. doi:10.1016/j.neuroimage.2012.05.060

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*, 813870. doi:10.1155/2011/813870

Carmeli, C., Knyazeva, M. G., Innocenti, G. M., & De Feo, O. (2005). Assessment of EEG synchronization based on state-space analysis. *NeuroImage*, *25*(2), 339–354. doi:10.1016/j.neuroimage.2004.11.049

Corradini, P. L., & Persinger, M. A. (2014). Spectral power, source localization and microstates to quantify chronic deficits from "mild" closed head injury: Correlation with classic neuropsychological tests. *Brain Injury*, *28*(10), 1317–1327. doi:10.3109/02699052.2014.916819

Custo, A., van der Ville, D., Wells, W. M., Tomescu, I. M., & Michel, C. (2017). EEG resting-state networks: Microstates source localization. *Brain Connectivity*, *7*(10), 671–682. doi:10.1089/brain.2016.0476

Das, A., Sampson, A. L., Lainscsek, C., Muller, L., Lin, W., Doyle, J. C., . . . Sejnowski, T. J. (2017). Interpretation of the precision matrix and its application in estimating sparse brain connectivity during sleep spindles from human electrocorticography recordings. *Neural Computation*, *29*(3), 603–642. doi:10.1162/NECO_a_00936

Diaz Hernandez, L., Rieger, K., Baenninger, A., Brandeis, D., & Koenig, T. (2016). Towards using microstate-neurofeedback for the treatment of psychotic symptoms in schizophrenia: A feasibility study in healthy participants. *Brain Topography*, *29*(2), 308–321. doi:10.1007/s10548-015-0460-4

Eke, A., Hermán, P., Bassingthwaighte, J., Raymond, G., Percival, D., Cannon, M., . . . Ikrényi, C. (2000). Physiological time series: Distinguishing fractal noises from motions. *Pflügers Archiv—European Journal of Physiology*, *439*(4), 403–415. doi:10.1007/s004249900135

Eke, A., Hermán, P., Kocsis, L., & Kozak, L. R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement*, *23*(1), 1–38. http://www.ncbi.nlm.nih.gov/pubmed/11876246

Faber, P. L., Travis, F., Milz, P., & Parim, N. (2017). EEG microstates during different phases of transcendental meditation practice. *Cognitive Processing*, *18*, 307–314. doi:10.1007/s10339-017-0812-y

Ferenets, R., Lipping, T., Anier, A., Jäntti, V., Melto, S., & Hovilehto, S. (2006). Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Transactions on Biomedical Engineering*, *53*(6), 1067–1077. doi:10.1109/TBME.2006.873543

Gärtner, M., Brodbeck, V., Laufs, H., & Schneider, G. (2015). A stochastic model for EEG microstate sequence analysis. *NeuroImage*, *104*, 199–208. doi:10.1016/j.neuroimage.2014.10.014

Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., . . . Schatzberg, A. F. (2007). Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, *62*(5), 429–437. doi:10.1016/j.biopsych.2006.09.020

Gschwind, M., Hardmeier, M., Van De Ville, D., Tomescu, M. I., Penner, I. K., Naegelin, Y., . . . Seeck, M. (2016). Fluctuations of spontaneous EEG topographies predict disease state in relapsing-remitting multiple sclerosis. *NeuroImage: Clinical*, *12*, 466–477. doi:10.1016/j.nicl.2016.08.008

Gschwind, M., Michel, C. M., & Van De Ville, D. (2015). Long-range dependencies make the difference—Comment on "A stochastic model for EEG microstate sequence analysis." *NeuroImage*, *117*, 449–455. doi:10.1016/j.neuroimage.2015.05.062

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., . . . Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, *80*, 360–378. doi:10.1016/j.neuroimage.2013.05.079

Kannathal, N., Choo, M. L., Acharya, U. R., & Sadasivan, P. K. (2005). Entropies for detection of epilepsy in EEG. *Computer Methods and Programs in Biomedicine*, *80*(3), 187–194. doi:10.1016/j.cmpb.2005.06.012

Khanna, A., Pascual-Leone, A., & Farzan, F. (2014). Reliability of resting-state microstate features in electroencephalography. *PLoS ONE*, *9*(12), 1–21. doi:10.1371/journal.pone.0114163

Khanna, A., Pascual-Leone, A., Michel, C. M., & Farzan, F. (2015). Microstates in resting-state EEG: Current status and future directions. *Neuroscience and Biobehavioral Reviews*, *49*, 105–113. doi:10.1016/j.neubiorev.2014.12.010

Kim, S.-g., Richter, W., & Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, *37*(4), 631–636. doi:10.1002/mrm.1910370427

Koenig, T., Kochi, K., & Lehmann, D. (1998). Event-related electric microstates of the brain differ between words with visual and abstract meaning. *Electroencephalography and Clinical Neurophysiology*, *106*(6), 535–546. doi:10.1016/S0013-4694(97)00164-8

Koenig, T., Prichep, L., Lehmann, D., Sosa, P. V., Braeker, E., Kleinlogel, H., . . . John, E. (2002). Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages. *NeuroImage*, *16*(1), 41–48. doi:10.1006/nimg.2002.1070

Lehmann, D. (1971). Multichannel topography of human alpha EEG fields. *Electroencephalography and Clinical Neurophysiology*, *31*(5), 439–449. doi:10.1016/0013-4694(71)90165-9

Lehmann, D. (1990). Brain electric microstates and cognition: The atoms of thought. In E. R. John, T. Harmony, L. S. Prichep, M. Valdés-Sosa, & P. A. Valdés-Sosa (Eds.), *Machinery of the mind* (pp. 209–224). Boston: Birkhäuser. doi:10.1007/978-1-4757-1083-0_10

Lehmann, D., Faber, P. L., Galderisi, S., Herrmann, W. M., Kinoshita, T., Koukkou, M., . . . Koenig, T. (2005). EEG microstate duration and syntax in acute, medication-naıve, first-episode schizophrenia: A multi-center study. *Psychiatry Research—Neuroimaging*, *138*(2), 141–156. doi:10.1016/j.pscychresns.2004.05.007

Lehmann, D., & Koenig, T. (1997). Spatio-temporal dynamics of alpha brain electric fields, and cognitive modes. *International Journal of Psychophysiology*, *26*(1–3), 99–112. doi:10.1016/S0167-8760(97)00758-7

Lehmann, D., Strik, W. K., Henggeler, B., Koenig, T., & Koukkou, M. (1998). Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *International Journal of Psychophysiology*, *29*(1), 1–11. doi:10.1016/S0167-8760(97)00098-6

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., & Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain.

Proceedings of the National Academy of Sciences of the USA, *104*(32), 13170–13175. doi:10.1073/pnas.0700668104

Mayer, A. R., & Bellgowan, P. S. F. (2014). Functional magnetic resonance imaging in mild traumatic brain injury. In S. M. Slobounov & W. J. Sebastianelli (Eds.), *Concussions in athletics*. New York: Springer. doi:10.1007/978-1-4939-0295-8

Mayhugh, R. E., Moussa, M. N., Simpson, S. L., Lyday, R. G., Burdette, J. H., Porrino, L. J., & Laurienti, P. J. (2016). Moderate-heavy alcohol consumption lifestyle in older adults is associated with altered central executive network community structure during cognitive task. *PLoS ONE*, *11*(8), e0160214. doi:10.1371/journal.pone.0160214

Meehan, T. P., & Bressler, S. L. (2012). Neurocognitive networks: Findings, models, and theory. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2232–2247. doi:10.1016/j.neubiorev.2012.08.002

Menon, R. S., & Kim, S.-g. (1999). Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends in Cognitive Sciences*, *3*(6), 207–216. doi:10.1016/S1364-6613(99)01329-7

Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, *15*(10), 483–506. doi:10.1016/j.tics.2011.08.003

Michel, C. M., & Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *NeuroImage*(November), 1–17. doi:10.1016/j.neuroimage.2017.11.062

Milz, P. (2016). Keypy: An open source library for EEG microstate analysis. *European Psychiatry*, *33*, S493. doi:10.1016/J.EURPSY.2016.01.1812

Milz, P., Faber, P., Lehmann, D., Koenig, T., Kochi, K., & Pascual-Marqui, R. (2015). The functional significance of EEG microstates: Associations with modalities of thinking. *NeuroImage*, *125*, 643–656. doi:10.1016/j.neuroimage.2015.08.023

Milz, P., Faber, P. L., Lehmann, D., Koenig, T., Kochi, K., & Pascual-Marqui, R. D. (2016). The functional significance of EEG microstates: Associations with modalities of thinking. *NeuroImage*, *125*, 643–656. doi:10.1016/j.neuroimage.2015.08.023

Milz, P., Pascual-Marqui, R., Achermann, P., Kochi, K., & Faber, P. (2017). The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. *NeuroImage*, *162*, 353–361. doi:10.1016/j.neuroimage.2017.08.058

Mizuno, T., Takahashi, T., Cho, R. Y., Kikuchi, M., Murata, T., Takahashi, K., & Wada, Y. (2010). Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. *Clinical Neurophysiology*, *121*(9), 1438–1446. doi:10.1016/j.clinph.2010.03.025

Nan, X., & Jinghua, X. (1988). The fractal dimension of EEG as a physical measure of conscious human brain activities. *Bulletin of Mathematical Biology*, *50*(5), 559–565. doi:10.1007/BF02458853

Natarajan, K., Acharya, U., R., Alias, F., Tiboleng, T., & Puthusserypady, S. K. (2004). Nonlinear analysis of EEG signals at different mental states. *Biomedical Engineering Online*, *3*(1), 7. doi:10.1186/1475-925X-3-7

Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., . . . Koenig, T. (2013). EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. *Clinical Neurophysiology*, *124*(6), 1106–1114. doi:10.1016/j.clinph.2013.01.005

Panda, R., Bharath, R. D., Upadhyay, N., Mangalore, S., Chennu, S., & Rao, S. L. (2016). Temporal dynamics of the default mode network characterize meditation-induced alterations in consciousness. *Frontiers in Human Neuroscience*, *10*, 372. doi:10.3389/fnhum.2016.00372

Parlar, M., Densmore, M., Hall, G. B., Frewen, P. A., Lanius, R. A., & McKinnon, M. C. (2017). Relation between patterns of intrinsic network connectivity, cognitive functioning, and symptom presentation in trauma-exposed patients with major depressive disorder. *Brain and Behavior*, *7*(5), e00664. doi:10.1002/brb3.664

Pascual-Marqui, R., Michel, C., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, *42*(7), 658–665. doi:10.1109/10.391164

Pascual-Marqui, R. D., Lehmann, D., Faber, P., Milz, P., Kochi, K., Yoshimura, M., . . . Kinoshita, T. (2014). *The resting microstate networks (RMN): Cortical distributions, dynamics, and frequency specific information flow*. arXiv:1411.1949, 1–14. doi:10.5167/uzh-100596

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology: Heart and Circulatory Physiology*, *278*(6), 2039–2049. doi:10.1152/ajpheart.2000.278.6.H2039

Rieger, K., Hernandez, L. D., Baenninger, A., & Koenig, T. (2016). Fifteen years of microstate research in schizophrenia: Where are we? A meta-analysis. *Frontiers in Psychiatry*, *7*(February), 1–7. doi:10.3389/fpsyt.2016.00022

Sadaghiani, S., Hesselmann, G., Friston, K. J., & Kleinschmidt, A. (2010). The relation of ongoing brain activity, evoked neural responses, and cognition. *Frontiers in Systems Neuroscience*, *4*(June), 20. doi:10.3389/fnsys.2010.00020

Seitzman, B. A., Abell, M., Bartley, S. C., Erickson, M. A., Bolbecker, A. R., & Hetrick, W. P. (2017). Cognitive manipulation of brain electric microstates. *NeuroImage*, *146*(2016), 533–543. doi:10.1016/j.neuroimage.2016.10.002

Shaw, S. (2017). Real-time filtering of gradient artifacts from simultaneous EEG-fMRI data. In *Proceedings of the 2017 International Workshop on Pattern Recognition in Neuroimaging*. doi:10.1109/PRNI.2017.7981510

Skarda, C. A., & Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, *10*(02), 161. doi:10.1017/S0140525X00047336

Sockeel, S., Schwartz, D., Pélégrini-Issac, M., & Benali, H. (2016). Large-scale functional networks identified from resting-state EEG using spatial ICA. *PLoS ONE*, *11*(1), e0146845. doi:10.1371/journal.pone.0146845

Stam, C. J. (2005). Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*, *116*(10), 2266–2301. doi:10.1016/j.clinph.2005.06.011

Sun, Y., Lim, J., Kwok, K., & Bezerianos, A. (2014). Functional cortical connectivity analysis of mental fatigue unmasks hemispheric asymmetry and changes in small-world networks. *Brain and Cognition*, *85*(1), 220–230. doi:10.1016/j.bandc.2013.12.011

Sun, Y., Yin, Q., Fang, R., Yan, X., Wang, Y., Bezerianos, A., . . . Sun, J. (2014). Disrupted functional brain connectivity and its association to structural connectivity in amnestic mild cognitive impairment and Alzheimer's disease. *PLoS ONE*, *9*(5). doi:10.1371/journal.pone.0096505

Suzuki, Y., Lu, M., Ben-Jacob, E., & Onuchic, J. N. (2016). Periodic, quasiperiodic and chaotic dynamics in simple gene elements with time delays. *Scientific Reports*, *6*, 1–13. doi:10.1038/srep21037

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411–423. doi:10.1111/1467-9868.00293

Uddin, L. Q., Kelly, A. M. C., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2009). Functional connectivity of default mode network components: Correlation, anticorrelation, and causality. *Human Brain Mapping*, *30*(2), 625–637. doi:10.1002/hbm.20531

Van Der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605. doi:10.1007/s10479-011-0841-3

Van de Ville, D., Britz, J., & Michel, C. M. (2010). EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proceedings of the National Academy of Sciences of the U.S.A.*, *107*(42), 18179–18184. doi:10.1073/pnas.1007841107

von Wegner, F., Tagliazucchi, E., & Laufs, H. (2017). Information-theoretical analysis of resting state EEG microstate sequences: Non-Markovianity, nonstationarity and periodicities. *NeuroImage*, *158*(June), 99–111. doi:10.1016/j.neuroimage.2017.06.062

Wackermann, J. (1999). Towards a quantitative characterisation of functional states of the brain: From the non-linear methodology to the global linear description. *International Journal of Psychophysiology*, *34*(1), 65–80. doi:10.1016/S0167-8760(99)00038-0

Wakin, M., Becker, S., Nakamura, E., Grant, M., Sovero, E., Ching, D., . . . Candes, E. (2012). A nonuniform sampler for wideband spectrally-sparse environments. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *2*(3), 516–529. doi:10.1109/JETCAS.2012.2214635

Wang, Y., Kang, J., Kemmer, P. B., & Guo, Y. (2016). An effcient and reliable statistical method for estimating functional connectivity in large scale brain networks using partial correlation. *Frontiers in Neuroscience*, *10*(March). doi:10.3389/fnins.2016.00123

Wolf, A. (1986). Quantifying chaos with Lyapunov exponents. In A. V. Holden (Ed.), *Chaos*. Princeton, NJ: Princeton University Press.

Yuan, H., Zotev, V., Phillips, R., Drevets, W. C., & Bodurka, J. (2012). Spatiotemporal dynamics of the brain at rest: Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. *NeuroImage*, *60*(4), 2062–2072. doi:10.1016/j.neuroimage.2012.02.031

Zhang, R., Xu, P., Chen, R., Li, F., Guo, L., Li, P., . . . Yao, D. (2015). Predicting intersession performance of SMR-based BrainComputer interface using the spectral entropy of resting-state EEG. *Brain Topography*, *28*(5), 680–690. doi:10.1007/s10548-015-0429-3

Received March 26, 2019; accepted July 9, 2019.

4.3 Discussion

The results of this paper identified major flaws in each of the three assumptions made during EEG microstate analysis, making it unsuitable for use as a biomarker for ICN activity. Most notably, this paper identified that assuming discontinuity of EEG microstates is incorrect, laying the foundation for a revised continuous-time interpretation of EEG microstates (Mishra et al., 2020). Follow-up work stemming from this study has also shown that the periodicity observed in EEG microstates can be a by-product of the rotating phase of alpha-oscillations (von Wegner et al., 2020).

Yet another finding from this paper was the inadequacy of the requirement for only one microstate being active at a given time. This was shown to be an artificially imposed requirement since there was considerable overlap between the EEG scalp topology at any given time and the pattern of more than one EEG microstate. This is also at odds with the ICN-linked interpretation of EEG microstates which posits that the distinct microstate topologies represent activation of distinct ICNs. However, co-activation of multiple ICNs has been well documented (Dosenbach et al., 2008), and can be a method for various brain regions to dynamically connect and communicate (Fries, 2015).

Furthermore, while EEG microstates have been used as EEG correlates of ICNs (Britz et al., 2010), there are inconsistencies in the correspondence of the clusters and the three core brain networks, especially the DMN.

Taken together, these shortcomings make it difficult to adequately infer network-level dynamics on the basis of EEG microstate dynamics, thereby eliminating EEG microstates as a viable EEG-based method to detect ICN activity and study their dynamics. Therefore, there is an apparent need for an EEG-based analysis method

that can detect ICN activity, in the context of the tri-network model.

Furthermore, in light of the identified shortcomings, the widespread clinical use of EEG microstates as a biomarker of the disease state, and even as a neurofeedback target (Diaz Hernandez et al., 2016), is particularly concerning. This further highlights the urgent need for a novel analysis method that can be clinically informative and practical.

Bibliography

Britz, J., Van De Ville, D., and Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, 52(4):1162–1170.

Corradini, P. L. and Persinger, M. a. (2014). Spectral power, source localization and microstates to quantify chronic deficits from 'mild' closed head injury: Correlation with classic neuropsychological tests. *Brain Injury*, 28(10):1317–1327.

Diaz Hernandez, L., Rieger, K., Baenninger, A., Brandeis, D., and Koenig, T. (2016). Towards Using Microstate-Neurofeedback for the Treatment of Psychotic Symptoms in Schizophrenia. A Feasibility Study in Healthy Participants. *Brain Topography*, 29(2):308–321.

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3):99–105.

Faber, P. L., Travis, F., Milz, P., and Parim, N. (2017). EEG microstates during different phases of Transcendental Meditation practice. *Cognitive Processing*, 18(3):307–314.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88(1):220–235.

Gschwind, M., Hardmeier, M., Van De Ville, D., Tomescu, M. I., Penner, I. K., Naegelin, Y., Fuhr, P., Michel, C. M., and Seeck, M. (2016). Fluctuations of

spontaneous EEG topographies predict disease state in relapsing-remitting multiple sclerosis. *NeuroImage: Clinical*, 12:466–477.

Gschwind, M., Michel, C. M., and Van De Ville, D. (2015). Long-range dependencies make the difference-Comment on "A stochastic model for EEG microstate sequence analysis". *NeuroImage*, 117:449–455.

Khanna, A., Pascual-Leone, A., Michel, C. M., and Farzan, F. (2015). Microstates in resting-state EEG: Current status and future directions. *Neuroscience and Biobehavioral Reviews*, 49:105–113.

Kikuchi, M., Koenig, T., Munesue, T., Hanaoka, A., Strik, W., Dierks, T., Koshino, Y., and Minabe, Y. (2011). EEG microstate analysis in Drug-Naive patients with panic disorder. *PLoS ONE*, 6(7):2–7.

Koenig, T., Prichep, L., Lehmann, D., Sosa, P. V., Braeker, E., Kleinlogel, H., Isenhart, R., and John, E. (2002). Millisecond by Millisecond, Year by Year: Normative EEG Microstates and Developmental Stages. *NeuroImage*, 16(1):41–48.

Milz, P., Faber, P. L., Lehmann, D., Koenig, T., Kochi, K., and Pascual-Marqui, R. D. (2016). The functional significance of EEG microstates-Associations with modalities of thinking. *NeuroImage*, 125:643–656.

Milz, P., Pascual-Marqui, R., Achermann, P., Kochi, K., and Faber, P. (2017). The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. *NeuroImage*.

Mishra, A., Englitz, B., and Cohen, M. X. (2020). EEG microstates as a continuous phenomenon. *NeuroImage*, 208:116454.

Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., Dierks, T., Strik, W., Kinoshita, T., and Koenig, T. (2013). EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. *Clinical Neurophysiology*, 124(6):1106–1114.

Rieger, K., Hernandez, L. D., Baenninger, A., and Koenig, T. (2016). 15 years of microstate research in schizophrenia - Where are we? A meta-analysis. *Frontiers in Psychiatry*, 7(FEB):1–7.

Santarnecchi, E., Khanna, A. R., Musaeus, C. S., Benwell, C. S. Y., Davila, P., Farzan, F., Matham, S., Pascual-Leone, A., and Shafi, M. M. (2017). EEG Microstate Correlates of Fluid Intelligence and Response to Cognitive Training. *Brain Topography*, 0(0):0.

Seitzman, B. A., Abell, M., Bartley, S. C., Erickson, M. A., Bolbecker, A. R., and Hetrick, W. P. (2017). Cognitive manipulation of brain electric microstates. *NeuroImage*, 146(April 2016):533–543.

Stevens, A., Lutzenberger, W., Bartels, D. M., Strik, W., and Lindner, K. (1997). Increased duration and altered topography of EEG microstates during cognitive tasks in chronic schizophrenia. *Psychiatry Research*, 66(1):45–57.

Strelets, V., Faber, P. L., Golikova, J., Novototsky-Vlasov, V., Koenig, T., Gianotti, L. R. R., Gruzelier, J. H., and Lehmann, D. (2003). Chronic schizophrenics with positive symptomatology have shortened EEG microstate durations. *Clinical Neurophysiology*, 114(11):2043–2051.

Strik, W. K., Chiaramonti, R., Muscas, G. C., Paganini, M., Mueller, T. J., Fallgatter,

A. J., Versari, A., and Zappoli, R. (1997). Decreased EEG microstate duration and anteriorisation of the brain electrical fields in mild and moderate dementia of the Alzheimer type. *Psychiatry Research - Neuroimaging*, 75(3):183–191.

von Wegner, F., Bauer, S., Rosenow, F., Triesch, J., and Laufs, H. (2020). EEG microstate periodicity explained by rotating phase patterns of resting-state alpha oscillations. *NeuroImage*, page 117372.

Chapter 5

Study of the Tri-network Model

5.1 Introduction

The tri-network model constitutes a powerful neuroscientifically-informed model of cognition and behaviour; it proposes three core networks, the CEN handling exteroceptively-driven cognitive functions, the DMN handling interoceptively-driven cognitive functions, and the SN gating the switching between these two networks (Menon, 2011). This model has improved our understanding of the network-level dysfunctions that underlie numerous psychopathologies (as discussed in Chapter 2), and the activity of the constituent ICNs have helped characterize sub-types of complex psychopathologies (Nicholson et al., 2020). Hence the detection of such ICN activity could provide clinically relevant biomarkers of pathological brain activity.

However, upon reviewing the current literature at the time of performing this study, it was apparent that a core aspect of this model remained to be tested, namely, the SN-based gating of DMN and CEN activation in the healthy brain, driven by interoceptive versus experoceptive task demands (as posited by the tri-network model).

The model was primarily derived from observations of dysfunctions in the three core ICNs in various psychopathologies (Menon, 2011), however the healthy network dynamics was not directly tested or experimentally characterized.

The study presented in this chapter addressed this gap in the literature by investigating the ICN dynamics in the context of a novel task-switching paradigm that randomly switched between interoceptive and exteroceptive processing task demands. A secondary goal of the study was to better understand the mechanisms underlying the gating function of the SN, since it is found to be disrupted in a wide range of mental health disorders.

The primary author (SBS) of the study conceptualized and designed the functional task used in this study with the last author (SB). The primary author (SBS) also performed all data collection, analysis and wrote the first draft of the manuscript. All co-authors contributed to the writing of the manuscript. MM, JH and SB contributed funding to this study.

5.2 Dynamic task-linked switching between brain networks - A tri-network perspective

Shaw, S. B., McKinnon, M. C., Heisz, J., & Becker, S. (2021). Dynamic task-linked switching between brain networks–A tri-network perspective. Brain and cognition, 151, 105725.

Brain and Cognition 151 (2021) 105725

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Dynamic task-linked switching between brain networks – A tri-network perspective

Saurabh Bhaskar Shaw [a,c,d], Margaret C. McKinnon [e,f,g,i], Jennifer Heisz [h,i], Suzanna Becker [a,b,c,d,i,*]

[a] Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada
[b] Department of Psychology Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada
[c] Vector Institute for Artificial Intelligence, Toronto, ON, Canada
[d] Centre for Advanced Research in Experimental and Applied Linguistics (ARiEAL), Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada
[e] Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada
[f] Mood Disorders Program, St. Joseph's Healthcare, Hamilton, ON, Canada
[g] Homewood Research Institute, Guelph, ON, Canada
[h] Department of Kinesiology, McMaster University, Hamilton, ON, Canada
[i] Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada

ARTICLE INFO

Keywords:
Tri-network model
Default mode network
Salience network
Central executive network
Connectome-multivariate pattern analyses (connectome-MVPA)
Multivariate Granger Causality (MVGC)
Graph theory

ABSTRACT

The highly influential tri-network model proposed by Menon integrates 3 key intrinsic brain networks — the central executive network (CEN), the salience network (SN), and the default mode network (DMN), into a single cohesive model underlying normal behaviour and cognition. A large body of evidence suggests that abnormal intra- and inter- network connectivity between these three networks underlies the various behavioural and cognitive dysfunctions observed in patients with neuropsychiatric conditions such as PTSD and depression. An important prediction of the tri-network model is that the DMN and CEN networks are anti-correlated under the control of the SN, such that if a task engages one of the two, the SN inhibits the activation of the other. To date most of the evidence surrounding the functions of these three core networks comes from either resting state analyses or in the context of a single task with respect to rest. Few studies have investigated multiple tasks simultaneously or characterized the dynamics of task switching. Hence, a careful investigation of the temporal dynamics of network activity during task switching is warranted. To accomplish this we collected fMRI data from 14 participants that dynamically switched between a 2-back working memory task and an autobiographical memory retrieval task, designed to activate the CEN, DMN and the SN. The fMRI data were used to 1. identify nodes and sub-networks within the three major networks involved in task-linked dynamic network switching, 2. characterize the temporal pattern of activation of these nodes and sub-networks, and finally 3. investigate the causal influence that these nodes and sub-networks exerted on each other. Using a combination of multivariate neuroimaging analyses, timecourse analyses and multivariate Granger causality measures to study the tri-network dynamics, the current study found that the SN co-activates with the task-relevant network, providing a mechanistic insight into SN-mediated network selection in the context of explicit tasks. Our findings also indicate active involvement of the posterior insula and some medial temporal nodes in task-linked functions of the SN and DMN, warranting their inclusion as network nodes in future studies of the tri-network model. These results add to the growing body of evidence showing the complex interplay of CEN, DMN and SN nodes and sub-networks required for adequate task-switching, characterizing a normative pattern of task-linked network dynamics within the context of Menon's tri-network model.

* Corresponding author at: Department of Psychology Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada.
E-mail address: becker@mcmaster.ca (S. Becker).

https://doi.org/10.1016/j.bandc.2021.105725
Received 4 October 2020; Received in revised form 15 March 2021; Accepted 26 March 2021
Available online 28 April 2021

1. Introduction

Traditionally the functional correlates of brain activity have been studied one region at a time. However, mounting evidence indicates that the brain is organized into functional networks of interacting brain regions. These clusters of sub-regions fire in a functionally correlated and synchronous manner, and are engaged for specific cognitive tasks, and are hence called functional networks (FN). Out of a large number of previously identified FNs (Yeo et al., 2011), three major brain networks are of particular interest to the cognitive neuroscience community. The central executive network (CEN), salience network (SN) and default mode network (DMN) form a tri-network system that is postulated to explain much of cognition and behaviour (Menon, 2011). The CEN is responsible for externally-directed cognitive tasks including working memory and other executive functions, while the DMN is involved in internally-directed cognitive tasks including autobiographical memory retrieval, imagining the future, spatial planning and navigation and self-reflection. Even while the person is not outwardly performing any explicit task (i.e. resting state), the brain is still active with internally-directed cognitive activities driven by the DMN, such as daydreaming, mind-wandering or more focused thought. According to Menon's tri-network model, the SN is involved in modulating the switching of attention between cognitive processes subserved by the CEN and DMN. This mediates the switch between externally-directed and internally-directed thoughts, respectively, a process required for successful cognition and balanced emotion regulation. Importantly, the inter- and intra-network connectivity between these three core networks is found to be dysregulated in psychiatric conditions such as depression (Hamilton et al., 2011; Yang et al., 2016), schizophrenia (Moran et al., 2013) and PTSD (Menon, 2011). Hence, studying the tri-network activity and its switching dynamics is essential to understanding the neurological underpinnings of the aberrant thought processes and emotional dysregulation observed in such complex neuropsychological disorders. This study investigates the temporal dynamics of various sub-components of the tri-network model with a composite memory task designed to activate the three networks.

1.1. The tri-network model

An important prediction of the tri-network model is that the DMN and CEN are anti-correlated under the control of the SN, such that if a task engages one of the two, the SN inhibits the activation of the other (Uddin, 2015).

There is a large body of work supporting SN's gating function, using dynamic causal modelling and Granger causality analysis to study major network nodes (Friston et al., 2014) during resting state (Sridharan et al., 2008, 2014, 2017), visual or auditory attention-switching (Sridharan et al., 2008), Go-No-Go and congruent/incongruent flanker tasks (Cai et al., 2016) and magnitude and parity judgement tasks (Sidlauskaite et al., 2014). These causal analyses support the gating function of the SN within tasks (Chand and Dhamala, 2016) and even more so when switching between externally-oriented tasks (Sidlauskaite et al., 2014).

More specifically, these studies identify the right anterior insula (rAI) as the node with the lowest latency of the event-related activity and the highest net causal "outflow" to other SN, CEN and DMN nodes, implicating a central role for rAI as a hub responsible for switching between different tasks, requiring a switch in the predominantly active network (Sridharan et al., 2008). Neurofeedback training is also known to improve AI-linked network regulation of the CEN and DMN, supporting the hub-like role of this region (Zhang et al., 2015). Furthermore, stimulating other cortical nodes of the salience and central executive networks using transcranial magnetic stimulation (TMS) causes deactivation of DMN nodes, suggesting that it is under the inhibitory control of CEN/SN nodes (Chen et al., 2013).

However, most of the evidence surrounding the functions of these three core networks comes from either resting state analyses (Goulden et al., 2014; Taghia et al., 2017; Ryali et al., 2016) or paradigms where brain network activation is studied in the context of a single task relative to rest (Sidlauskaite et al., 2014). Few studies have investigated multiple tasks simultaneously (see Spreng et al. (2010)), and even fewer have attempted to characterize the dynamics of task switching (see Bréchet et al. (2019, 2014)). Bréchet et al. (2019) characterized the spatial regions activated across two different tasks, while describing the temporal dynamics using EEG microstates, a method that makes some faulty assumptions (Shaw et al., 2019). Furthermore, most findings supporting the gating function of SN has either been in the context of resting state (Chand et al., 2017), or in the context of suppressing DMN and recruiting CEN during externally-directed tasks (Sridharan et al., 2008; Chen et al., 2013; Chand and Dhamala, 2016), and has not assessed the network switching dynamics when recruiting DMN for self-related tasks in healthy participants. Additionally, these studies do not look at the temporal sequence of activity while switching between tasks that rely on different networks, that can provide a dynamic view of the switching process. Hence, a detailed investigation of the temporal dynamics of network activity during task switching is warranted.

This study attempts to address these shortcomings by studying three major aspects of CEN, SN and DMN functioning in the context of a paradigm that involves dynamic task switching designed to activate these three networks. This involves

1. identification of the nodes and sub-networks within the three major networks that are involved in such dynamic network switching,
2. characterization of the temporal pattern of activation of these nodes and sub-networks, and finally
3. investigation of the causal influence that these nodes and sub-networks exert on each other over the course of the tasks.

To fulfill these aims, the spatiotemporal pattern of activating each of the networks while performing the tasks is first described using data-driven patterns of voxel-to-voxel connectivity and event-related activity timecourses, followed by multivariate Granger causality analysis to probe the patterns of causality between the network nodes.

2. Methods

In this study, fMRI data were collected from 14 healthy participants (8 females, mean age 22 ± 3 years) while they performed a complex memory task expected to evoke dynamic switching between intrinsic networks. The complex memory task consisted of randomly interleaved blocks of autobiographical memory (ABM) retrieval trials and working memory (WM) trials, expected to activate the DMN and CEN respectively. At the end of some blocks, a task-switching cue prepared participants to switch from an AM to a WM block or vice versa, and was expected to activate the SN. All participants provided informed consent prior to their participation in the study, in accordance with the approved Hamilton Integrated Research Ethics Board (HiREB) protocol.

2.1. Memory assessment

Each participant engaged in a pre-scanning phase, during which detailed autobiographical memories were recorded by the participant. Following this, participants completed a 1 h 20 min long memory assessment in the MRI scanner, comprised of randomly ordered 30-s blocks of either cued AM retrieval (remembering previous autobiographical memories - ABM) or a 2-back WM task (remembering which word they saw 2 words ago in a stream of words). These two block types were predicted to activate the DMN and CEN respectively, while a cue to a pending task switch between blocks was expected to activate the SN. Participants completed as many blocks as possible within 80 min, up to a maximum of 64 blocks (32 ABM, 32 WM), in chunks of 16 blocks. Each block was followed by a 60 s rest period to prevent fatigue.

Autobiographical memory (ABM) task Our ABM task was modeled

after a similar protocol used by Addis et al. (2004), with some modifications to balance the visual stimulus processing and motor response demands between the ABM and WM tasks. Prior to entering the scanner, the participant was asked to recall and record within an Excel sheet up to 10 positive or neutral autobiographical memories in vivid detail. The participant was then asked to identify descriptive words corresponding to each memory, that would serve as cues during the ABM retrieval task. During each ABM block within the scanner, the participant was shown a series of cue words pertaining to one of the previously described autobiographical memories and was instructed to vividly imagine the corresponding memory in detail. Each ABM block included 10 cues pertaining to the same memory, shown for 2 s each. In each different ABM block, a different memory was cued. To ensure that the participant was staying on task, at random points within each block, they were asked to perform a 4-alternative forced choice trial in which they should select the word that represents the memory they were currently recalling from a selection of four words.

Working Memory (WM) task: During the WM blocks, the participant was presented with a sequence of neutral words (not expected to elicit the autobiographical memories recorded) and was asked to remember the word that was presented two words back. At random points within each block, they were asked to perform a 4-alternative forced choice trial in which they should choose the word they saw two words ago from a selection of four words.

Task cue period: The words "Word Memory" and "Autobiographical Memory" were shown for 2 s before the onset of the WM and ABM blocks, respectively. A schematic illustrating the task paradigm is shown in Fig. 1.

2.2. MRI acquisition

A GE MR750 3T MRI scanner and an 8 channel RF coil (General Electric Healthcare, Milwaukee, WI) were used for data acquisition. FMRI data were acquired using an axial 3D fSPGR pulse sequence (2D gradient echo EPI, FA = 90°, TE/TR = 35/2000 ms, 64x64 matrix, 39 interleaved 3.8 mm slices, 1200 temporal points).

2.3. MRI data analysis

The raw fMRI data were first bandpass filtered (0.008 Hz - 0.09 Hz), motion corrected, and aligned with the MNI standard space after applying a 2 mm Gaussian blur. These pre-processed fMRI data were used for all subsequent analyses.

To accomplish the three major study aims described at the end of Section 1, we conducted a series of analyses probing the spatio-temporal patterns of activity in the three networks of interest. These analyses are discussed below, in the same order as the corresponding study objectives they are accomplishing.

1. We first identified brain regions differentially activated by the ABM and WM tasks, and during task switching, using whole-brain connectome multivariate pattern analysis (connectome-MVPA), followed by a group-wide independent component analysis (group-ICA). Both of these analyses are data-driven and hence avoided any a priori biases or assumptions around the network nodes and sub-networks that might participate in the tasks being studied. The group-ICA components corresponding to independent CEN, DMN and SN sub-networks were then identified for further analysis.
2. Next, temporal patterns of task-linked network activation were studied using two complementary approaches. First, an ICA-based data-driven approach was used to identify and study the temporal activation patterns of the three networks. This was followed by a region of interest (ROI)-driven analysis, based on evidence from the

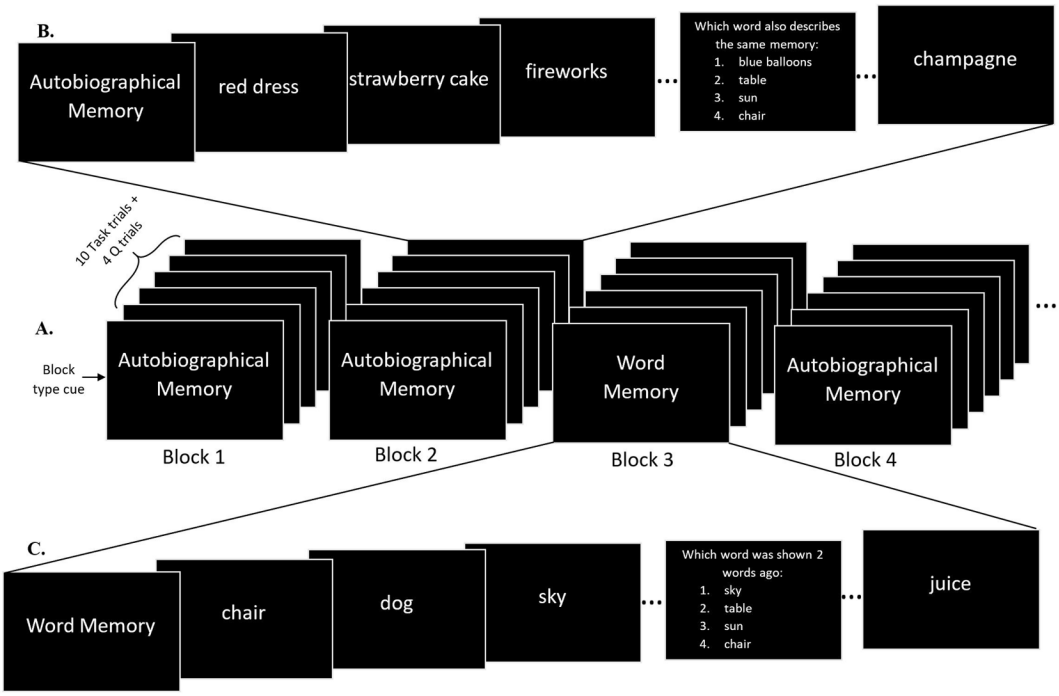

Fig. 1. A. Schematic of the dynamic switching task. Each block is either an autobiographical memory retrieval task (ABM task) or a 2-back working memory (WM) task. Each block consists of exactly 10 task cue trials and 4 four alternative forced choice (4-AFC) trials placed at random locations within the block. Each trial is presented to the participant for 2 s, except the 4-AFC trials that persist until the participant has made a choice. The structure of an ABM block and a WM block are shown in more detail in panels B (top) and C (bottom) respectively. During the ABM blocks (panel B), the task cues are memory keywords that correspond to one of the participant's autobiographical memories, thereby cuing that particular memory. The trials within a block consists of different memory cues from a single autobiographical memory. On the other hand, the task cues during the WM blocks (panel C) are a sequence of words from which the participant needs to remember the word shown 2 words ago.

literature to identify temporal patterns of activity within key nodes of the CEN, DMN and SN.
3. Finally, these activity time courses were analyzed using multivariate Granger causality (MVGC) analysis to identify the causal influence of these nodes and sub-networks on one another.

These analyses were carried out using the SPM12 and CONN toolboxes (Whitfield-Gabrieli and Nieto-Castanon, 2012a), along with custom MATLAB scripts and are described in detail in the following subsections.

2.3.1. Connectome Multivariate Pattern Analysis (connectome-MVPA)

In order to identify brain regions activated by the ABM and WM tasks without any a priori biases (study aim 1), we used data-driven whole-brain connectome multivariate pattern analysis (connectome-MVPA). This analysis identifies whole-brain multivariate patterns of functional connectivity (*MCOR*), allowing for discovery of task-linked patterns of voxel-to-voxel functional connectivity (Beaty et al., 2015; Thompson et al., 2015; Whitfield-Gabrieli et al., 2016).

Mathematically, the *MCOR* connectivity maps refer to the top M spatial principal component scores of the seed-based correlation matrix containing the correlation between each seed voxel and all other voxels in the brain, aggregated across all participants and conditions.

The whole-brain connectome-MVPA maps are identified by repeating the following procedure for each of the K voxels, x, in the brain. First, for each voxel x, the $1 \times (K)$-dimensional seed-based correlation map is computed using x as the seed, for each participant-condition pair. Assuming a total of N participant-condition pairs, these seed-based correlation maps are aggregated into a $N \times K$ seed-based correlation matrix ($\mathbf{R}$) for each voxel x.

Next, the top M principal spatial components of $\mathbf{R}$ are identified using principal component analysis (PCA), as shown in Eq. (1).

$$\mathbf{R} = \mathbf{Q} \cdot \Lambda \cdot \mathbf{Q}^T \tag{1}$$

where Λ is the $M \times M$ square diagonal matrix of eigenvalues corresponding to the principal spatial components (eigenvectors) contained in the $K \times M$ dimensional $\mathbf{Q}$ matrix. These components form an orthogonal basis set that span the space of the seed-based correlation patterns observed across all participants and conditions.

The $N \times M$ dimensional component scores matrix (*MCOR*) is computed using the $\mathbf{R}$ matrix and $\mathbf{Q}$ as shown in Eq. (2), giving a $1 \times M$ dimensional component score for each of the N participant-condition pairs, for each voxel x:

$$\mathbf{MCOR}(x) = \mathbf{R} \cdot \mathbf{Q} \tag{2}$$

Consequently, each element $r(x,y)$, of the n^{th} participant-condition pair seed-based correlation matrix ($\mathbf{R}$) can be written as,

$$\mathbf{R}_n : r_n(x,y) = \sum_{i=1}^{M} MCOR_{n,i}(x) Q_i(x,y) \tag{3}$$

The *MCOR* maps provide a multivariate low-dimensional representation of the between-subjects/conditions variance of the seed-based functional connectivity between each voxel x and the rest of the brain. Task-linked differences in these whole-brain patterns of connectivity can then be identified by performing a second-level between condition/participant analysis by performing an F-contrast on the *MCOR* spatial maps between the task conditions of interest.

First-level connectome-MVPA analyses were run using the *CONN* toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012b) with $M = 4$ top components, followed by a second-level contrast between the ABM and WM task conditions to identify the multivariate voxel-to-voxel patterns of connectivity that differ between these task conditions.

2.3.2. Independent Component Analysis (group-ICA)

We studied the temporal dynamics of the 3 networks (study aim 2) using two different approaches, a data-driven approach to identifying the three networks and a more traditional ROI-driven approach. For the data-driven approach, following connectome-MVPA, we used group wide independent component analysis (group-ICA) for the identification of spatio-temporally independent CEN, DMN and SN sub-networks without any a priori biases (study aim 1).

The group-ICA decomposition (Calhoun et al., 2009) was performed on the pre-processed voxel data using *tanh* as the non-linear contrast function with the iterative FastICA algorithm to identify 20 group-ICA spatial components. The group-level ICA components were labeled based on their spatial overlap (quantified using the Dice coefficient) with known functional networks (described in Shirer et al. (2012) and downloaded from https://findlab.stanford.edu/functional_ROIs.html) and back-projected to individual participants' data, using GICA3 back-projection (Erhardt et al., 2011), to obtain the activity timeseries of each ICA component. These timeseries were used for the temporal analysis of CEN, DMN and SN sub-network activity (study aim 2).

2.3.3. Regions-of-Interest (ROI) Definitions

The temporal dynamics of CEN, DMN and SN node activity (study aim 2) were also studied using an ROI approach, based on the average BOLD activity within major nodes of the three networks. Two network configurations were studied, each using a different set of nodes to define the three networks, to identify the appropriate collection of nodes that adequately described tri-network dynamics in the context of a task paradigm. The first network configuration included an abridged set of 10 nodes that has been extensively used to define the three networks in most prior studies investigating the tri-network model. Additionally, since this widely used set of tri-network nodes did not include key DMN and SN nodes, such as medial-temporal lobe (MTL) nodes and the posterior insula (PI), a second network configuration was defined by adding 8 nodes that span the MTL and PI, to the earlier set of nodes comprising the DMN and the SN respectively.

The nodes for the abridged and extended network configurations were defined using a combination of standard and custom regions of interest (ROI), as described in Tables S1 and S2, respectively. The standard ROIs from the Harvard-Oxford atlas (H-O atlas) were used to define the posterior cingulate cortex (PCC), precuneus (PreC), right and left posterior parietal cortex (r/l PPC), hippocampal (r/l HC) and parahippocampal (r/l pHC) ROIs. The rest of the ROIs were created using a sphere of radius 5 mm, centered at pre-defined MNI coordinates, given in Tables S1 and S2.

2.3.4. Multivariate Granger Causality (MVGC) Analyses

In order to assess the causal interactions between the networks (study aim 3), the causal influence between the identified sub-networks and network nodes were studied using multivariate causality analyses on the ICA component timeseries and the ROI timeseries, respectively.

Granger causality has been extensively used to characterize data-driven estimates of directional causality between two sources of activity in EEG and fMRI (Iwabuchi et al., 2017; Seth et al., 2015). It relies on information-theoretic principles of causality, asserting that a signal x_i causally influences signal x_j if predictions about the future samples of x_j can be improved by including past samples of x_i in addition to past samples of x_j itself. The multivariate version of Granger causality (Barnett and Seth, 2014) extends this univariate description by also considering activity from nodes in the system that may indirectly influence the causality from $i \rightarrow j$. Within this framework, two vector autoregressive (VAR) models are fitted - a 'full' or 'unrestricted' VAR model that models the activity of all system nodes ($\mathbf{x}(t)$) as a function of their past activity; and a 'reduced' or 'restricted' VAR model that estimates future activity of the system excluding node i ($\mathbf{x}(t)|x_i$), as shown in Eqs. 4 and 5 respectively.

$$\mathbf{x}(t) - \widetilde{\mathbf{A}}L^{P}(\mathbf{x}(t)) = \widetilde{\epsilon} \qquad (\text{full}/\text{unrestrictedmodel}) \tag{4}$$

$$(\mathbf{x}(t)|x_i) - \mathbf{A}L^{P}(\mathbf{x}(t)|x_i) = \epsilon \qquad (\text{reduced}/\text{restrictedmodel}) \tag{5}$$

where $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ are $(n-1) \times P$ and $n \times P$ dimensional coefficient matrices respectively, for a system with n nodes. L^P is the lag operator such that $L^P(\mathbf{x}(t))$ gives a vector with the last P values in $\mathbf{x}(t)$, and ϵ and $\widetilde{\epsilon}$ are white time-uncorrelated noise processes with covariance $\mathbf{D}$ and $\widetilde{\mathbf{D}}$ respectively (Duggento et al., 2018).

The MVGC strength for the connection from $i \rightarrow j$ is given by estimating the additional information gained about future x_j samples by including previous x_i samples, ignoring the information provided by the other nodes in the system (Eq. (6)):

$$MVGC_{i,j} = \log \frac{\mathbf{D}_{j,j}}{\widetilde{\mathbf{D}}_{j,j}} \tag{6}$$

The optimal value for P of the lag operator L^P can be estimated from the time series data using the Akaike information criterion (AIC) or Bayesian information criterion (BIC), to maximize the fit of the VAR model while minimizing overfitting (Barnett and Seth, 2014). Optimal values of $P = 3$ have been previously observed when applying MVGC analysis on BOLD fMRI data (Duggento et al., 2018).

MVGC analyses were performed on the ICA and ROI timecourses using the MVGC Toolbox (Barnett and Seth, 2014). The AIC was used to estimate the optimal lag for fitting the VAR model, followed by splitting the time courses into 14 s long sliding windows with a step size of 2 s (1 TR). This provided a dynamic estimate of the MVGC values over time, that was compared against the null MVGC estimate using time-permuted surrogate signals. The connections that were significantly different from their surrogate estimates were retained. A unique directed graph was constructed from the MVGC estimates for each sliding window, generating a series of graphs.

To characterize the changing role of each network node in this series of graphs, the net causal outflow was estimated for each network node by subtracting the weighted sum of incoming connections to the node, (i.e. importance weighted in-degree of the node) from the weighted sum of outgoing connections from the node (i.e. importance weighted out-degree of the node). A hubness score was also estimated for each node of these graphs by identifying the nodes that have high node degree, high node betweenness, low clustering coefficients and low average path length. This combination of four sub-measures is thought to indicate hub-like behaviour within a network, and has been used to define a composite hubness score (van den Heuvel et al., 2010). This composite hubness score was derived from sub-scores of each of the four sub-measures. A sub-score of 1 was assigned to each node if it ranked in the top 20% of nodes that matched the aforementioned hub-like pattern of the four sub-measures, while a sub-score of 0 was assigned to all nodes that failed to meet this criteria. The sub-scores for each of the four sub-measures were then added together to identify the composite hubness score for each node. Consequently, this score ranged from 0, representing no hub-like behaviour, to 4, indicating the presence of all four hub-like characteristics. Further details on this composite hubness score can be found in van den Heuvel et al. (2010).

3. Results

This section first discusses the results of the connectome-MVPA and group-ICA analyses in sub-section 3.1, accomplishing the goals of study aim 1 to spatially characterize the tri-network activity by identifying sub-networks and dominant regions of interest. Next, the results pertaining to study aim 2 are discussed by exploring the temporal characteristics of the ICA and ROI time courses in sub-section 3.2. Finally, the results of the MVGC analyses performed to investigate study aim 3 are presented in sub-section 3.3, characterizing the patterns of causality between the identified nodes and sub-networks using multivariate Granger causality analysis.

3.1. Spatial characterization of tri-network activity

3.1.1. Connectome-MVPA Analysis

The connectome-MVPA analysis for the contrast between the ABM trials and WM trials revealed that when contrasted against the WM trials, the ABM trials more strongly activated the right and left parietal cortices and the posterior cingulate cortex (PCC), while the WM trials more strongly activated supplementary motor areas, right and left middle frontal gyri, and some medial frontal cortical and right cerebellar regions. These results, shown in Fig. 2 and Table 1, are consistent with previous findings of activation within the medio-temporal subnetwork of the DMN in autobiographical memory recall (Andrews-Hanna et al., 2014), and activation of phonological loop structures in working memory (Yaple et al., 2019).

3.1.2. ICA Networks - Spatial Pattern Analysis

Of the 20 identified ICA components, 11 were found to have significant overlap with various functional networks (Shirer et al., 2012), with 5 of these 11 external to the networks of interest (spanning language, visual, auditory, cerebellar and somato-motor networks), while the remaining 9 components appeared to be artefactual or noise components. The remaining 6 ICA components were identified as sub-networks of CEN, DMN and SN and were isolated for further analyses. These include the.

- Left CEN comprised of left dlPFC and parietal structures
- Bilateral CEN comprised of bilateral dlPFC and parietal structures,
- Dorsal DMN comprised of posterior cingulate cortex (PCC) and MPFC nodes,
- Ventral DMN comprised of retrosplenial cortex and medial temporal lobe structures,
- Anterior SN comprised of anterior insula and dorsal ACC nodes, and
- Posterior SN comprised of posterior insula.

The spatial patterns of these components are shown in Fig. 3.

3.2. Temporal characterization of tri-network activity

3.2.1. ICA Networks - Timecourse Analysis

The ICA timecourses for the CEN, DMN and SN networks are shown in Figs. 4a and 4b, averaged over all ABM and WM trials respectively.

During the ABM trials (Fig. 4a), as predicted, the SN was observed to coactivate with the DMN, jointly increasing after t = 0, and peaking at 12s post-task onset. At the same time, the CEN activity decreased through the first half of the ABM trials, with its lowest point coinciding with the DMN/SN activity peak. Surprisingly, after the point of peak DMN and minimum CEN activity at 12 s, DMN activity decreased while CEN activity increased, reaching its peak at 18s, suggesting there may be greater demands on executive functions during the latter half of the trial. This peak in CEN activity, however, was not accompanied by a peak in SN activity.

During WM trials, unexpectedly the SN did not seem to coactivate with the CEN. Instead, the activity of all three networks increased after t = 0s, reaching their maximum levels at different times. The SN was the first to reach its peak at 5s, followed by the DMN at 8s, ending with the CEN peak at 13s.

The observed pattern of SN network activity was unexpected in both ABM and WM trials. We predicted that the SN would control the selection between the CEN and DMN according to task demands and therefore expected SN activity to rise at the start of each block, followed by activation of either the CEN or DMN in a task dependent manner. Additionally, we expected the DMN to be activated during ABM trials but not WM trials and the CEN to be activated during WM trials but not

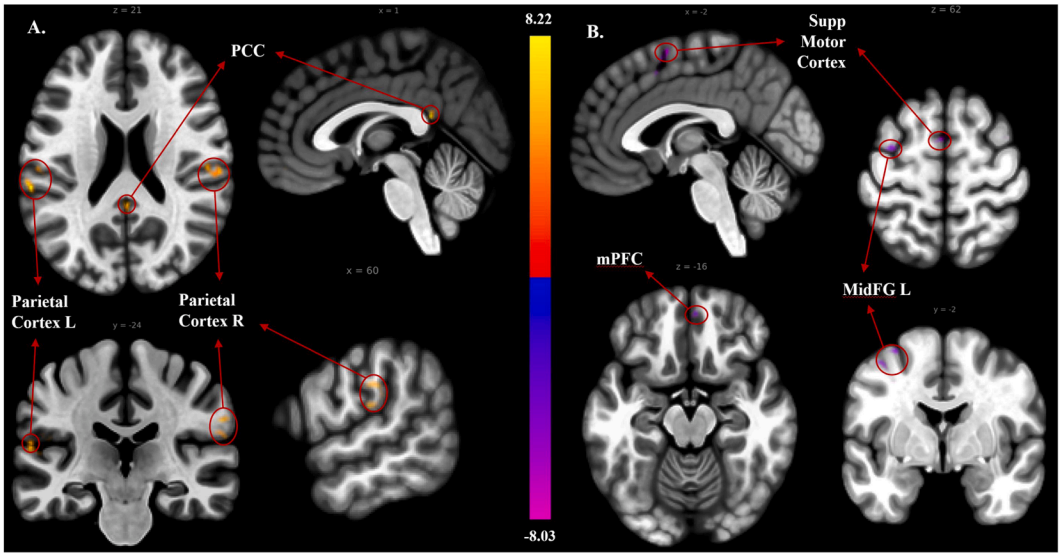

Fig. 2. Connectome-MVPA activation for the contrast between ABM>WM with positive contrast clusters (ABM > WM) shown in red/yellow (panel A on the left) and negative contrast clusters (WM > ABM) shown in blue/purple (panel B on the right). The identified clusters of activity, along with their MNI coordinates are listed in Table 1. Abbreviations: PCC - posterior cingulate cortex, mPFC - medial prefrontal cortex, MidFG - Middle Frontal Gyrus, Supp Motor Cortex - Supplementary Motor Cortex. A cluster-discovery threshold of p-uncorrected < 0.001 and cluster-visualization threshold of FDR-corrected p < 0.05 was used to generate this image.

Table 1
Clusters of voxels identified in a two-sided ABM > WM contrast of connectome-MVPA values. The abbreviations used are aSMG = Anterior Supramarginal Gyrus, PostCG = Post Central Gyrus, PO = Parietal Opercular Cortex, PT = Planum Temporale, CO = Central Opercular Cortex, MidFG = Middle Frontal Gyrus, SFG = Superior Frontal Gyrus, SMA = Supplementary Motor Cortex, LG = Lingual Gyrus.

MNI coordinates ofactivation clustercentroids [x y z]	Brain regions activated	Number of Voxels
ABM > WM		
−60 −32 + 24	Parietal Cortex L	182
	aSMG L	68
	PO L	53
	PT L	39
	PostCG L	22
+62 −18 + 32	Parietal Cortex R	155
	aSMG R	80
	PostCG R	32
	PO R	30
	CO R	13
+00 −44 + 20	Posterior Cingulate	55
WM > ABM		
−26 + 02 + 58	Middle Frontal Gyrus L	73
	MidFG L	36
	PreCG L	15
	SFG L	11
−02 + 04 + 64	Supplementary Motor Cortex	53
	SMA L	22
	SMA R	19
	ParaCingG L	12
+08 −92 −02	Occipital Pole R	46
+10 −80 −16	Cerebellum	34
	LG R	18
	Cereb 6 R	11
	Vermis 6	4
+36 + 04 + 66	Middle Frontal Gyrus R	16
+00 + 50 −16	Frontal Medial Cortex	17

ABM trials, however, as noted above, there was cross-network activation in both types of trials. We postulated that this pattern of tri-network activity could be due to differential activity within its constituent sub-networks and we therefore investigated the anterior vs posterior sub-networks of the SN, alongside the CEN and DMN sub-networks identified by the group-wide ICA analysis, described in Section 3.1.2. The activation timecourses of these SN, CEN and DMN subnetworks are shown in Fig. S1.

The correlated increase in DMN and SN activity during ABM trials observed in Fig. 4a was found to be due to the correlated increase in activity of the ventral DMN and anterior SN sub-networks (seen in panels b and c respectively), whereas the later increase in CEN activity in Fig. 4a was due to an increase in activity of the bilateral CEN network (panel a). In contrast, during the WM trials, the posterior SN (panel f) and the dorsal DMN activity (panel e) seemed to underlie the early SN and DMN peaks, while the LCEN subnetwork (panel d) contributed to the later CEN peak at 13s observed in Fig. 4b. An early peak in bilateral CEN activity was also observed, coinciding with the increase in dorsal DMN and posterior SN activity in panels d, e and f of Fig. S1, respectively.

In sum, two key patterns emerged in these sub-network analyses. Firstly, the anterior SN subnetwork showed co-activation with the task-relevant network (ventral DMN during ABM trials and LCEN during WM trials). In contrast, the posterior SN subnetwork was anti-correlated with the anterior SN, and instead consistently co-activated with the bilateral CEN and dorsal DMN subnetworks during both ABM and WM trials. Secondly, the LCEN co-activated with the bilateral CEN during the ABM trials, however, the LCEN showed a distinctly different activation pattern during WM trials, increasing in activity during the latter half of the trials (Fig. S1d).

The network dynamics of these sub-networks were further studied using multivariate Granger causality (MVGC) analyses and are discussed later in Section 3.3.1.

While the above analyses are based on the data-driven (ICA) identification of the networks and sub-networks, we also examined the tri-network activation in terms of ROI analyses in the next section.

3.2.2. ROI-to-ROI Timecourse Analysis

In addition to ICA based tri-network definitions, we investigated the tri-network activation using two different sets of ROIs to describe the CEN, DMN and SN. The first network configuration utilized an abridged set of ROIs widely used in the literature to study the tri-network model, while the second network configuration extended the set of ROIs defining the DMN and SN to include medial temporal lobe (MTL) and posterior insula (PI) nodes. The results from these ROI analyses are

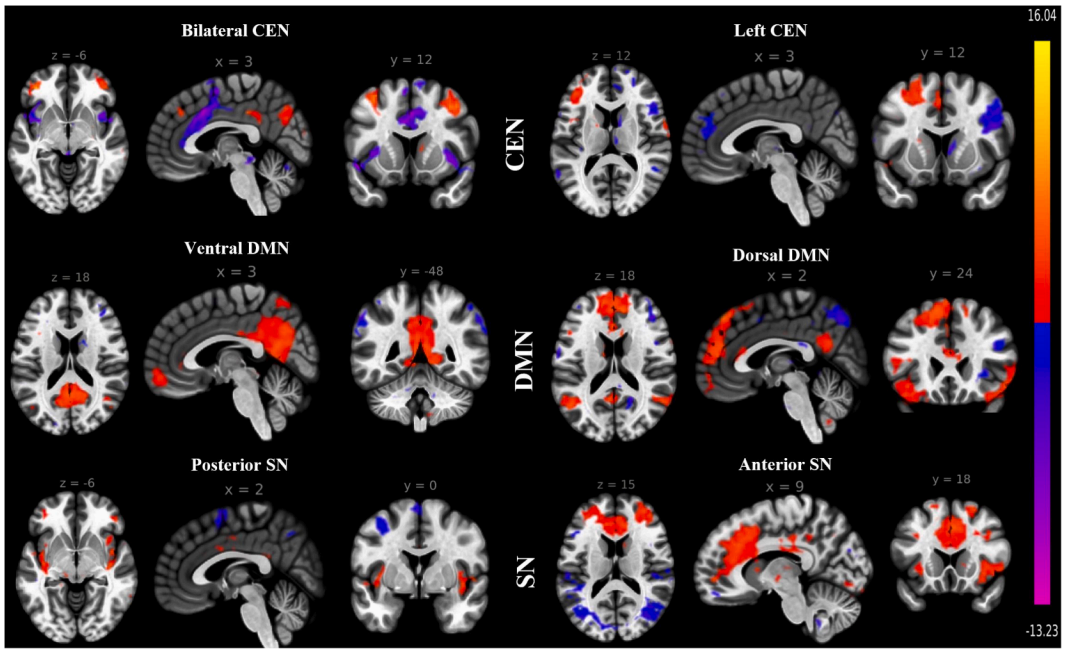

Fig. 3. Spatial maps of ICA components with significant overlap with CEN, DMN and SN. Two ICA components were identified to have significant overlap with each one of the three networks.

discussed in this section.

The average BOLD activity timecourses of each of the three major networks, gathered by summing together the activity of the corresponding nodes, is shown in Figs. 4c and 4d, averaged across all ABM and WM task trials respectively.

During the ABM trials, consistent with our findings from the ICA-based sub-network analyses, SN activity was the first to increase, reaching its peak at 8 s, followed by an increase in DMN activity which peaked at 11 s. As expected, CEN activity decreased by the same amount during the same time frame, and increased only after the DMN activity had returned to baseline around 18 s post-task onset. This change in CEN activity was accompanied by a decrease in SN activity at 18 s post-task onset. During the WM trials, the DMN activity increased rapidly, peaking around 6 s, while the SN and CEN activity decreased during the same time frame. This was followed by an increase in SN and CEN activity, peaking at 17 s and 19 s respectively. This was accompanied by a large decrease in DMN activity, reaching its lowest activity level at 20 s post-task onset. Across both ABM and WM trials, the SN was found to co-activate with the activity of the network predominantly expected to activate, i.e. DMN during the ABM trials and CEN during the WM trials. Furthermore, SN activity was observed to lead the activity of the activated network, consistent with the activation pattern expected for a network gating the onset of another network.

The activity of each network was further broken down into the activity of its constituent nodes for more detailed analysis, as shown in Fig. S2. During the ABM trials, there was an increase in vmPFC activity at 5 s post-task onset (panel b), followed by dACC, right AI and left AI activity around 8 s post-task onset (panel c), after which PCC activity increased to its maximum value around 10 s post-task onset (panel b). The amPFC activity followed an opposite trend to that of PCC (panel b). During the latter half of the trial, right and left dlPFC activity were found to increase around 18 s post-stimulus onset (panel a). During the WM trials, the activity of SN and CEN nodes was found to be biphasic. First, the dACC activity increased at 14 s post-stimulus onset (panel f), immediately followed by an increase in right PPC activity (panel d). In the second phase, the right and left AI activity increased to their maximum values at 16 s and 20 s post-task onset respectively (panel f). This was followed by an increase in left PPC, left dlPFC and finally right dlPFC, peaking at 18, 20 and 22 s respectively (panel d). Activity of the DMN nodes during the WM trials was dominated by an early increase in PCC activity around 7 s (panel e), which was accompanied by a decrease in right and left AI node activity (panel f).

Taken together, the results described above and illustrated in Fig. 4 suggest that different nodes dominated the activity of each network during the two different tasks. The observed increase in DMN activity during the ABM task might have been due to an increase in vmPFC and PCC activity driven by rAI, lAI and dACC activity, whereas, the early increase in DMN activity during the WM trials was primarily due to a corresponding increase in PCC and amPFC activity. This increase in PCC activity during WM trials was accompanied by a decrease in SN node activity indicating that this activation might be negatively correlated to the SN. Furthermore, the similar patterns of PCC activity observed during both ABM and WM trials might indicate some common component processes required to adequately complete both tasks. The late increase in CEN activity during the ABM task was dominated by an increase in right and left dlPFC activity, and was also accompanied by a decrease in SN node activity. However, the increase in net CEN activity during the WM task was found to be due to an underlying increase in right PPC activity and that of the dACC node of the SN, followed by an increase in left PPC and bilateral dlPFC activity, shortly preceded by an increase in bilateral AI activity. Consequently, this CEN activation might have been driven by the SN. To better investigate the causal patterns underlying the observed network dynamics, multivariate Granger causality (MVGC) analysis was performed using the described network nodes, and its results are discussed in Section 3.3.2.

While the above described results rely on a small subset of nodes used to describe the CEN, DMN and SN in prior studies investigating the tri-network model (abridged ROIs), they omit a key set of medial temporal lobe (MTL) nodes implicated in DMN-linked memory tasks, in addition to the posterior insula (PI), which is thought to participate in the bottom-up salience detection functions of the SN. To investigate the impact of including these key nodes within the network definitions on tri-network activity, we added right and left hippocampus (r/l HC), right and left parahippocampus (r/l pHC) and the retrosplenial cortex (RSC) to the set of DMN nodes. Additionally, we also included the right and left posterior insula (r/l PI) in the set of SN nodes. The impact of including the extended set of nodes on tri-network activity is discussed below.

During the ABM trials, the addition of the MTL nodes in the DMN

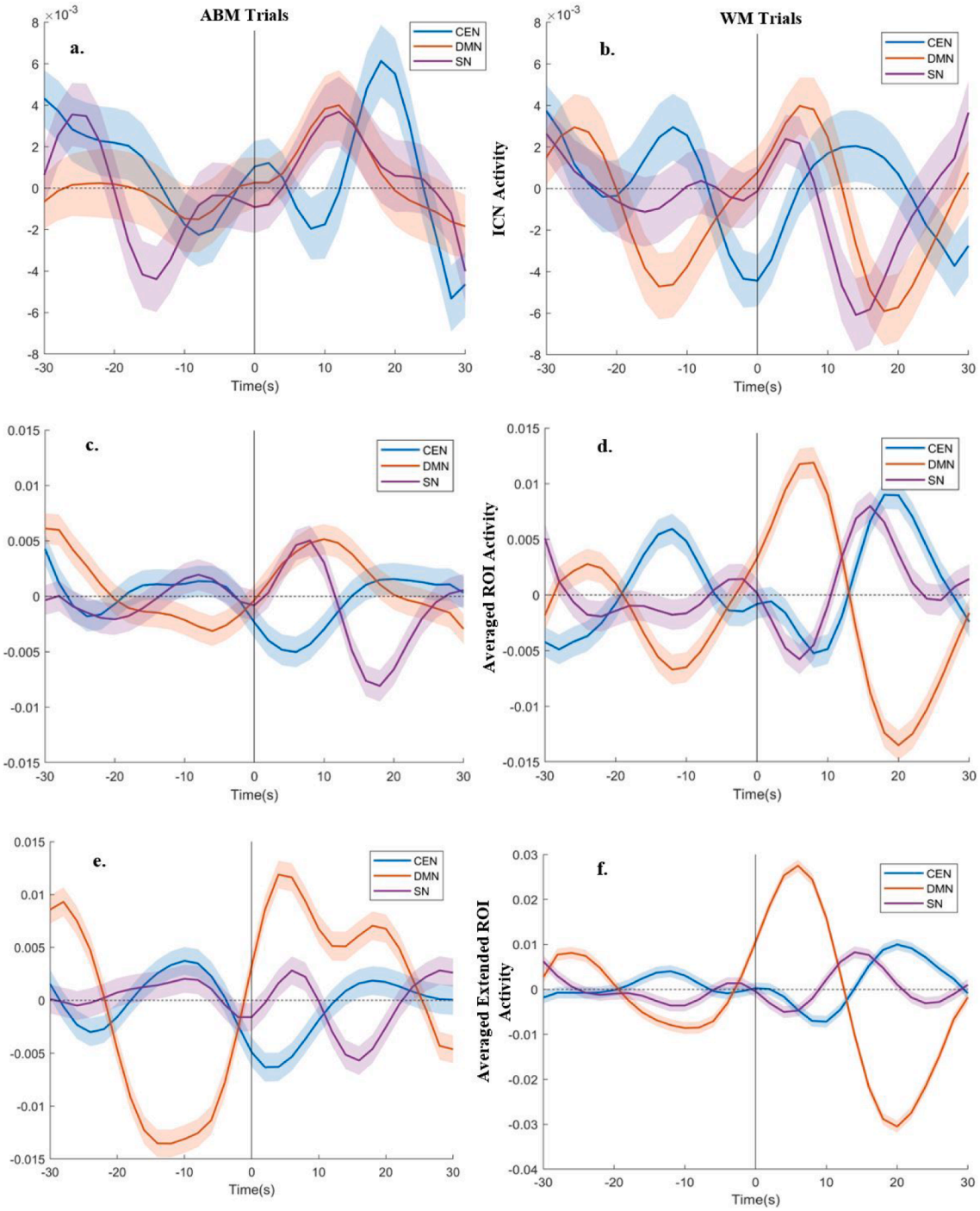

Fig. 4. Activity timecourses of the ICA networks corresponding to the CEN (blue), DMN (orange) and SN (purple), averaged over all a. ABM trials, and b. WM trials. Time is shown with respect to the task onset (time = 0s). BOLD activity timecourses of the CEN (blue), DMN (purple) and SN (orange), computed from activity of the corresponding nodes, shown averaged over all c. ABM trials, and d. WM trials. BOLD activity timecourses of the CEN (blue), DMN (purple) and SN (orange), computed from activity of the corresponding nodes using the extended set of ROIs, averaged over all e. ABM trials, and f. WM trials. BOLD activity is shown in units of percent signal change from mean activity (represented by the dotted line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

network definition introduced a second peak in DMN activity in the latter half of the trials around 18s post-task onset (Fig. 4e). This peak coincided with the peak in CEN activity during the ABM trials and was not accompanied by a corresponding SN peak. This second peak could be attributed to the increase in activity within the MTL nodes, as seen in Fig. S3b. Interestingly, this late DMN peak was also accompanied by an increase in PI activity (panel c), although this increase was not sufficient to increase the overall SN activity seen in Fig. 4e. In addition to this late peak in DMN activity, the MTL nodes also showed an increase in activity during the first 5-6 s of the trial, co-activating with the increase in PCC activity during the same time period (panel b).

This early increase in MTL node activity was also observed during the WM trials, however, in contrast to the ABM trials, there was no increase in MTL activity in the latter half of the trials (panel e). The PI activity was observed to be elevated during the early increase in MTL node activity, followed by a decrease in activity during the latter half of the WM

trials dominated by bilateral AI activity (panel f).

In conclusion, the findings including the extended set of DMN and SN nodes indicate the involvement of key memory-linked MTL nodes in the cognitive processes required during the ABM and WM trials. A common pattern of increased MTL activity was observed during the first half of both ABM and WM trials, that correlated with a similar increase in PCC and PreC activity, potentially indicating that some of the common component processes between the two tasks relied on key MTL nodes and could be related to memory processing.

3.3. Causality analysis of tri-network activity

To appropriately understand the causal influence of the above described sub-networks and ROIs, we conducted multivariate Granger causality analyses and describe them below.

3.3.1. ICA Networks - Multivariate Granger Causality (MVGC) Analysis

MVGC analyses revealed that the optimal model order for estimating the VAR was 3–4 time points using an AIC criterion, implying that lagging signals most likely had a lag of 6 to 8 s (model order x TR). Furthermore, the MVGC connectivity patterns generally supported the observations made in the previous sections.

ABM Trials - During the onset of the ABM Trials, the strength of the pairwise-conditional causal outflow from the anterior SN sequentially increased to the left CEN, bilateral CEN, the ventral DMN and then the dorsal DMN. The anterior SN to ventral DMN direction (Fig. 5) causality was at its peak around the same time as the peak in ventral DMN activity (Fig. S1b) post-task onset, also seen in the sub-network graphs from time t = 1s through t = 7s in Fig. 5. This observed increase in causality accompanying the correlated increase in anterior SN and ventral DMN activity could be indicative of the anterior SN recruiting the ventral DMN sub-network post-ABM task onset.

Net causal inflow to the ventral DMN increased during the first few seconds (t = 1s to t = 15s) post-task onset, primarily due to an increase in anterior SN to ventral DMN causality. It was also supplemented by an early increase in input from the posterior SN (t = 1s to 3s), followed by an increase in bilateral CEN input (t = 3s to 7s). The direction of net causal flow to the ventral DMN switched from primarily causal inflow to primarily causal outflow around t = 7s, and lasted until t = 17s. This pattern of causal flow might represent the recruitment of the ventral DMN sub-network by the anterior SN during the first few seconds of the ABM trials, followed by DMN-generated recruitment of other regions during AM memory retrieval and elaboration.

The peak in anterior SN to ventral DMN causality was preceded by a biphasic peak in anterior SN to bilateral CEN causality, followed by a peak in anterior SN to dorsal DMN causality. However, the activity of these sub-networks were anti-correlated with that of the anterior SN (Fig. S1), potentially indicative of the anterior SN's role in suppressing dorsal DMN and bilateral CEN activity when recruiting the ventral DMN during the ABM trials.

The dorsal DMN sub-network was found to have a mix of incoming and outgoing connections to other sub-networks shortly after ABM-task onset, followed by predominantly incoming connections from the left CEN, anterior SN and posterior SN during the latter half of the ABM trials (t = 17s onwards), as seen in the corresponding panels in Fig. 6. These causal patterns for the dorsal DMN are distinctly different from that of the ventral DMN, indicating different ABM-linked activity in the nodes comprising these sub-networks.

WM Trials - After WM task onset, the pairwise conditional causality from anterior SN, ventral DMN and posterior SN to left CEN increased. The causality from left CEN to bilateral CEN also increased along with a concomitant increase in bilateral CEN activity (Fig. S1d). Furthermore, the causality from posterior SN to dorsal DMN also increased along with an increase in dorsal DMN activity. These changes in pairwise causality

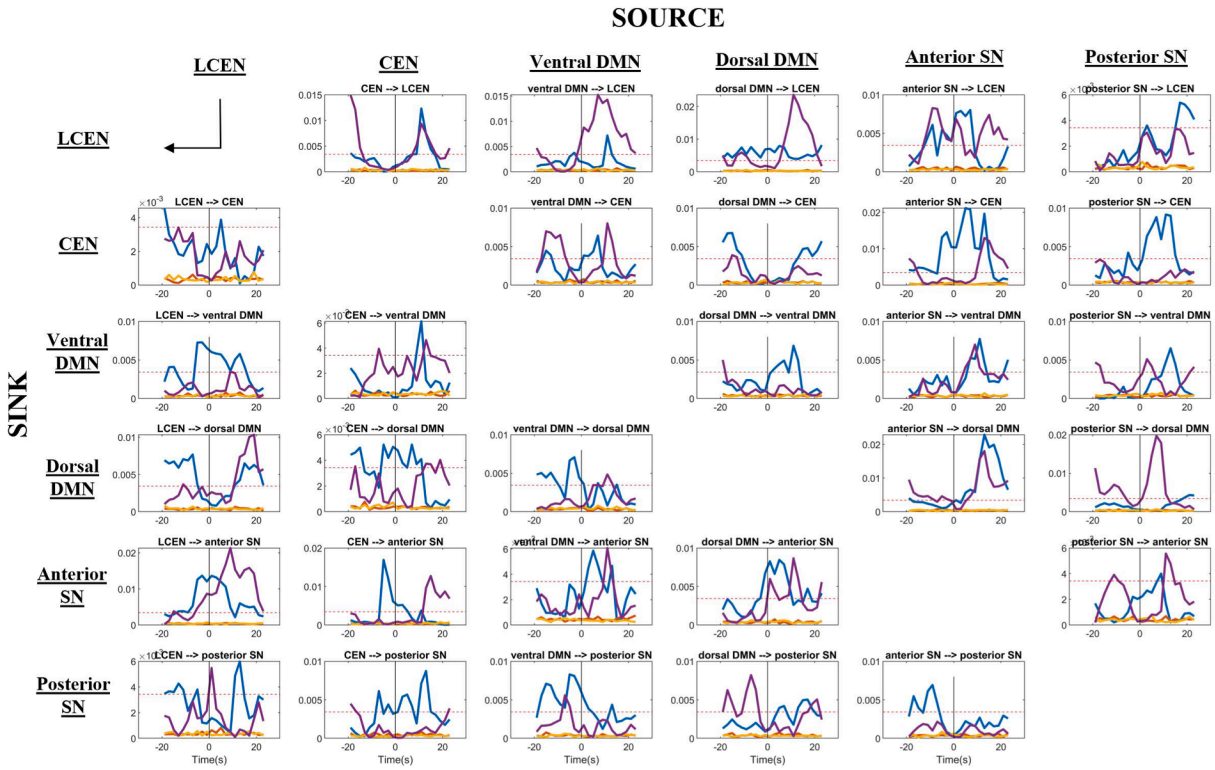

Fig. 5. MVGC values between the six sub-networks, as they vary across the ABM and WM trials, shown in blue and purple respectively. Each MVGC value is estimated using sliding window of width 14 s and step 2 s. Null MVGC values are also shown (orange) along with the threshold for FDR-corrected significance at $p < 0.05$ (dotted red line). Each column is a source node, while each row is a sink node, with each row-column intersection representing the connection going from the corresponding source node to the sink node. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

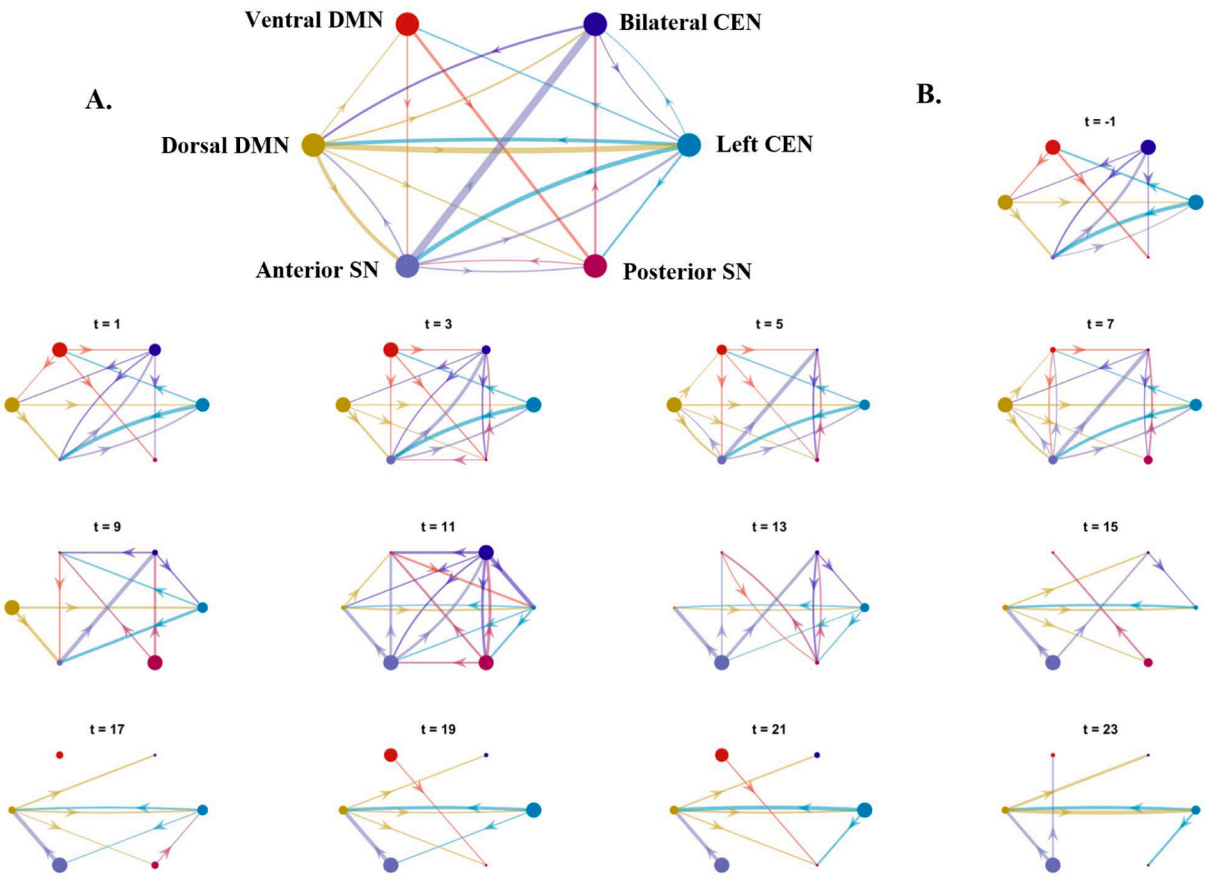

Fig. 6. MVGC connections between the six sub-networks, shown for all ABM trials. Connections passing FDR correction at $p < 0.05$ level are shown. A. The MVGC connections estimated using the entire trial duration; B. The MVGC connections estimated using a sliding window of width 14 s and step 2 s. The time label for each graph corresponds to the center of each window. The sizes of each node is scaled according to the MVGC weighted out-degree/in-degree. The thickness of each edge is scaled according to the MVGC value representing the edge.

contributed to the co-activation of bilateral CEN, dorsal DMN and the posterior SN during the first 10 s of WM trials (Fig. S1 d, e and f). This pattern is also observed in the latter half of the ABM trials, potentially indicative of CEN-mediated working memory maintenance processes common to both ABM and WM trials.

The pairwise-conditional causality from ventral DMN to left CEN also increased close to the task onset, continuing to increase until 10 s. Accompanied by an increase in causal inflow from dorsal DMN and bilateral CEN, this peak in causal input to the left CEN coincided with an increase in left CEN activity (Fig. S1d), which persisted through the rest of the trial. This causal pattern is indicative of left CEN activation, potentially for the maintenance of working memory.

In sum the above results suggest the anterior SN sub-network influenced the key task-relevant networks, while the posterior SN sub-network consistently interacted with working memory-linked sub-networks common to both ABM and WM tasks. The observed causal patterns also emphasise the different functional roles of the bilateral CEN and left CEN sub-networks, potentially involved in initializing working memory processes and the maintenance of working memory, respectively.

3.3.2. ROI-to-ROI - Multivariate Granger Causality (MVGC) Analysis

In addition to exploring the causality of data-driven ICA sub-networks, we conducted MVGC analysis on the ROI-to-ROI data to understand causal interactions between an abridged set of widely described nodes of tri-network model, as well as an extended set of nodes representing the tri-network model.

The AIC criterion analysis found 4 time points to be the optimal lag for the VAR model used in the MVGC analysis. This corresponded to an average time lag of 8 s for the pairwise conditional Granger causality between two nodes. The timecourses of the MVGC estimates between each pair of nodes for the ABM and WM trials are shown in Fig. S4. The connectivity between this abridged set of tri-network nodes is further visualized from t = -1 s to t = 23 s post-task onset for the ABM and WM trials in Figs. 8 and 9 respectively. This is contrasted with the temporal changes in connectivity between the extended set of tri-network nodes, shown in Figs. S5 and S6 for the ABM and WM trials respectively. The size of each node in these graphs corresponds to their net causal outflow (outflow - inflow), which is shown separately for the ABM and WM trials in Fig. 10. The hubness score of each node is also shown in Fig. 11. Key observations from these results are discussed below.

ABM Trials - The DMN node with maximal outgoing causal influence during the ABM trials was the vmPFC for the model using the abridged set of tri-network nodes, with outgoing connections to dACC and left AI of the SN during the first 8–10 s post-task onset (solid lines in Fig. 10b). During this time period, a bi-directional connection between vmPFC and the right PPC of the CEN was also observed, with the vmPFC to right PPC direction being much stronger than the rPPC to vmPFC direction. Combined with the peak in vmPFC activity (Fig. S2b) and the high hubness score (solid line in Fig. 11b) during the same time period, these causal patterns might represent vmPFC-driven memory processes that also rely on posterior parietal cortex and left anterior insular nodes.

While the amPFC showed similar, albeit weaker causal influence on other nodes, the observed casual pattern of the PCC was markedly different. The PCC was the DMN node with maximal incoming causal influence during the first 8–10 s of the ABM trials, receiving input from

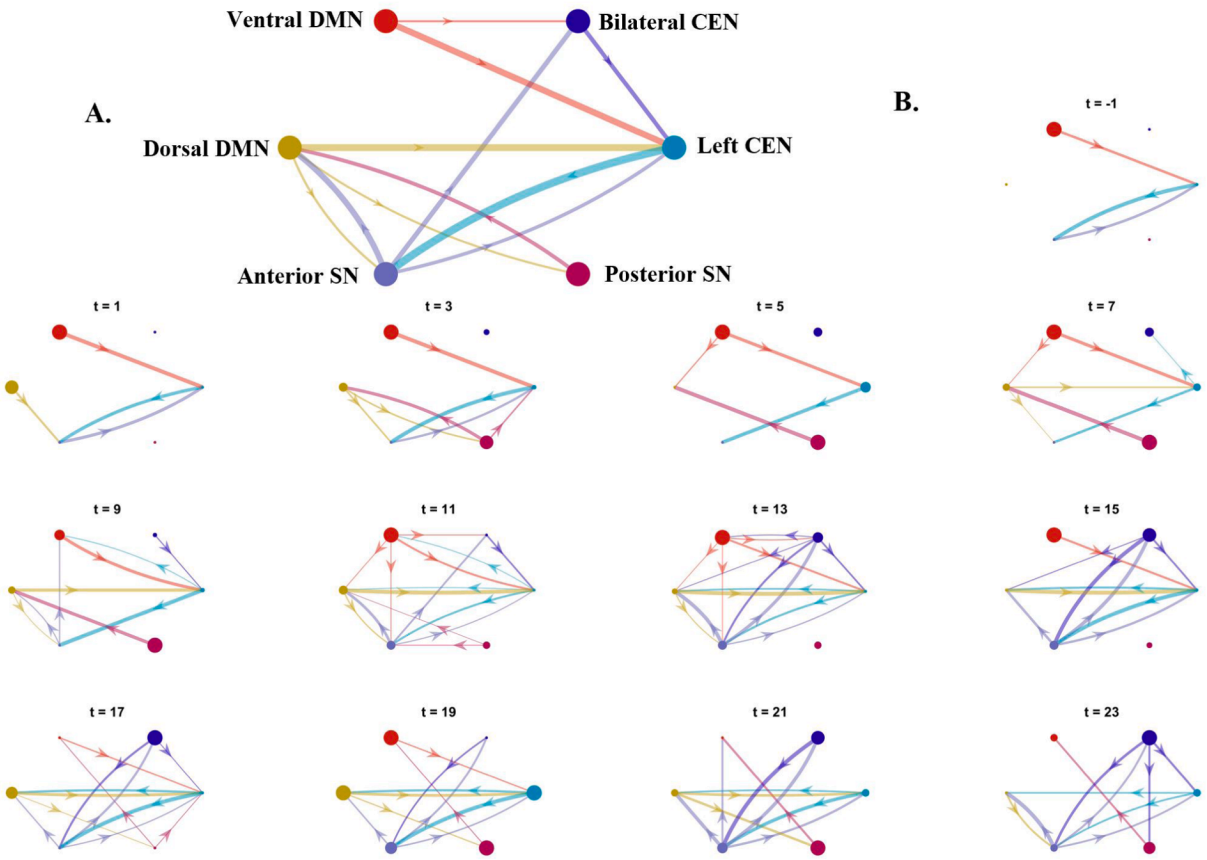

Fig. 7. MVGC connections between the six sub-networks, shown for all WM trials. Connections passing FDR correction at $p < 0.05$ level are shown. A. The MVGC connections estimated using the entire trial duration; B. The MVGC connections estimated using a sliding window of width 14 s and step 2 s. The time label for each graph corresponds to the center of each window. The sizes of each node is scaled according to the MVGC weighted out-degree/in-degree. The thickness of each edge is scaled according to the MVGC value representing the edge.

the left AI of the SN and the left PPC of the CEN. This is also reflected in the low hubness score of the PCC (solid line in Fig. 11b) and follows the predominantly anterior to posterior pattern of causal influence during this time period.

The SN nodes showed mostly incoming connections during the first 8 s post-task onset, compared to primarily outgoing connections during the rest of the trial. The early incoming connections were predominantly from the vmPFC, adding further evidence for vmPFC-driven memory processes during the early period of ABM trials. The CEN nodes were sparsely connected to other nodes during the first half of the ABM trials, however, showed increased PPC-driven causal influence during the latter half of the ABM trials, potentially indicative of the involvement of these nodes in maintenance of ABM.

These causal patterns were further investigated using the extended set of tri-network nodes, revealing that vmPFC remained the dominant DMN hub with maximal causal outflow during the first half of the ABM trials, and was followed by extensive causal outflow from the amPFC towards the end of the ABM trials (solid lines in Fig. 10e). The causal influence from these mPFC nodes was also accompanied by high causal outflow from the right HC node throughout the ABM trials. Additionally, the newly added MTL nodes also received strong incoming connections from other DMN and SN nodes during the first 10 s of ABM trials (Fig. S5). This is seen in Fig. 10e, in the form of negative net causal flow into these nodes, and in Fig. 11e, as high hubness scores for bilateral HC, pHC and RSC nodes during this time period. Furthermore, the PCC is found to be heavily connected to other tri-network nodes during the same time period, resulting in hub-like behaviour shortly after ABM task onset (Fig. 11e). These results indicate highly connected hippocampal and parahippocampal hubs, that work together with a PCC hub and a mPFC hub during the first half of the ABM task, switching to a predominantly mPFC hub during the second half of the ABM trials.

The SN nodes showed numerous connections to other DMN and CEN nodes during the first half of the ABM trials, followed by a widespread increase in the hub-like properties of all SN nodes during the latter half of the ABM trials (solid lines in Fig. 11d). This included the newly added PI nodes and coincided with the increase in PI activity and the second peak in MTL activity observed in Fig. S3c and b, respectively. These findings might represent extensive PI-mediated processing throughout the ABM trials, and show its collaborative integration with other SN nodes.

Lastly, the addition of the MTL and PI nodes to the tri-network model led to an increase in hub-like behaviour within CEN nodes, with right PPC and dlPFC showing increased hubness score during the first half of the ABM trials, followed by a second peak in the hub-like properties of the right PPC during the latter half of the ABM trials (solid lines in Fig. 11f). The continued involvement of these posterior parietal CEN nodes throughout the ABM trials might indicate a multi-functional role of these nodes in working memory and autobiographical memory recall.

Collectively, the MVGC results support an mPFC-based hub involved in autobiographical memory retrieval, that dynamically interacts with MTL nodes and other hubs anchored within SN nodes, and the regions within the parietal cortex, such as the PPC, RSC and PCC.

WM Trials - Similar to the ABM trials, PCC of the DMN received significant causal input from the left AI of the SN during the first 8–10 s of the task trials, corresponding with an increase in PCC activity during this time (Fig. S2e), potentially representative of common memory-linked processes required during both ABM and WM trials.

However, contrary to the ABM trials, all SN nodes of bilateral AI and

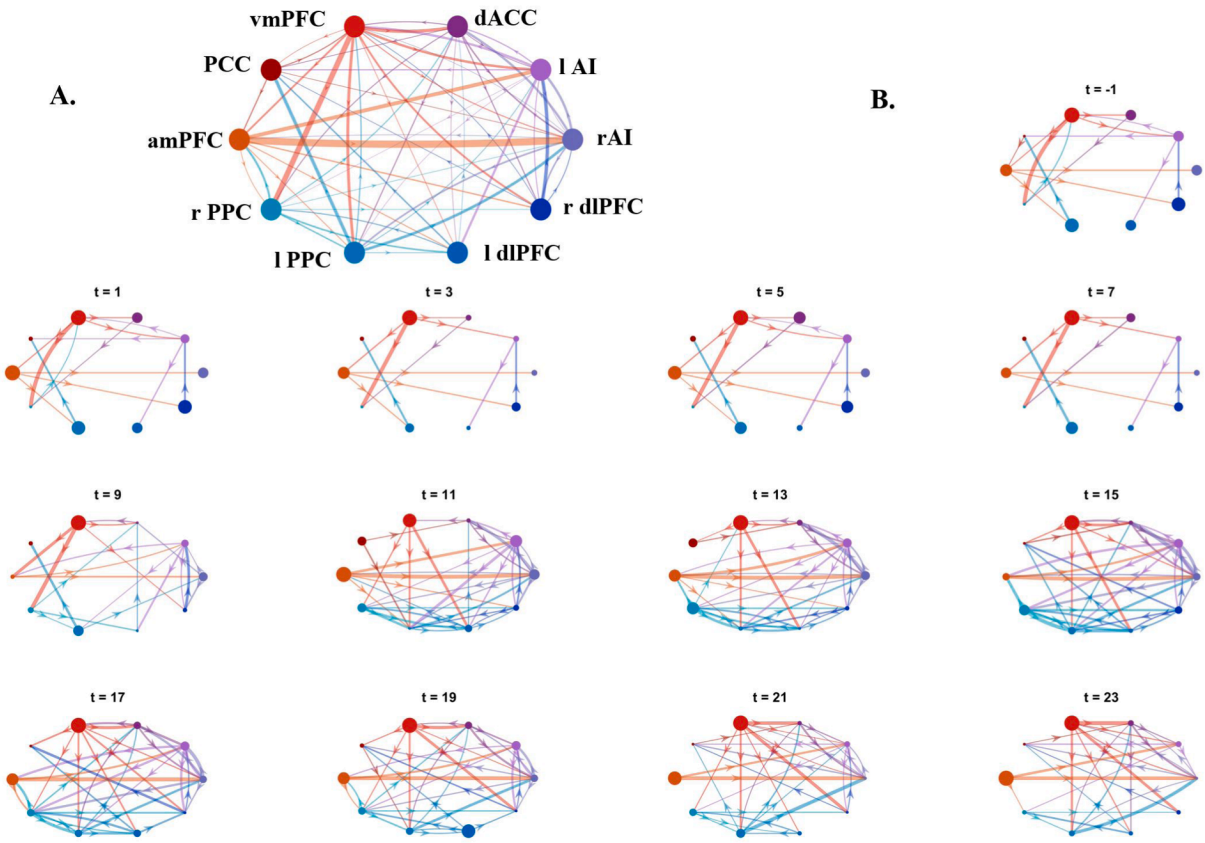

Fig. 8. MVGC connections between the network nodes, shown for all ABM trials. Only connections passing FDR correction at $p < 0.05$ level are shown. A. The MVGC connections estimated using the entire trial duration; B. The MVGC connections were estimated using a sliding window of width 14 s and step 2 s. The time label for each graph corresponds to the center of each window. The size of each node is scaled according to the MVGC weighted out-degree/in-degree, whereas the thickness of each edge is scaled according to the magnitude of the corresponding MVGC value. Furthermore, the color of each edge corresponds to that of the source node and, along with the direction of the arrowhead, represents the edge direction.

dACC exerted extensive causal influence on left CEN structures, such as PPC and dlPFC, and frontal DMN nodes such as amPFC and vmPFC during the first half of the WM task. This is reflected in the high net causal outflow from SN nodes, seen in Fig. 10a (dotted lines), and high hubness score of SN nodes during this time period (dotted lines in Fig. 11a). Combined with the subsequent decrease in amPFC activity after the first 8 s of the WM trials (Fig. S2e), such causal patterns might represent the inhibition of DMN nodes, while the CEN nodes were recruited for executive processing.

Furthermore, in the first half of the WM trials, the CEN nodes of left dlPFC and left PPC received causal input from bilateral AI and dACC of the SN respectively. This was followed by extensive outgoing causal influence from bilateral PPC and right dlPFC during the latter half of the WM trials, leading to an increase in its net causal outflow and hub-like behaviour shown using dotted lines in Figs. 10c and 11c, respectively. This might be representative of working memory processes mediated by the PPC and dlPFC nodes of the CEN, reaching its peak during the latter half of the WM trials.

Reanalyzing these causal patterns with the extended set of tri-network nodes revealed some finer details in the dynamic causality of the tri-network nodes. Similar to the ABM trials, the RSC was found to be a major hub in the first half of the WM trials (dotted line in Fig. 11e), however, unlike the ABM trials, this was the result of primarily outgoing causality from the RSC to other tri-network nodes, as seen in Figs. S6 and 10e. This was also accompanied by some hub-like behaviour from right HC and the mPFC nodes during this time period. This was followed by extensive incoming causal connections to the PreC, which in turn connected with other MTL nodes and posterior parietal CEN nodes during the second half of the WM trials, seen in Figs. S6 and 10e. This led to an increase in the hub-like behaviour of the PreC, left HC and left pHC, potentially indicating a shift from RSC-centric processing early in WM trials, to more PreC-centric processing within the DMN in the latter half of the WM trials.

The newly added PI nodes also showed extensive hub-like behaviour shortly after WM task onset, alongside the bilateral AI and dACC nodes, that continued throughout the WM trials (dotted line in Fig. 11d). High causal outflow was observed from these PI nodes to CEN nodes such as the dlPFC, and DMN nodes including the MTL and mPFC nodes, seen in Fig. S6. These causal properties of the PI nodes added to the previously observed patterns of task-linked SN influence on DMN and CEN nodes and might warrant its inclusion in the standard set of nodes used to define the SN within the tri-network model.

Lastly, the causal patterns of the CEN nodes using the extended set of tri-network nodes were similar to that using the abridged set of tri-network nodes, with the right PPC and dlPFC showing extensive causal outflow during the first half of the WM trials, switching to mostly incoming causal links in the second half. In contrast, the left PPC and dlPFC displayed mostly incoming causal connections in the first half of the WM task, followed by some outgoing connections in the latter half of the WM trials (dotted lines in Fig. 10f). This resulted in high left dlPFC hubness score shortly after WM task-onset, followed by an increase in the hub-like properties of the remaining CEN nodes during the last 15 s of the WM trials, coinciding with a similar increase in hub-like properties of the PreC and left HC and pHC nodes, as seen in Figs. 11f and e. These results show significant functional lateralization of working memory processing within CEN nodes, and the dynamically shifting hub-like properties within the CEN nodes, which coincides with that of other DMN and MTL nodes during the second half of the WM trials.

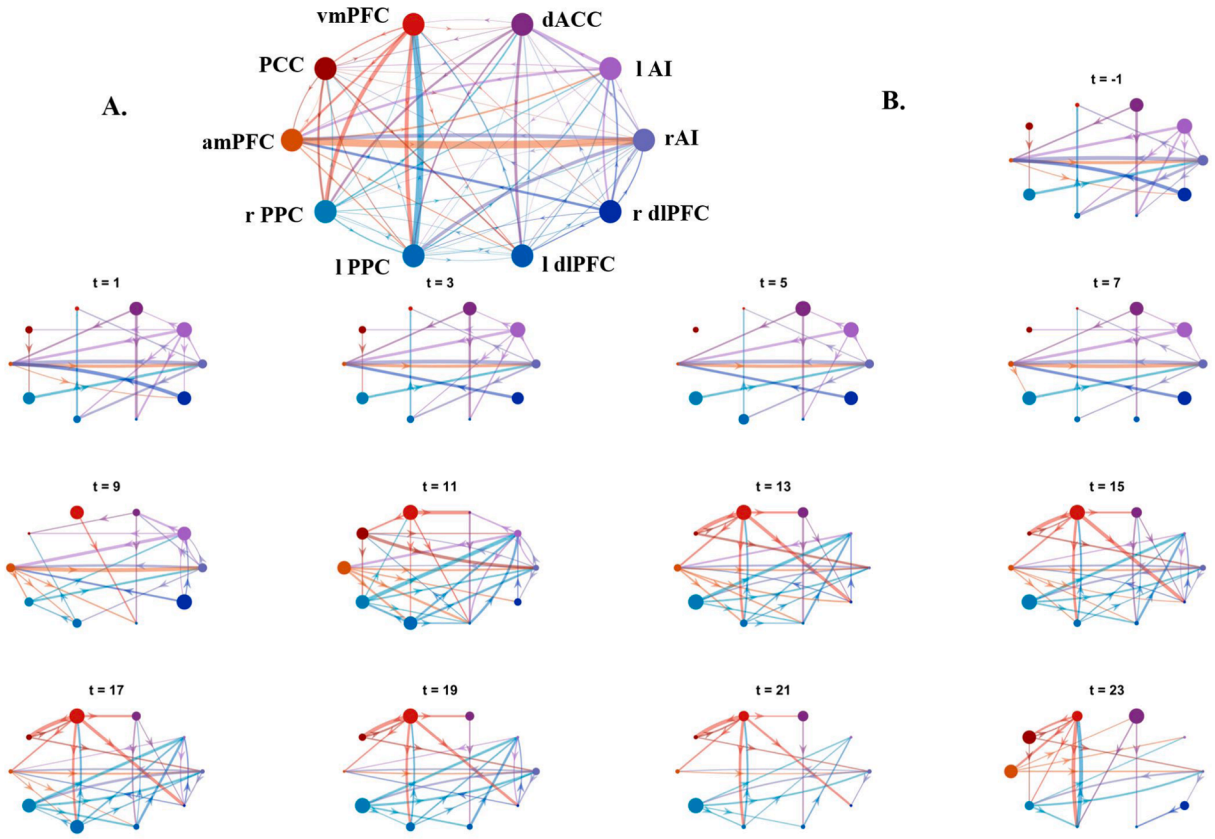

Fig. 9. MVGC connections between the network nodes, shown for all WM trials. Only connections passing FDR correction at $p < 0.05$ level are shown. A. The MVGC connections estimated using the entire trial duration; B. The MVGC connections were estimated using a sliding window of width 14 s and step 2 s. The time label for each graph corresponds to the center of each window. The sizes of each node is scaled according to the MVGC weighted out-degree/in-degree, whereas the thickness of each edge is scaled according to the magnitude of the corresponding MVGC value. Furthermore, the color of each edge corresponds to that of the source node and, along with the direction of the arrowhead, represents the edge direction.

In conclusion, unlike the ABM trials, bilateral AI and PI showed hub-like causal properties indicative of some DMN node suppression while recruiting CEN nodes for working memory processes shortly after WM task-onset, accompanied by a left dlPFC-centric hub and RSC-centric hub. This transitioned to a highly connected network structure during the latter half of the WM trials, with some HC, pHC, mPFC, PreC and all CEN nodes displaying hub-like properties, potentially indicative of working memory maintenance processes. The results also support lateralization of WM processes between left and right CEN nodes.

4. Discussion

The triple network model of Menon (2011) postulates that network switching between the default mode network (DMN) and central executive network (CEN), gated by the salience network (SN), is essential for everyday tasks requiring switching between internally directed and externally directed thought processes. This study adds to the growing body of evidence showing the gating function of SN while dynamically switching between two tasks specifically designed to engage these three networks. It additionally investigates the differential role of the various SN sub-networks and sub-nodes in the context of a complex switching task that utilized an autobiographical memory recall task and a 2-back working memory task to activate the DMN and CEN respectively.

The data-driven analyses performed to characterize spatial patterns of tri-network activity during the ABM and WM trials (study aim 1) identified a set of brain regions and sub-networks that are known to be associated with a wide range of functions necessary for adequate task performance, as discussed in the following paragraphs.

Autobiographical memory (ABM) retrieval is a complex multiphasic process with an initial memory instantiation/retrieval phase that is subserved by the right hippocampus, right/medial prefrontal cortices and retrosplenial cortex, followed by an elaboration phase that is found to be associated with increased activity in the left prefrontal cortex, precuneus and visual cortex (Daselaar et al., 2008). Each of these phases can take up to 12 s (Daselaar et al., 2008), with memory retrieval being faster for general autobiographical memories compared to specific ones (Addis et al., 2004). Furthermore, according to Kim (2012)'s dual-subsystem model of the DMN, cortical midline nodes such as amPFC and PCC are involved in the self-related processing aspects of autobiographical memory (ABM) recall, while the parieto-temporal nodes, including inferior parietal lobule and medial/lateral temporal cortices, are associated with the memory retrieval aspects of the ABM task. The precuneus is also implicated in ABM retrieval (Svoboda et al., 2006; Addis et al., 2004). More specifically, the BBB model postulates a central role for the precuneus in representing the products of memory retrieval and mental imagery (Byrne et al., 2007). Our MVPA results (Fig. 2 and Table 1) are in agreement with these findings showing activation of the medio-temporal subnetwork of the DMN, including the inferior parietal lobule of the supramarginal gyrus (SMG) and the PCC, during autobiographical memory recall (Andrews-Hanna et al., 2014) trials. The involvement of the parietal cortical structures such as bilateral anterior supramarginal gyri (aSMG) in autobiographical memory recall align with its involvement in self-referential processing tasks vs non self-referential processing tasks (Axelrod et al., 2017), and redirecting attention on internal representations (Sestieri et al., 2011). Furthermore, the observed ABM-linked activation in parietal opercular structures show involvement of secondary somatosensory areas during autobiographical memory recall, providing evidence for distributed

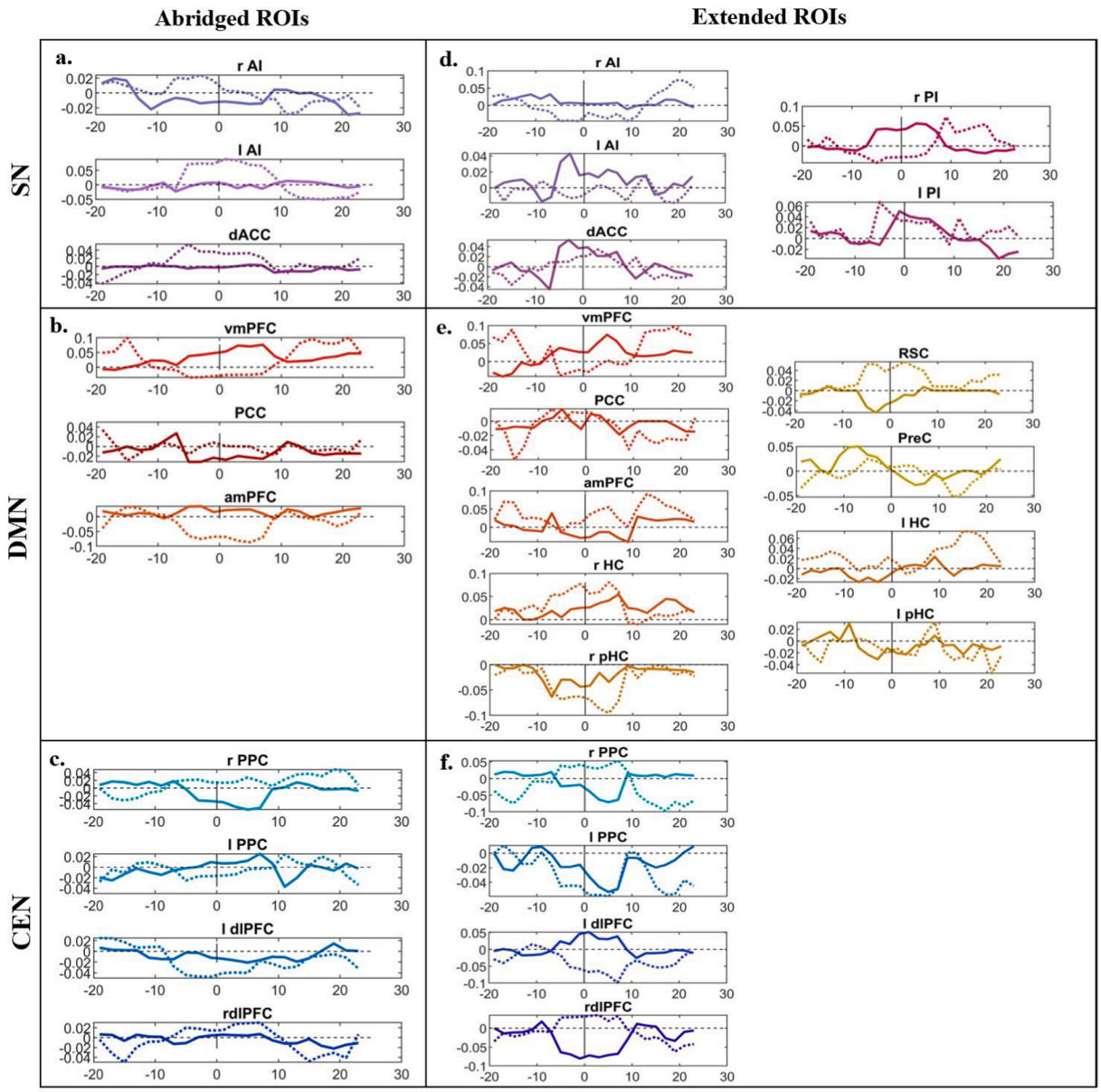

Fig. 10. The net pairwise causal outflow from each network node (causal outflow - causal inflow) for the abridged set of a. SN, b. DMN and c. CEN nodes, alongside the causal outflow from the extended set of d. SN, e. DMN and f. CEN nodes. Each panel shows the time course of the causal outflow from a single network node, during the ABM (solid line) and WM trials (dashed line). A positive value corresponds to a new causal outflow and a negative value corresponds to a net causal inflow. Note the different range of causal outflow values for each network node. This allows for adequate visualization of the temporal patterns in causal outflow for all network nodes.

constructive processes during autobiographical memory recall.

On the other hand, the N-back task is a popular working memory task that relies on the multi-component executive process of temporal coding (Collette and Van der Linden, 2002; Smith and Jonides, 1997), involving frontal brain regions such as the frontal poles, dlPFC, vlPFC, and latero-medial premotor cortices, in addition to posterior parietal regions such as medio-lateral posterior parietal cortices, dorsal cingulate (Owen et al., 2005; Yaple et al., 2019), and the precuneus (Yaple et al., 2019). The observed activation of the dlPFC in the MidFG (Fig. 2 and Table 1) during the working memory trials indicates activity of this core fronto-parietal CEN subsystem (Yaple et al., 2019). Furthermore, the engagement of supplementary motor areas might indicate the involvement of sub-vocal rehearsal structures (Smith and Jonides, 1997; Fegen et al., 2015) for maintaining the presented words in working memory through articulatory rehearsal processes (Vallar and Pagano, 2002).

The ICA analyses identified data-driven differences in activity, finding two functionally distinct sub-networks within each of the three networks (Fig. 3). These sub-networks significantly overlapped with the networks presented in Shirer et al. (2012) and showed distinct activation patterns between the ABM and WM trials, indicating functionally distinct activation of various nodes within each of the networks to accomplish the complex ABM and WM tasks.

Characterization of the temporal and causal patterns of tri-network (study aims 1 and 2 respectively) yielded five key findings, which are discussed below.

4.1. *Different DMN and CEN nodes and sub-networks activated during both task conditions*

Both DMN and CEN were observed to be active during the ABM and WM trials, although different nodes and sub-networks of the networks were predominantly active during each task (Fig. 4).

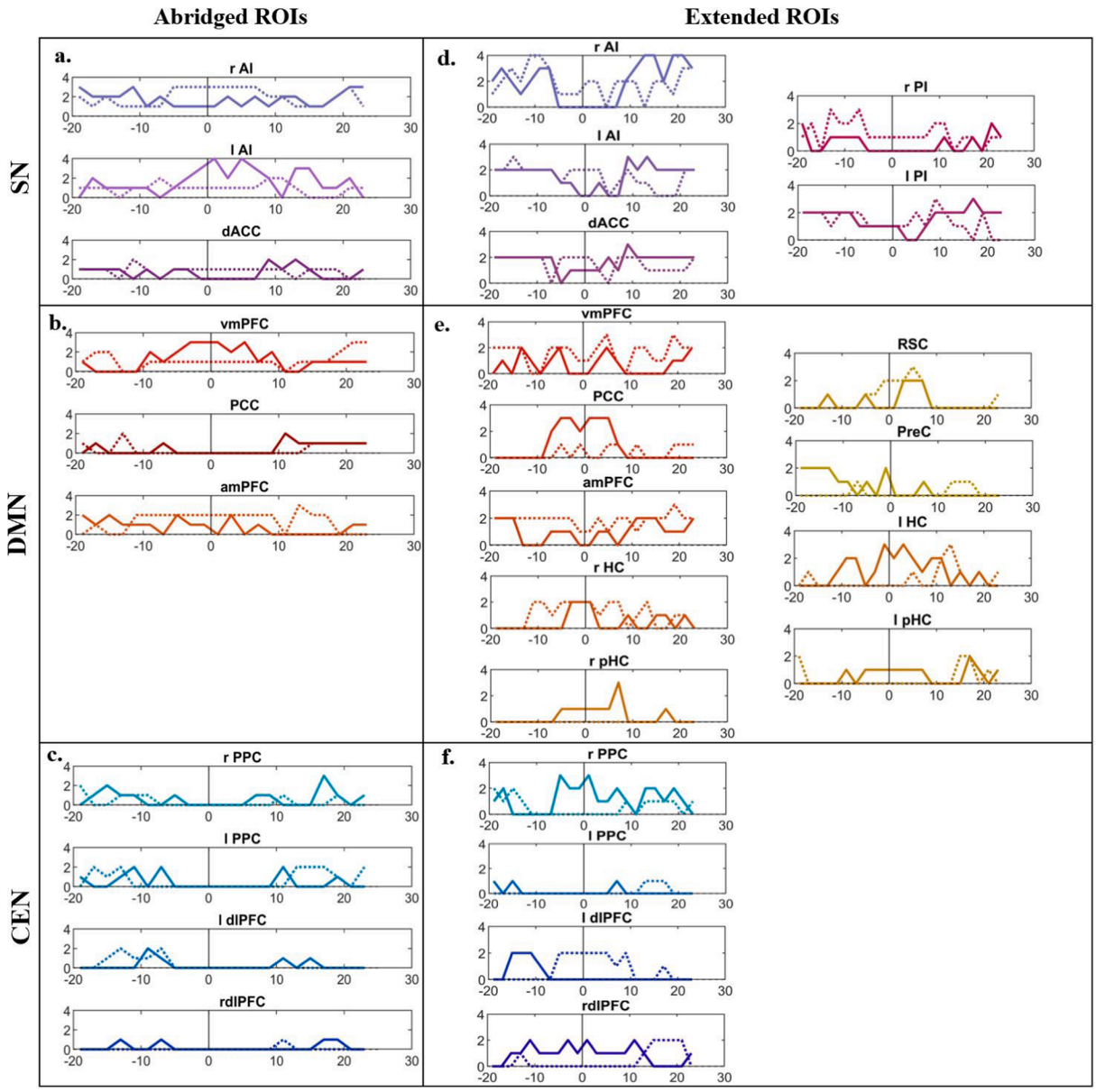

Fig. 11. The hubness score for each network node for the abridged set of a. SN, b. DMN and c. CEN nodes, alongside the hubness score for the extended set of d. SN, e. DMN and f. CEN nodes. Each panel shows the time course of the hubness score from a single network node, during the ABM (solid line) and WM trials (dashed line). The hubness scores range from 0 to 4 and is scored according to the procedure described in Section 2.3.4.

DMN activity during ABM trials was expected due to its key involvement in self-related processing and memory retrieval (Spreng et al., 2010), however, contrary to expectations, the CEN was also observed to be active during the ABM trials, as seen in Figs. 4a and S1a, where the activity of the CEN network and its sub-networks increased during the latter half of the ABM trials. Fig. S2a shows that the observed increase in CEN activity was due to an underlying increase in dlPFC activity which might have been due to the working memory needed to maintain the retrieved memory throughout the trial duration, once it had been recalled. This increase in dlPFC activity began around 8–12 s post-task onset, aligning with the time required to fully retrieve an ABM (Addis et al., 2004; Daselaar et al., 2008). This could also have been due to the involvement of executive processing during autobiographical memory recall (Unsworth et al., 2012) used for searching autobiographical information. Such a pattern of DMN recruitment at the onset of a DMN-linked task, followed by CEN recruitment is also seen in creative idea production (Beaty et al., 2015).

Furthermore, the hippocampus and associated medial temporal lobe structures (MTL) co-activated with the aforementioned peak in dlPFC and PPC activity (Fig. S3a and b). This was accompanied by an increase in hub-like behaviour within the HC, pHC and RSC nodes shortly after ABM task-onset, followed by a second increase in hub-like activity of the right PPC and some MTL nodes during the latter half of the ABM trials (Fig. 11e and f). These results are consistent with the BBB model that implicates the hippocampus and associated MTL structures in the retrieval and reconstruction of details associated with a particular autobiographical memory, followed by its representation within parietal lobe structures, forming mental imagery representations within the "parietal window" that enables conscious access to these products of memory retrieval (Byrne et al., 2007).

During WM trials, the DMN was also found to be active in addition to the expected CEN activity. More specifically, the dorsal DMN subnetwork was found to be active during the first half of the WM trials, compared to the ventral DMN being predominantly active during the ABM trials (Fig. S1e). This difference can be attributed to the higher activity of the amPFC (a node of the dorsal DMN) during the WM trials,

in contrast to the ABM trials (Fig. S2e). Another contributing factor could have been the participation of posterior parietal nodes of the dorsal DMN sub-network (Fig. 3), which are also associated with n-back working memory tasks (Owen et al., 2005). These posterior parietal nodes also showed increased hub-like behaviour towards the latter half of the WM trials (Fig. 11f), a causal pattern that was also observed within the precuneus and some other MTL nodes (Fig. 11e), consistent with the role of these parietal structures in maintaining mental images of items in working memory (Byrne et al., 2007).

Interestingly, the PCC node of the DMN also showed increased activity during the initial phase of the WM task, followed by a large decrease in activity when DMN nodes increased their activity later in the WM trials (Fig. S2e). PCC activity was also observed to increase during the initial 8–10 s of the ABM trials, followed by a subsequent decrease in activity, however, not decreasing to the extent seen during the WM trials. Although seemingly contrary to the expected inactivation of PCC during WM, the observed pattern is consistent with the role of PCC in self-related processing, while also participating in cognitive control tasks under high task load (Leech et al., 2011).

4.2. The SN recruited the task-appropriate network by synchronizing its activity with the desired network

Despite both DMN and CEN being active during both conditions, global SN activity was found to correlate with the task-appropriate network, and anti-correlate with the "task-opposite" network. Fig. 4c/d show that SN activity was correlated with the early peak in DMN activity during ABM trials and the late peak in CEN activity during WM trials and that it was anti-correlated with the later peak in CEN activity during ABM trials and the early peak in DMN activity during WM trials. The peaks in global SN activity were also observed to lead the corresponding peaks in DMN or CEN activity by 2-4 s (Fig. 4c/d). These findings, combined with the observed causal patterns of SN nodes driving the DMN and CEN nodes (Fig. S4), indicate that the SN might have been gating the activity of the corresponding network by synchronizing its activity with that of the task-appropriate network. Interestingly, this pattern of network co-activation is similar to that observed within the dual-network model of cognitive control, wherein the salience network (cingulo-opercular network) is observed to co-activate with the central executive network (fronto-parietal network) during executive task performance (Dosenbach et al., 2008). The dual-network model ascribes this pattern of co-activation to the set-maintenance role of the SN, which when combined with its role in salience assignment (Menon, 2011) could explain its gating function within the tri-network model. In this paradigm, the SN would use key salience signals to identify the task-appropriate network based on the interoceptive or exteroceptive nature of the task demands, and then co-activate with the identified network to stably maintain task sets during task performance. This effectively selects the task-appropriate network by providing stable set-maintenance to the network deemed appropriate for the task at hand, following the general principle of 'connection through coherence' (Fries, 2015), albeit in the context of network connectivity.

4.3. Different sub-networks of the SN showed different causal properties

Data driven ICA analysis revealed two distinct sub-networks that have significant overlap with SN nodes (Fig. 3), forming anterior and posterior salience sub-networks, similar to those described by Shirer et al. (2012). The anterior SN seemed to co-activate with the task-appropriate network in both ABM and WM trials, however, posterior SN seemed to correlate with bilateral CEN and dorsal DMN activity (Fig. S1). While the anterior SN causally influenced most of the other sub-networks, owing to its hub-like properties, the posterior SN network causally inhibited ventral DMN activity during the latter half of the ABM trials, while causally stimulating dorsal DMN activity during the first half of the WM trials (Fig. 6).

Posterior SN was also closely linked to the LCEN activity, providing causal input during both WM and ABM trials. This could be owing to its role in maintaining a representation of the passage of time (Wittmann et al., 2010), which could be important for successfully performing an N-back task by keeping track of the word seen 2 words ago. Posterior SN deactivation has also been observed in the case of 2-back verbal working memory (Sweet et al., 2008). This combination of posterior SN activation, owing to its time-keeping role, and its deactivation observed in n-back working memory tasks might explain the activation-inactivation pattern of posterior SN activity observed in Fig. S1c and f.

The posterior insula (PI) is one of the major constituent nodes of the posterior SN subnetwork, as previously described (Shirer et al., 2012) and seen in Fig. 3. However, its role in network switching has been relatively understudied compared to that of the AI (Menon and Uddin, 2010) and dACC (Crottaz-Herbette and Menon, 2006). The high hubness score of the PI and extensive integration with other SN, DMN and CEN nodes observed in this study (Figs. S5, S6, and 11d) indicate the dynamic role played by the PI in autobiographical memory retrieval and working memory processes. The PI is well positioned to integrate various aspects of salient information and guide network switching, given its multisensory inputs (Björnsdotter et al., 2009), and its connectivity with the emotionally salient ventral AI, and the cognitively salient dorsal AI (Davidovic et al., 2019). Dysfunctional PI activity has also been observed with abnormal cognitive behaviour, as seen in autism spectrum disorders (Ebisch et al., 2011) and PTSD (Nicholson et al., 2020).

4.4. Distinct causal roles of the right AI and the left AI

As has been extensively discussed in the literature, the right AI is a key hub region that modulates network switching by recruiting the CEN and inhibiting the DMN during externally directed tasks (Sridharan et al., 2008), as also seen in this study by the patterns of rAI to amPFC/vmPFC causality (Fig. 10, panel a and d, and Fig. 9). However, this study also found that the right AI was not the primary causal outflow hub during self-directed processing, and that the left AI alongside the dACC and vmPFC played a much larger role in recruiting other DMN nodes when required in the ABM trials (Fig. 10, panel a and d, and Fig. 8).

In light of the results from Sections 4.3 and 4.4, the bilateral insula seems to form a set of key nodes that modulate inter-network connectivity in a task-linked manner, and has been also observed to couple with the DMN at the onset of creative idea production (Beaty et al., 2015).

4.5. Left CEN nodes formed a separate cluster (from Bilateral CEN) that was uniquely activated for each of the two tasks

The data-driven ICA analyses revealed that left CEN nodes engaged in distinct connectivity patterns, compared to the rest of the CEN (Fig. 7), and was hence identified as an independent spatiotemporal component (Fig. 3) by the data-driven group-ICA algorithm. This was further supported by the ROI-to-ROI analyses revealing primarily left CEN nodes receiving incoming connections during the onset of WM trials (Figs. S4 and 10c/f). Left lateralization has been extensively documented in some key working memory sub-systems, most notably the subvocal rehearsal sub-systems of the phonological loop (Smith and Jonides, 1997; Fegen et al., 2015). This might also be the underlying reason for the observed increase in hubness score of the left dlPFC during the first half of the WM trials (Fig. 11f), nd the subsequent increase in activity of the left CEN sub-network during the latter half of the WM trials (Fig. S1).

These findings consolidate a normative view of network dynamics within Menon (2011)'s tri-network model, that had originally emerged from a large body of clinical research, suggesting that many psychopathologies can be viewed as a dysregulation of these three key networks. For example, an overly dominant DMN is associated with maladaptive rumination in depression (Hamilton et al., 2011), while dysregulated SN nodes lead to aberrant network switching in schizophrenia (Moran et al.,

2013), and Post-traumatic stress disorder (PTSD) (Lanius et al., 2015). PTSD is an especially relevant example with patients suffering from improper recruitment of the DMN during a working memory task (?), while their DMN was under-recruited during an autobiographical memory task (St. Jacques and De Brigard, 2015). Given the critical role of the SN and its sub-networks in healthy network-switching, such PTSD-linked network dysfunctions might be due to an aberrant SN causing disruptions in normal salience detection (Rabellino et al., 2015). This might lead to a breakdown of the SN-linked co-activation dynamics described in Section 4.2, which can provide a mechanistic explanation of how abnormal recruitment of task-opposite networks can lead to poor task performance in patients with PTSD. For example, the abnormal recruitment of DMN during a working memory task observed in PTSD patients (Daniels et al., 2010), is also associated with an absence of correlated activity between SN and CEN. This absence of SN-linked stable set-maintenance might result in the rapid decay of the working memory items maintained by the CEN, ultimately resulting in poor working memory performance in patients with PTSD. Such improved understanding of the mechanisms underlying healthy and dysfunctional SN-gated switching between DMN and CEN could inform the development of specialized treatment plans that directly target the dysfunctional tri-network dynamics (Lanius et al., 2015) in patients with various psychopathologies.

This study used MVGC analysis to probe the patterns of causal influence between the SN, CEN and DMN. However, some considerations should be kept in mind when using MVGC analysis with BOLD fMRI signals due to the poor time resolution of the BOLD signal, reliance of the BOLD signal on the haemodynamic response, and the susceptibility of GC estimates to poor signal-to-noise ratio (SNR) (Ramsey et al., 2010). Despite these concerns, GC estimates have been found to be robust in the context of haemodynamic response function confounds (Schippers and Keysers, 2011), and its multivariate version is found to appropriately represent the connectivity patterns in low density networks (Duggento et al., 2018). The impact of the other factors can be minimized by using sliding windows to maximize signal stationarity, performing “vertical regression” across trials that are expected to have similar stochastic generative processes, and comparing the MVGC estimates of the signal with those of the time-permuted surrogate signals to minimize spurious connectivity estimates, as performed here.

One caveat to the ROI analyses performed in this study is that the network membership assigned to inherently multi-functional nodes such as vmPFC and PPC is based on the definitions used by the current work surrounding this tri-network model (Menon, 2011). However, as evident from the BBB model, there is a complex interplay between fronto-parietal areas in encoding, maintaining and manipulating the products of memory retrieval (Byrne et al., 2007). For example, in addition to its role within the DMN, vmPFC is also an integral component of reward processing circuits and can be heavily involved in some executive tasks (Gerlach et al., 2013). The PPC, similarly, has a broader role in representing mental imagery, which is important for both working memory and maintaining the retrieved components of autobiographical memory (Byrne et al., 2007). This is supported by our results showing activation of these structures during the ABM trials (Fig. 2) and WM trials (Fig. S2). Thus nodes such as vmPFC and PPC can be expected to dynamically shift network membership, as a function of dynamic cognitive processing, and should be considered as cross-network nodes (Pedersen et al., 2018). As such, these multi-functional nodes are considered within the context of the tri-network model in this study, however, future studies could look into a dynamic version of the tri-network model with flexible node-membership.

In conclusion, results from the current study add to the growing body of evidence showing the complex interplay of CEN, DMN and SN nodes and sub-networks required for adequate task-switching. Furthermore, the discovery that the SN co-activates with the task-relevant network provides mechanistic insight into SN-mediated network selection in the context of explicit tasks. Finally, the results from this study indicate active involvement of the posterior insula and some MTL nodes in task-linked functions of the SN and DMN, warranting their inclusion as network nodes in future studies of the tri-network model.

Declarations of Competing Interest

None.

Acknowledgements

This research was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to SB, an Alexander Graham Bell Canada Graduate Scholarship (CGS-D) from NSERC to SBS, a Canadian Institutes of Health Research (CIHR) grant to MCM and an NSERC Discovery Discovery grant to JH. MCM is also supported by the Homewood Chair in Mental Health and Trauma at McMaster University.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.bandc.2021.105725.

References

Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: A partial least squares approach. *NeuroImage, 23*(4), 1460–1471.

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences, 1316*(1), 29–52.

Axelrod, V., Rees, G., & Bar, M. (2017). The default network and the combination of cognitive processes that mediate self-generated thought. *Nature Human Behaviour, 1*(12), 896–910.

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods, 223*, 50–68.

Beaty, R. E., Benedek, M., Barry Kaufman, S., & Silvia, P. J. (2015). Default and Executive Network Coupling Supports Creative Idea Production. *Scientific Reports, 5*, 1–14.

Björnsdotter, M., Löken, L., Olausson, H., Vallbo, k., & Wessberg, J. (2009). Somatotopic organization of gentle touch processing in the posterior insular cortex. *Journal of Neuroscience, 29*(29), 9314–9320.

Bréchet, L., Brunet, D., Birot, G., Gruetter, R., Michel, C. M., & Jorge, J. (2019). Capturing the spatiotemporal dynamics of self-generated, task-initiated thoughts with EEG and fMRI. *NeuroImage, 194*(March), 82–92.

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review, 114*(2), 340–375.

Cai, W., Chen, T., Ryali, S., Kochalka, J., Li, C. S. R., & Menon, V. (2016). Causal Interactions Within a Frontal-Cingulate-Parietal Network During Cognitive Control: Convergent Evidence from a Multisite-Multitask Investigation. *Cerebral Cortex, 26*(5), 2140–2153.

Calhoun, V. D., Liu, J., & Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage, 45*(1), S163–S172.

Chand, G. B., & Dhamala, M. (2016). Interactions among the Brain Default-Mode, Salience, and Central-Executive Networks during Perceptual Decision-Making of Moving Dots. *Brain Connectivity, 6*(3), 249–254.

Chand, G. B., Wu, J., Hajjar, I., & Qiu, D. (2017). Interactions of the Salience Network and Its Subsystems with the Default-Mode and the Central-Executive Networks in Normal Aging and Mild Cognitive Impairment. *Brain Connectivity, 7*(7), 401–412.

Chen, A. C., Oathes, D. J., Chang, C., Bradley, T., Zhou, Z.-W., Williams, L. M., Glover, G. H., Deisseroth, K., & Etkin, A. (2013). Causal interactions between fronto-parietal central executive and default-mode networks in humans. *Proceedings of the National Academy of Sciences, 110*(49), 19944–19949.

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews, 26*(2), 105–125.

Crottaz-Herbette, S., & Menon, V. (2006). Where and When the Anterior Cingulate Cortex Modulates Attentional Response: Combined fMRI and ERP Evidence. *Journal of Cognitive Neuroscience, 18*(5), 766–780.

Daniels, J. K., McFarlane, A. C., Bluhm, R. L., Moores, K. A., Clark, C. R., Shaw, M. E., … Lanius, R. A. (2010). Switching between executive and default mode networks in posttraumatic stress disorder: alterations in functional connectivity. *Journal of Psychiatry & Neuroscience: JPN, 35*(4), 258.

Daselaar, S. M., Rice, H. J., Greenberg, D. L., Cabeza, R., LaBar, K. S., & Rubin, D. C. (2008). The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall, emotional intensity, and reliving. *Cerebral Cortex, 18*(1), 217–229.

Davidovic, M., Starck, G., and Olausson, H. (2019). Processing of affective and emotionally neutral tactile stimuli in the insular cortex. Developmental Cognitive Neuroscience, 35(December 2017), 94–103.
Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences, 12*(3), 99–105.
Duggento, A., Passamonti, L., Valenza, G., Barbieri, R., Guerrisi, M., & Toschi, N. (2018). Multivariate Granger causality unveils directed parietal to prefrontal cortex connectivity during task-free MRI. *Scientific Reports, 8*(1), 1–11.
Ebisch, S. J., Gallese, V., Willems, R. M., Mantini, D., Groen, W. B., Romani, G. L., Buitelaar, J. K., & Bekkering, H. (2011). Altered intrinsic functional connectivity of anterior and posterior insula regions in high-functioning participants with autism spectrum disorder. *Human Brain Mapping, 32*(7), 1013–1028.
Erhardt, E. B., Rachakonda, S., Bedrick, E. J., Allen, E. A., Adali, T., & Calhoun, V. D. (2011). Comparison of multi-subject ICA methods for analysis of fMRI data. *Human Brain Mapping, 32*(12), 2075–2095.
Fegen, D., Buchsbaum, B. R., & D'Esposito, M. (2015). The effect of rehearsal rate and memory load on verbal working memory. *NeuroImage, 105*, 120–131.
Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron, 88*(1), 220–235.
Friston, K. J., Bastos, A. M., Oswal, A., van Wijk, B., Richter, C., & Litvak, V. (2014). Granger causality revisited. *NeuroImage, 101*, 796–808.
Gerlach, K. D., Spreng, R. N., Madore, K. P., & Schacter, D. L. (2013). Future planning: Default network activity couples with frontoparietal control network and reward-processing regions during process and outcome simulations. *Social Cognitive and Affective Neuroscience, 9*(12), 1942–1951.
Goulden, N., Khusnulina, A., Davis, N. J., Bracewell, R. M., Bokde, A. L., McNulty, J. P., & Mullins, P. G. (2014). The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. *NeuroImage, 99*(June), 180–190.
Hamilton, J. P., Furman, D. J., Chang, C., Thomason, M. E., Dennis, E., & Gotlib, I. H. (2011). Default-mode and task-positive network activity in major depressive disorder: Implications for adaptive and maladaptive rumination. *Biological Psychiatry, 70*(4), 327–333.
Iwabuchi, S. J., Raschke, F., Auer, D. P., Liddle, P. F., Lankappa, S. T., & Palaniyappan, L. (2017). Targeted transcranial theta-burst stimulation alters fronto-insular network and prefrontal GABA. *NeuroImage, 146*, 395–403.
Kim, H. (2012). A dual-subsystem model of the brain's default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. *NeuroImage, 61*(4), 966–977.
Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., & McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology, 6*(1), 27313.
Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: Distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *Journal of Neuroscience, 31*(9), 3217–3224.
Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences, 15*(10), 483–506.
Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, 1–13.
Moran, L. V., Tagamets, M. A., Sampath, H., O'Donnell, A., Stein, E. A., Kochunov, P., & Hong, L. E. (2013). Disruption of anterior insula modulation of large-scale brain networks in schizophrenia. *Biological Psychiatry, 74*(6), 467–474.
Nicholson, A.A., Harricharan, S., Densmore, M., Neufeld, R.W., Ros, T., McKinnon, M.C., Frewen, P.A., Théberge, J., Jetly, R., Pedlar, D., and Lanius, R.A. (2020). Classifying Heterogeneous Presentations of PTSD via the Default Mode, Central Executive, and Salience Networks with Machine Learning. NeuroImage: Clinical, Submitted, 102262.
Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25*(1), 46–59.
Pedersen, M., Zalesky, A., Omidvarnia, A., and Jackson, G.D. (2018). Multilayer network switching rate predicts brain performance. Proceedings of the National Academy of Sciences of the United States of America, 115(52), 13376–13381.
Rabellino, D., Tursich, M., Frewen, P. A., Daniels, J. K., Densmore, M., Théberge, J., & Lanius, R. A. (2015). Intrinsic Connectivity Networks in post-traumatic stress disorder during sub- and supraliminal processing of threat-related stimuli. *Acta Psychiatrica Scandinavica, 132*(5), 365–378.
Ramsey, J., Hanson, S., Hanson, C., Halchenko, Y., Poldrack, R., & Glymour, C. (2010). Six problems for causal inference from fMRI. *NeuroImage, 49*(2), 1545–1558.
Ryali, S., Supekar, K., Chen, T., Kochalka, J., Cai, W., Nicholas, J., Padmanabhan, A., & Menon, V. (2016). Temporal Dynamics and Developmental Maturation of Salience, Default and Central-Executive Network Interactions Revealed by Variational Bayes Hidden Markov Modeling. *PLoS Computational Biology, 12*(12), 1–29.
Schippers, M. B., & Keysers, C. (2011). Mapping the flow of information within the putative mirror neuron system during gesture observation. *NeuroImage, 57*(1), 37–44.
Sestieri, C., Corbetta, M., Romani, G. L., & Shulman, G. L. (2011). Episodic Memory Retrieval, Parietal Cortex, and the Default Mode Network: Functional and Topographic Analyses. *Journal of Neuroscience, 31*(12), 4407–4420.
Seth, A. K., Barrett, A. B., & Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *Journal of Neuroscience, 35*(8), 3293–3297.
Shaw, S. B., Dhindsa, K., Reilly, J. P., & Becker, S. (2019). Capturing the Forest but Missing the Trees: Microstates Inadequate for Characterizing Shorter-Scale EEG Dynamics. *Neural computation, 31*(11), 2177–2211.
Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex, 22*(1), 158–165.
Sidlauskaite, J., Wiersema, J. R., Roeyers, H., Krebs, R. M., Vassena, E., Fias, W., Brass, M., Achten, E., & Sonuga-Barke, E. (2014). Anticipatory processes in brain state switching - Evidence from a novel cued-switching task implicating default mode and salience networks. *NeuroImage, 98*, 359–365.
Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*(1), 5–42.
Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage, 53*(1), 303–317.
Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences, 105*(34), 12569–12574.
St. Jacques, P. L. De Brigard, F. (2015). Neural Correlates of Autobiographical Memory. In The Wiley Handbook on the Cognitive Neuroscience of Memory, pages 265–286. John Wiley & Sons, Ltd.
Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia, 44*(12), 2189–2208.
Sweet, L. H., Paskavitz, J. F., Haley, A. P., Gunstad, J. J., Mulligan, R. C., Nyalakanti, P. K., & Cohen, R. A. (2008). Imaging phonological similarity effects on verbal working memory. *Neuropsychologia, 46*(4), 1114–1123.
Taghia, J., Ryali, S., Chen, T., Supekar, K., Cai, W., & Menon, V. (2017). Bayesian switching factor analysis for estimating time-varying functional connectivity in fMRI. *NeuroImage, 155*(February), 271–290.
Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(3), 1125–1165.
Thompson, W. H., Thelin, E. P., Lilja, A., Bellander, B. M., & Fransson, P. (2015). Functional resting-state fMRI connectivity correlates with serum levels of the S100B protein in the acute phase of traumatic brain injury. *NeuroImage: Clinical, 12*, 1004–1012.
Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience, 16*(1), 55–61.
Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). The role of working memory capacity in autobiographical retrieval: Individual differences in strategic search. *Memory, 20*(2), 167–176.
Vallar, G., & Pagano, C. (2002). *In Handbook of memory Disorders (eds: Baddeley AD.* Wilson BA: Kopelman MD.
van den Heuvel, M. P., Mandl, R. C. W., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2010). Aberrant Frontal and Temporal Complex Network Structure in Schizophrenia: A Graph Theoretical Analysis. *Journal of Neuroscience, 30*(47), 15915–15926.
Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012a). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity, 2*(3), 125–141.
Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012b). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*.
Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., Reynolds, G. O., Hofmann, S. G., Pollack, M. H., & Gabrieli, J. D. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry, 21*(5), 680–685.
Wittmann, M., Simmons, A. N., Aron, J. L., & Paulus, M. P. (2010). Accumulation of neural activity in the posterior insula encodes the passage of time. *Neuropsychologia, 48*(10), 3110–3120.
Yang, Y., Zhong, N., Imamura, K., Lu, S., Li, M., Zhou, H., Li, H., Yang, X., Wan, Z., Wang, G., Hu, B., & Li, K. (2016). Task and resting-state fMRI reveal altered salience responses to positive stimuli in patients with major depressive disorder. *PLoS ONE, 11*(5), 1–19.
Yaple, Z. A., Stevens, W. D., & Arsalidou, M. (2019). Meta-analyses of the n-back working memory task: fMRI evidence of age-related changes in prefrontal cortex involvement across the adult lifespan. *NeuroImage, 196*(March), 16–31.
Zhang, Q., Zhang, G., Yao, L., & Zhao, X. (2015). Impact of real-time fMRI working memory feedback training on the interactions between three core brain networks. *Frontiers in Behavioral Neuroscience, 9*(september), 1–9.

5.3 Discussion

The results of this study provided the first direct empirical validation of the SN-based gating of the DMN and CEN during interoceptive and exteroceptive processing in a cohort of healthy participants. While both the DMN and CEN were active during both tasks, the SN was found to co-activate with the task-relevant network. This finding indicated that the SN's gating behaviour might be actualized by inter-network coherence. While such co-activation of the SN and CEN was previously observed during working memory and executive tasks (Dosenbach et al., 2008), this study shows that SN co-activation generalizes to other ICNs in a task-relevant manner.

This mechanistic insight provided further evidence of network co-activation being an important feature of ICN dynamics, further invalidating a major assumption of EEG microstate-based analyses, namely, that only one microstate is active at a time, as discussed in chapter 4.

Critically, this study characterized healthy patterns of fMRI-derived ICN activity during interoceptive and exteroceptive processing, an important step that could inform the development of an EEG-based ICN analysis tool, as described in the next chapter. Understanding such healthy network dynamics is also necessary for understanding the trajectory of network changes representative of an improvement or deterioration in network dynamics. This is especially relevant for assessing the impacts of treatments and interventions on ICNs, as shown in chapter 7, and could lead to improved objective treatment tracking in clinical settings.

Bibliography

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3):99–105.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, 15(10):483–506.

Nicholson, A. A., Harricharan, S., Densmore, M., Neufeld, R. W., Ros, T., McKinnon, M. C., Frewen, P. A., Théberge, J., Jetly, R., Pedlar, D., and Lanius, R. A. (2020). Classifying Heterogeneous Presentations of PTSD via the Default Mode, Central Executive, and Salience Networks with Machine Learning. *NeuroImage: Clinical*, Submitted:102262.

Chapter 6

EEG-based tracking of ICNs

6.1 Introduction

The need for an alternate EEG-based method for assessing ICN activity and dynamics is apparent from the scrutiny of EEG microstates in chapter 4. Ideally, such a method would leverage the pattern of healthy network dynamics identified in chapter 5 to detect any deviations indicative of network dysfunctions in the context of the tri-network model. This approach could predict cognitive and behavioural deficits due to changes in ICN dynamics involved in interoceptive and exteroceptive processing. The ability to monitor the pattern of activity between these networks can also provide a better network-level understanding of the neural mechanisms underlying the mood and cognitive impairments observed in psychological disorders and characterize the network-level changes caused by current clinical therapies such as exposure therapy and cognitive behavioural therapy, known to improve mood and cognitive processing in such psychological disorders (Lanius et al., 2015). Furthermore, the use of a cheap and accessible neuroimaging modality, such as EEG, to monitor the activity of these

three networks will greatly increase the clinical accessibility of brain network-based biomarkers in the treatment planning and tracking of psychopathologies. Additionally, such a system could lead to the development of more promising interventions such as network-based neurofeedback training (Sitaram et al., 2016).

While others have attempted to develop such a system, most other approaches overlook the spatio-temporal hallmarks of each ICN that could be predictive of its activation. This chapter discusses a unique signal processing and machine learning pipeline that uses simultaneous EEG-fMRI data to first identify such spatio-temporal signatures of the three ICNs within the tri-network model (CEN, DMN, SN), followed by training a classifier to predict ICN activation based on these spatio-temporal signatures of each ICN.

The primary author (SBS) of the study conceptualized and designed this study, along with the last author (SB). SBS and SB also developed the functional task used in cohort 1 of this study. SBS also performed all data collection of the 1st cohort of participants. AHH and JFC designed the functional task used in cohort 2 of this study, and collected data from all participants within this cohort. SBS and SB co-developed the novel analysis pipeline discussed in this paper and wrote the first draft of the manuscript. All co-authors contributed to the writing of the manuscript. MM, JH and SB contributed funding to this study.

6.2 Tracking the Brain's Intrinsic Connectivity Networks in EEG

Shaw, S. B., McKinnon, M. C., Heisz, J., Harrison, A. H., Connolly, J. F. & Becker, S. (2021). Tracking the Brain's Intrinsic Connectivity Networks in EEG. bioRxiv Available online at: https://doi.org/10.1101/2021.06.18.449078

Tracking the Brain's Intrinsic Connectivity Networks in EEG

Saurabh Bhaskar Shaw[a,c,d], Margaret C. McKinnon[e,f,g], Jennifer J. Heisz[h], Amabilis H. Harrison[i,j], John F. Connolly[a,b,c,d,k], Suzanna Becker[a,b,c,d]

[a]*Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada*
[b]*Department of Psychology Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada*
[c]*Vector Institute for Artificial Intelligence, Toronto*
[d]*Centre for Advanced Research in Experimental and Applied Linguistics (ARiEAL), Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada*
[e]*Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada*
[f]*Mood Disorders Program, St. Joseph's Healthcare, Hamilton, ON, Canada*
[g]*Homewood Research Institute, Guelph, ON, Canada*
[h]*Department of Kinesiology, McMaster University, Hamilton, ON, Canada*
[i]*Neuroscience Program, Hamilton Health Sciences, Hamilton ON*
[j]*Imaging Research Centre, St. Joseph's Heathcare Hamilton, Hamilton ON*
[k]*Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada*

Abstract

Functional magnetic resonance imaging (fMRI) has identified dysfunctional network dynamics underlying a number of psychopathologies, including post-traumatic stress disorder, depression and schizophrenia. There is tremendous potential for the development of network-based clinical biomarkers to better characterize these disorders. However, to realize this potential requires the ability to track brain networks using a more affordable imaging modality, such as Electroencephalography (EEG). Here we present a novel analysis pipeline capable of tracking brain networks from EEG alone, after training on supervisory signals derived from data simultaneously recorded in EEG and fMRI, while people engaged in various cognitive tasks. EEG-based features were then used to classify three cognitively-relevant brain networks with up to 75% accuracy. These findings could lead to affordable and non-invasive methods to objectively diagnose brain disorders involving dysfunctional network dynamics, and to track and even predict treatment responses.

Keywords: Intrinsic Connectivity Networks (ICN), Default Mode Network (DMN), Central Executive Network (CEN), Salience Network (SN), Simultaneous EEG-fMRI, Machine Learning

1. Introduction

A large body of neuroimaging research over the past decade indicates that the brain is organized into functional networks of interacting brain regions, called intrinsic connectivity networks (ICNs). The study of large-scale ICNs has provided considerable insight into the neural basis of human cognition and behaviour in the healthy and diseased brain [52, 23]. There is immense potential to use features of ICN dynamics as clinical biomarkers in patients with various psychopathologies [23] and even track their response to treatments. However, the predominant imaging modality used to study ICNs, functional magnetic resonance imaging (fMRI), is expensive, lacks the necessary temporal resolution, and is not readily available for performing routine neurocognitive assessments in patients with brain disorders. A major breakthrough would be to track ICNs using a non-invasive and more widely accessible modality such as electroencephalography (EEG). Here we present a significant advancement towards this goal. We apply machine learning methods to simultaneously recorded EEG and fMRI data, to derive supervisory signals for learning EEG-based ICN features, permitting highly accurate classification and tracking of ICNs from EEG alone.

Three predominant networks have been extensively studied [23] using fMRI: the central executive network (CEN) generally involved in exteroceptive processing, *i.e.* tasks involving attention to external stimuli; the default mode network (DMN) involved in interoceptive processing tasks, *e.g.* autobiographical memory retrieval, imagining the future, spatial planning and nav-

igation, and self-reflection; and the salience network (SN), thought to modulate switching between CEN-mediated exteroceptive and DMN-mediated interoceptive cognitive processes [44]. Appropriate network switching dynamics between these three core networks is thought to be critical for healthy cognitive functioning. Disruptions in normal inter- and intra-network connectivity in these ICNs have been observed in numerous neuropsychological conditions affecting emotion and cognition. For instance, those with Major Depressive Disorder (MDD) [7] exhibit deficits in down-regulating activity within the DMN in association with persistent rumination [29]. Those with Post-Traumatic Stress Disorder (PTSD) exhibit disruptions in the activation and functional connectivity of DMN, CEN and SN [2, 34, 20, 48]. For example, abnormal activation of the DMN has been observed in PTSD patients while switching to a working memory task that normally recruits the CEN [8]. Disrupted ICN dynamics are also observed in numerous other psychological disorders [23], such as bipolar disorder [56, 51], schizophrenia [57] and mild cognitive impairment (MCI) [55]. Therefore, it is extremely promising to track the network dynamics within these ICNs as clinical markers of brain disorders. The ability to monitor such network activity will be particularly useful for tracking progress in the treatment of such neuropathologies, where network dynamics is dysregulated, and may also lead to the development of novel individualized treatments such as network-based neurofeedback interventions.

Unfortunately, doing so using fMRI would be prohibitively expensive. Moreover, fMRI lacks the temporal resolution to track the temporal dynamics of the networks on a millisecond timescale. An appealing alternative is electroencephalography (EEG), a cheaper and more widely accessible imaging modality with excellent temporal resolution.

In previous work some EEG features of these brain networks have been identified. For instance, working memory load on the CEN is found to modulate fronto-parietal EEG power in theta and upper alpha frequency bands [37], fronto-parietal phase-based functional connectivity graphs [10], and common spatial patterns [5]. Frontal lobe EEG activity in the theta frequency band is also found to negatively correlate with the DMN [39], and when combined with delta and alpha band powers, is capable of discriminating the DMN from the sensorimotor network [41]. Furthermore, theta-gamma coupling is a key mechanism driving hippocampal memory processes required by the DMN during autobiographical memory retrieval and is found to be dysfunctional in patients with memory impairments [28]. Hence, a combination of within-frequency and cross-frequency phase-based and amplitude-based connectivity measures could capture various component neural processes inherent to the CEN, DMN and SN. An open question is whether unique signatures of each of these networks can be identified, allowing us to track each of the three networks as distinct from the other two.

Rather than focusing on such temporal dynamics, most previous attempts to identify EEG features of ICNs have relied on spatial filtering analyses such as beamforming and blind source separation [46]. These analyses are especially susceptible to signal leakage due to volume conduction [32], and specific data acquisition parameters [21], limiting their utility as clinical tools for tracking ICN dynamics. One alternate approach is to track EEG microstates [17], which are spatial correlates of ICNs identified by spatial clustering. However, despite its increasing popularity in probing dysfunctional ICN dynamics in numerous psychological conditions [26], this analysis is riddled with flawed assumptions that lead to inaccuracies at finer temporal scales [43, 27].

To date no one has identified unique signatures that permit classification and tracking of ICNs using EEG alone. To accomplish this, we developed a machine learning model that learns EEG signatures of the ICNs using supervisory labels derived from simultaneously recorded fMRI. The simultaneous EEG-fMRI data were collected from two cohorts of participants that performed two different multi-task paradigms: a dual task-switching paradigm designed to activate the DMN, CEN and SN [44]; and a multi-task paradigm that cycles through a series of seven tasks, a subset of which rely on the three networks of interest [16]. A large battery of approximately 40M amplitude and phase-based features were computed from the EEG data collected during these tasks. Classification labels for the DMN, CEN and SN were derived from the thresholded activity of these ICNs, identified using group-wide ICA analysis of the simultaneously acquired fMRI data. These labels were used to select the optimal feature set for a multiclass SVM classifier, using a hierarchical minumum-redundancy-maximum-relevance (mRMR) [33] feature selection algorithm. Importantly, having identified these features from the simultaneous EEG-fMRI data, there is the potential to track the ICNs using EEG data alone. The ability of the EEG features alone to classify the predominantly active network was validated on testing sets from the above described datasets using 20-fold cross-validation.

2. Results

2.1. FMRI Networks as labels

The first step in our analysis pipeline (see Figure 1C) was to identify the "ground-truth" ICN activity, *i.e.* the correct network labels, as identified using fMRI. This involved a group-wide independent component analysis (group-ICA) to discover components that overlap with the ICNs of interest.

The group-ICA analysis identified a total of 20 components with distinct spatio-temporal patterns of activity. Of these, 11 components showed significant spatial overlap (high Dice coefficient) with well known ICNs [45], while the remaining 9 components represented unwanted noise and artefacts. We focused on the 6 components representing CEN, DMN and SN subnetworks (Figure 1C.iii), and extracted the time-courses of overall ICN activity by averaging the activity of each ICNs' component sub-networks (Figure 1C.v). The temporal dynamics of these components matched those expected from the CEN, DMN and SN within the dynamic task-switching paradigm used in cohort A, with the SN causally influencing the CEN and DMN in a task-linked manner. The fMRI temporal dynamics of the ICNs lie outside the scope of this paper and are further explored in Shaw *et al.* [44].

The component time courses for the CEN, DMN and SN were partitioned into overlapping windows of 5 fMRI time points, using a sliding-window approach and advancing by 1 fMRI time point for each window. Each time window was labeled with the most active ICN during the corresponding time period, creating the class labels for feature selection and classification of the EEG data to predict the predominantly active ICN during each time window.

2.2. Learning EEG features from fMRI-derived labels

The next step in our processing pipeline (see figure 1C) involved using the fMRI labels identified in section 2.1 to select relevant EEG features.

2.2.1. Generalized EEG features

A generalized feature set was identified using mRMR on the features extracted from the EEG data, using labels derived from the simultaneously acquired fMRI data. Here mRMR tries to find EEG features that are maximally relevant to the fMRI-derived class labels across all participants, while being minimally redundant.

2.2.2. Individualized EEG features

An individualized EEG feature set was also identified using a similar procedure for each individual participant. Here mRMR tries to find separate EEG features for each participant that are maximally relevant to the fMRI-derived class labels for that participant, while being minimally redundant.

Having obtained these features, they can now be applied to EEG data from other datasets. We now investigate the ability of these features to predict ICN activity. Researchers can then utilize the trained classifiers with the corresponding EEG features dervied from their own datasets to assess ICN activity using EEG data only.

2.3. ICN activation can be identified using EEG features alone

Using the optimal feature sets identified in the previous section, we trained three versions of the classifier to predict CEN, DMN and SN activation, representing three scenarios of data availability while using this pipeline. The first two scenarios (generalized classifier and semi-supervised individualized classifier) represent a use case where the researcher has access to only EEG data from their participants. They can use our models trained on EEG-fMRI derived features to predict ICN activity using only EEG data from their participants. The third scenario (fully-supervised individualized classifier) represents a use case where the researcher has access to simultaneous EEG-fMRI data from their own participants. The observed classification performance for each of the three scenarios is detailed below.

2.3.1. Generalized classifer

The first classifier was trained on the generalized feature set from all but one participant. This classifier was then tested on the left-out participant's EEG data, comparing the predicted labels to the "ground-truth" labels derived from their corresponding fMRI data. These classifiers achieved an average classification accuracy of 58% ± 6% for cohort A and 61% ± 5% for cohort B, and performed significantly better than the corresponding surrogate null models ($p < 0.001$), trained using randomly permuted training labels.

2.3.2. Semi-supervised individualized classifier

We then explored a semi-supervised approach to boost the classification performance of the generalized classifier, by training a custom classifier for each participant's EEG data using the predicted class labels from section 2.3.1. To accomplish this, the time points with a confidently predicted label (maximum posterior

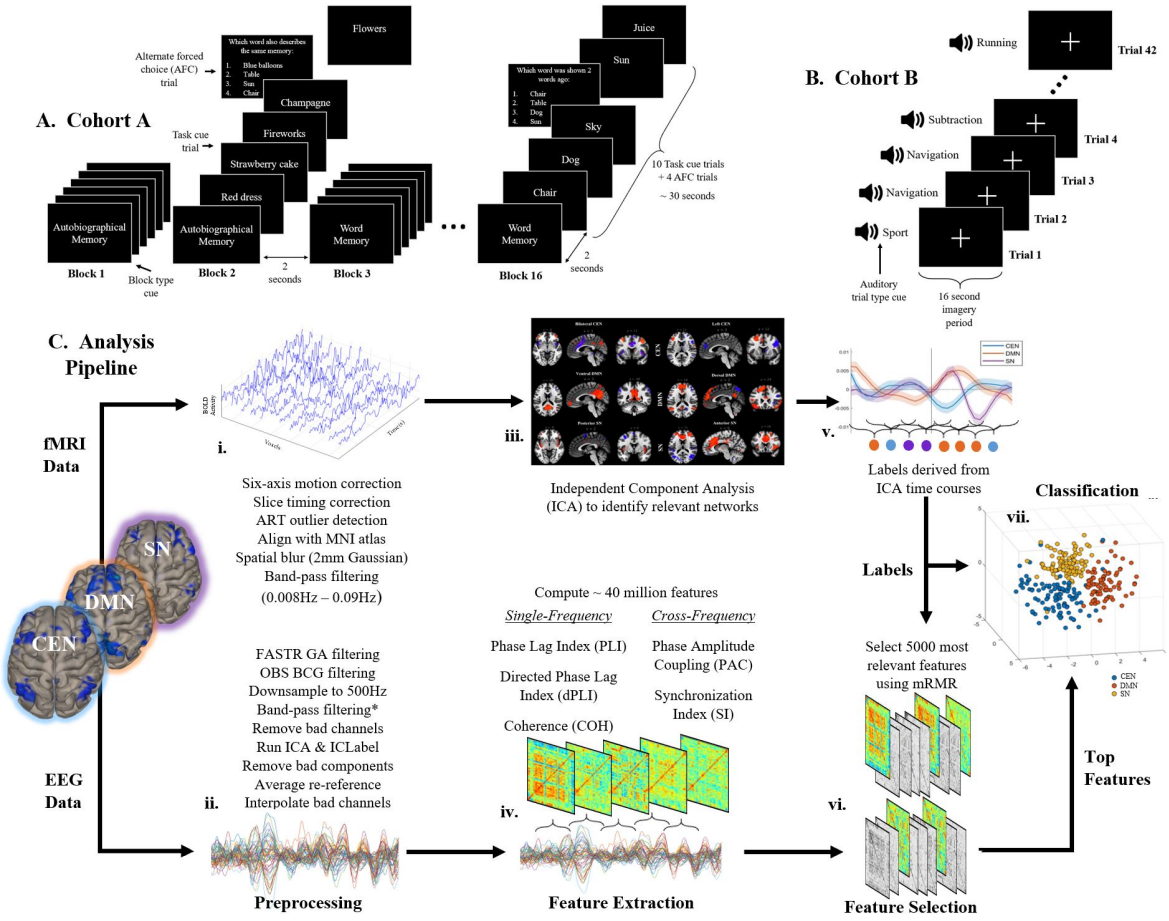

Figure 1: The simultaneous EEG-fMRI data used in this study were collected from two cohorts of participants that performed two different multi-task paradigms, shown in panels A and B. Cohort A used a dual task-switching paradigm (shown in panel A) designed to activate the DMN, CEN and SN, identical to the paradigm used in Shaw *et al.* [44]; and Cohort B performed a multi-task paradigm (shown in panel B) that cycled through a series of seven tasks [16], a subset of which relied on the three networks of interest. Refer to the methods section for a detailed description of the tasks. Panel C details the analysis pipeline used in this study to analyze the simultaneous EEG-fMRI data collected from Cohorts A and B. The top row illustrates the analysis steps for fMRI data (sub-panels i., iii., and v.), while that for EEG data is shown in the bottom row (sub-panels ii., iv., and vi.). *Top row (fMRI data analysis):* i. Preprocessing - The fMRI scans were first realigned and unwarped, followed by motion correction, slice timing correction (STC), ART-based outlier identification and scrubbing, normalization to the MNI152 atlas, spatial smoothing using a 2mm Gaussian kernel, and band-pass filtering between 0.008Hz−0.09Hz.; iii. The fMRI data was then decomposed into 20 independent components using ICA, of which 6 components were found to be relevant to the DMN, CEN and SN (shown here). v. Classification labels for the DMN, CEN and SN were derived from the thresholded activity of these three ICNs. *Bottom row(EEG data analysis):* ii. The concurrently collected EEG data was preprocessed by first removing the gradient artefact (GA) using a parallel optimized version of the FASTR gradient artefact removal toolbox Shaw [42], followed by ballistocardiogram (BCG) filtering using optimal basis set filtering (OBS). The artefact-free EEG data were then downsampled from 5000Hz to 500Hz, followed by temporal band-pass filtering into six different frequency bands - full (1-50Hz), delta (1-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), and gamma (30-50Hz). Bad channels were then detected and removed based on their spectral characteristics, followed by an ICA decomposition to identify and remove artefacutal components such as ocular artefacts, eye blinks and muscle artefacts. Finally, the EEG data were referenced to average EEG channel data, after which the previously removed bad channels were interpolated using spherical interpolation. iv. A large battery of approximately 40M amplitude and phase-based features were computed from the preprocessed EEG data. vi. The previously derived classification labels were used to select the optimal feature set for a multiclass SVM classifier, using a hierarchical minumum-redundancy-maximum-relevance (mRMR) [33] feature selection algorithm. vii. These features were used to classify the three ICNs. The clear separation of the three classes shown results in high classification accuracy.

probability > 75%) from the generalized classifier were picked as the labelled time points, while the rest of the time points were considered unlabelled. When comparing these EEG-only derived labels were then compared to the "ground-truth" labels from the participants' fMRI data. the confidently labeled time points boosted the average classification accuracy to 82% ± 6%, while accounting for only 25% of the participant's time points. Therefore, this process split each participants' data into a small labelled dataset with "expert" labels from the generalized classifier, and a larger unlabelled dataset.

Optimal features were selected using a weighted-mRMR approach that used labels predicted by the generalized classifier, and a weighted mutual information estimate for identifying minimally redundant and maximally relevant features. The weighted mutual information estimate weighted each time point by the maximum posterior probability of the predicted label, giving more importance to the time points with more confident predictions. A semi-supervised approach was then used to predict the labels of the unlabelled data, achieving an average accuracy of 57.9% ± 6% for cohort A and 65.5% ± 6% for cohort B.

2.3.3. *Fully-supervised individualized classifier*

Lastly, a fully-supervised individual classifier was trained on a subset of each participants' data, using the individualized feature set identified in section 2.2.2 with their fMRI-derived labels. This classifier was tested using a held-out subset of the same participants' data, achieving an average classification accuracy of 98% ± 3% for cohort A and 96% ± 4% for cohort B.

Interestingly, the expansion of the EEG signal from 64 channels to an extremely high dimensional feature space (40M), and its subsequent reduction to a 5000 dimensional space, made the classification task much easier by transforming the EEG feature space from an inseparable domain to a readily separable domain, shown in Figure1C.vii. This allowed a simple classifier, a multi-class support vector machine (SVM) to achieve the highest classification accuracy, substantially outperforming several much more complicated deep neural network classifiers (see supplementary figure). These deep neural network methods perform notoriously poorly in domains where the number of observations is not significantly larger than the number of features, where they are susceptible to overfitting.

2.4. *Characterizing the features of each network*

MRMR feature selection identified the top 5000 most relevant features for discriminating between the CEN, DMN and SN. To gain a more intuitive understanding of the EEG signatures uniquely representative of each ICN's activity, a Shapley additive explanations (SHAP) analysis [22] was conducted for each feature within the identified feature space for the generalized classifier. The SHAP analysis explains the contribution of each feature to the classification, by identifying the relative change in log-odds of each ICN label, given an increase in the corresponding feature value. The SHAP values of each feature for the correct ICN label, averaged across all trials, are shown in Figure 3 for the single-frequency features and in Figure 4 for the cross-frequency features. These are further discussed in the following sections.

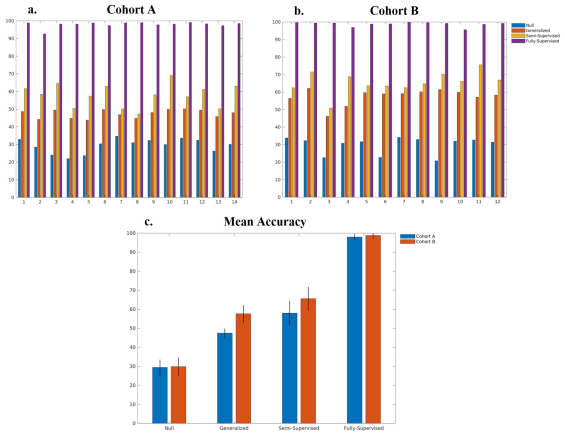

Figure 2: The test-set classification accuracies for predicting the activation of CEN, DMN and SN networks using EEG data alone across cohorts A and B, shown in (a.) and (b.) respectively. The accuracy of the null model is shown in blue, that of the generalized model is shown in orange, and that of the fully-supervised model is shown in yellow. while that of the surrogate null model, trained using random labels, is shown in red. The average classification accuracy for the three classifiers is shown in (c.), with cohort A shown in blue and cohort B shown in orange.

Five connectivity features were included in our feature set - three single frequency features, and two cross frequency features. The single frequency features identified the connectivity between pairs of EEG channels within the same frequency band, and included phase lag index (PLI), directed phase lag index (dPLI), and coherence (COH). The cross-frequency features describe the connectivity between pairs of EEG channels across different frequency bands, and include phase amplitude coupling (PAC) and synchronization index (SI).

2.5. *CEN Features*

All five connectivity features contributed towards the classification of the CEN network, with the PLI

Pt. Number	Cohort A				Pt. Number	Cohort B			
	Null Class.	Gen. Class.	Semi-sup. Ind. Class.	Fully-sup. Ind. Class.		Null Class.	Gen. Class.	Semi-sup. Ind. Class.	Fully-sup. Ind. Class.
1	32.9	48.7	61.6	98.9	1	33.8	56.4	62.4	99.7
2	28.5	44.2	58.4	92.6	2	32.3	62.1	71.5	99.4
3	24.0	49.5	64.6	98.2	3	22.6	46.2	50.8	99.4
4	22.0	44.8	50.3	98.1	4	30.8	51.9	68.8	96.9
5	23.7	43.8	57.4	98.8	5	31.7	59.7	63.6	98.8
6	30.4	49.8	62.9	97.3	6	22.7	59.0	63.5	98.9
7	34.7	46.9	50.1	98.9	7	34.2	59.1	62.5	99.8
8	31.0	44.8	47.2	99.0	8	32.9	60.2	64.7	99.6
9	32.3	48.1	58.1	97.8	9	20.8	61.4	70.1	99.2
10	30.0	49.9	69.1	98.1	10	32.0	59.9	66.1	95.6
11	33.6	50.2	57.0	99.1	11	32.7	57.1	75.6	98.7
12	32.4	49.5	61.2	98.3	12	31.4	58.3	66.9	99.2
13	26.3	45.8	50.3	97.3	-	-	-	-	-
14	30.1	48.0	63.0	98.5	-	-	-	-	-

Table 1: Three-way classification accuracy (test set) for predicting the activation of the CEN, DMN and SN across cohort A and B. The reported values are the cross-validated (20-fold) mean accuracy of the multiclass-SVM model trained using a 75-25 train-test split of the data. All classification accuracies of the trained model are significantly higher than that of the surrogate null model, trained using random labels.

and dPLI being the highest contributors (Figure 3a). Among the single-frequency features, high connectivity between fronto-temporal, intra-parietal and intra-frontal electrodes (Figure 3g) across theta, alpha, beta and delta bands (Figure 3d) were found to be indicative of CEN activity.

Interestingly, the features most strongly influencing the CEN classification showed two unique lateralization patterns (Figure 3g). The fronto-temporal and intra-parietal connectivity features were left-lateralized, alongside bilateral and cross-hemispheric fronto-temporal and intra-frontal connectivity features. These two lateralization patterns are consistent with the bilateral and left-lateralized sub-networks of the CEN, seen with the fMRI results in Figure 1C.iii.

A wide range of cross-frequency coupling (CFC) features were also found to influence CEN classification, as shown in Figure 4a. Among these, the connectivity features between theta band and the full frequency band were particularly predictive of CEN activity (Figure 4d). Furthermore, the channel-pairs communicating within these frequency bands included inter-hemispheric frontal-frontal and fronto-temporal channels pairs, along with intra-hemispheric fronto-parietal and fronto-occipital channels. The direction of the fronto-temporal connections followed a theta to full-band direction, with the phase of the theta frequency at the frontal electrodes synchronizing with the full frequency band activity at the temporal electrodes. This pattern is consistent with frontal theta-driven processes during executive tasks such as working memory [38] and mental arithmetic [36].

2.6. DMN Features

Similar to the pattern seen in CEN classification, all five connectivity features were informative in predicting DMN activity, with dPLI, COH and SI being more informative than PLI and PAC. However, in contrast to the pattern seen in CEN classification, the most informative single-frequency features included gamma, theta and alpha band activity within parietal-parietal channel pairs. Furthermore, the intra-parietal connectivity within the single-frequency features was restricted to ipsi-lateral channel pairs with minimal inter-hemispheric connections. Such parietal-driven gamma activity could be indicative of self-related processing occurring within key parietal DMN nodes such as the posterior cingulate cortex (PCC) and the Precuneus.

Additionally, fronto-parietal theta-gamma coupling features were found to be the most informative CFC feature, following a fronto-parietal direction, with the theta phase of frontal electrode activity synchronizing with the gamma activity of the parietal electrodes. These included bilateral cross-hemispheric fronto-parietal con-

nections, and a right-lateralized fronto-parietal connection, with the latter being most informative of DMN activity. These data are consistent with a large body of research implicating theta-gamma coupling in the various DMN-linked memory processes [12, 35, 31].

2.7. SN Features

In contrast to the pattern seen for CEN and DMN classification, the cross-frequency coupling features (SI and PAC) dominated the prediction of SN activity. The single-frequency features that contributed to SN classification included a wide range of fronto-temporal, temporal-temporal, occipito-frontal, parietal-parietal and parieto-occipital connections. Furthermore, these connections were within the beta, alpha, delta and full frequency band, with no single-frequency features within the gamma and theta bands.

Among the CFC features, coupling between alpha-delta, gamma-full band and alpha-full band was found to be most predictive of SN activity, across a wide combination of channel-pairs. This is consistent with the integrative role of the SN that involves communication with numerous brain regions across a wide range of frequency bands.

One notable observation is the concentration of gamma and delta activity at the right frontal electrodes, that might be linked to the activity of the right anterior insula (rAI) node of the SN. This node is particularly important for task-linked switching between the CEN and DMN, and is well connected with other frontal, parietal and temporal brain regions [47, 25, 24].

Taken together, the identified EEG feature set captures critical spatio-temporal characteristics of the CEN, DMN and SN, that is consistent with their functional roles and previously observed dynamics.

3. Discussion

In this study, we used cutting-edge machine learning methods to classify and track the activity of three major ICNs using EEG data, which was previously only possible using fMRI data. We focused on three core ICNs within Menon's tri-network model [23] - the CEN, DMN and SN, given their relevance in characterizing a wide range of psychopathologies. FMRI studies indicate that these intrinsic networks are dysregulated in psychopathologies including PTSD, and can even be used to predict patients' PTSD subtype [30]. Moreover, characterizing ICN dynamics has the potential to track response to treatments and inform individualized treatment planning decisions. However, despite its potential as a clinical assessment tool, probing ICN dynamics using fMRI is prohibitively expensive. Other barriers to routine clinicial use of fMRI in psychiatric disorders include the distress caused by the confined enclosure and loud sounds made by the scanner, both of which can be triggering for those with PTSD. EEG, by comparison, is much more affordable, relatively portable, and does not carry with it the risk of claustrophobia- and noise-related distress. Therefore, the potential to monitor brain networks using EEG, as afforded by the present study, can greatly improve the clinical accessibility of ICN-based assessment.

To accomplish this, we used fMRI-derived ICN labels to select relevant features from simultaneously-acquired EEG data and classified the active ICNs using EEG data alone. Three distinct classification scenarios were explored, providing investigators with three levels of performance, depending on the type of data available to them. The first scenario involves applying the generalized model trained on our dataset to EEG data from other participants to predict the active ICN. The second scenario extends this approach by additionally applying semi-supervised learning to train a custom classifier for each participant, improving classification accuracy by 5% to 15%. These two approaches represent the most significant contribution of this work, allowing investigators to probe ICN activity using the proposed methods with EEG data from their participants. Lastly, the third scenario demonstrates that higher classification accuracy is achievable by fully-supervised learning, if the investigator has access to simultaneously acquired EEG-fMRI data from their participants.

It is important to note that the performance of the classifiers in the three scenarios was not significantly different between the two cohorts (p=0.6), demonstrating that our methods can be used to discriminate between the activity of the CEN, DMN and SN, across a wide range of cognitive tasks.

This study also investigated the EEG features contributing to the detection of each ICN, identifying EEG signatures that uniquely characterize the CEN, DMN and SN. Interestingly, the data-driven approach used in this study identified EEG signatures that aligned with major findings in the literature.

Frontal theta band activity, found to be predictive of both CEN and DMN activation in this study (Figure 4), has been implicated in both executive processes (CEN functioning) and autobiographical memory encoding and retrieval processes (DMN functioning). In the context of executive tasks, frontal theta is thought to synchronize the pre-frontal cortex with a wide range of other brain regions [14], signalling the need for cogni-

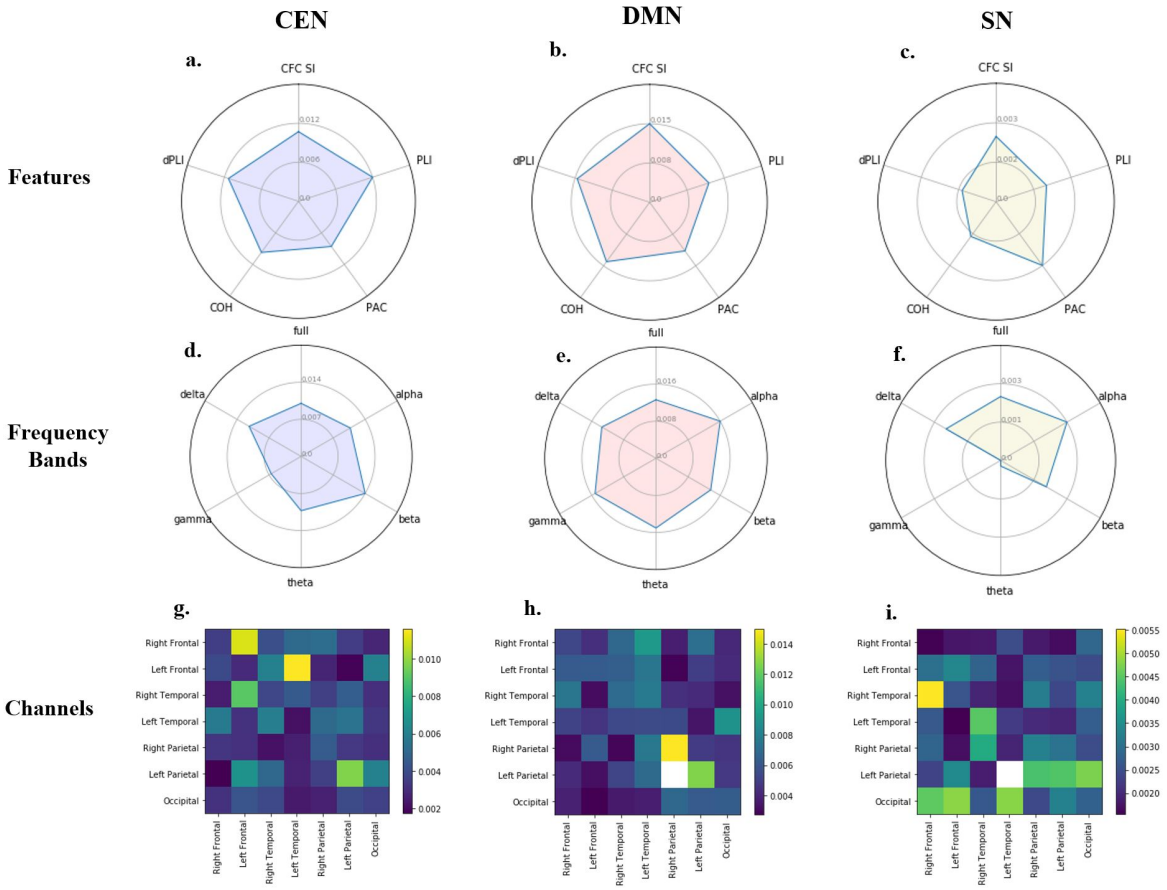

Figure 3: The influence of the identified feature set in increasing the log-odds of each network class (CEN, DMN and SN). The contributions of the five features across different frequency bands and channel pairs are shown in the top, middle and bottom rows respectively.

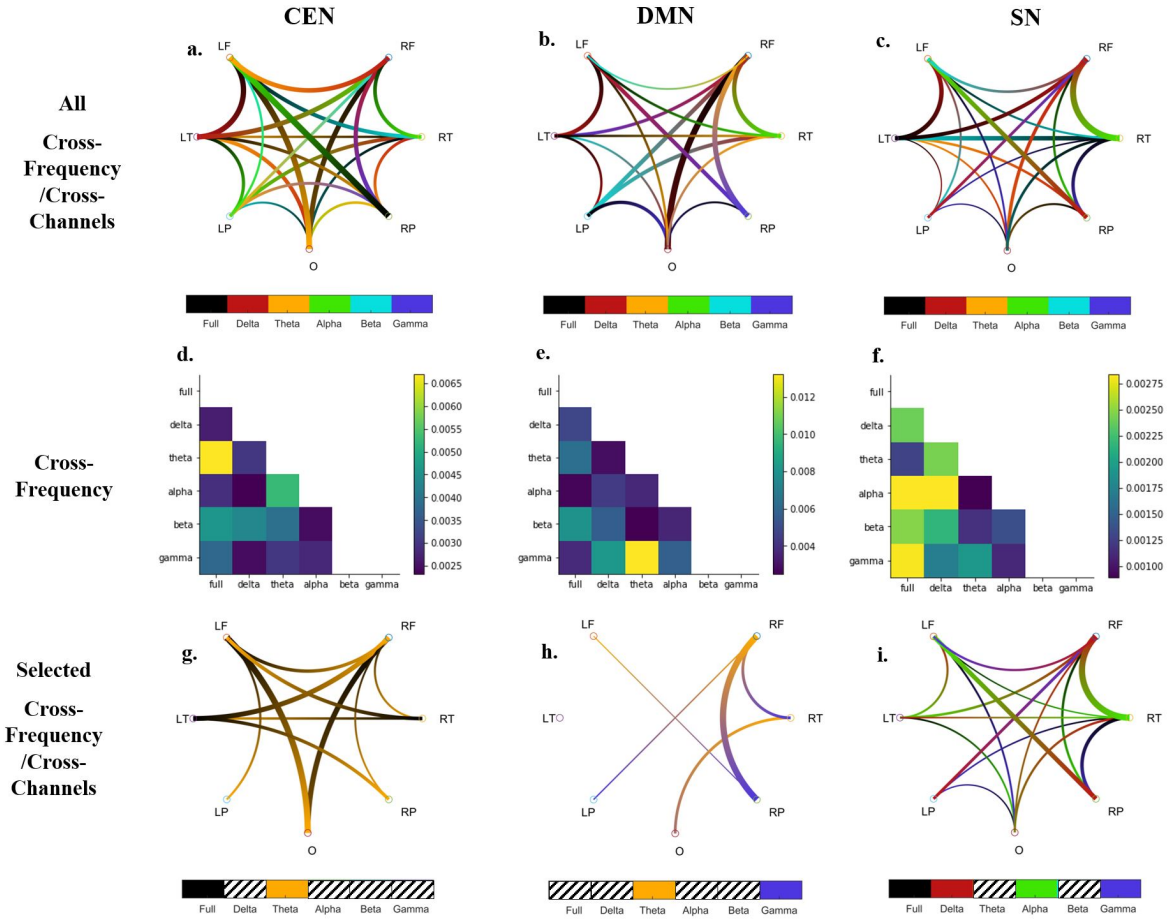

Figure 4: The influence of the identified feature set in increasing the log-odds of each network class (CEN, DMN and SN), shown for the cross-frequency features (SI and PAC).

tive control processes during periods of high risk and/or high levels of uncertainty [4]. Doing so in a phase-linked manner, it acts as an alarm signal for instantiating cognitive control processes to better learn from the higher error rates encountered during such scenarios. Adaptations in theta band dynamics also assist with optimizing this process, as observed with the reduction in peak theta frequency during higher cognitive control loads. Simulated models show the slower theta frequency increases the probability of success in more difficult scenarios [40], better adapting to the task demands. Given the diverse roles of theta band activity and phase synchrony in executive tasks, it is not surprising that it is highly predictive of CEN activity across numerous executive tasks, as seen in Figure 4.

Theta band activity also plays a major role in the retrieval of autobiographical memories, increasing in amplitude during autobiographical memory recall [18]. It also provides a mechanism for a vmPFC-linked schema instantiation model, suggesting that the vmPFC modulates more posterior long-term memory representations [15]. Additionally, cross-frequency theta-gamma coupling between medial temporal lobe (MTL) structures and cortical regions is critical for the recall of autobiographical memories, and is disrupted in individuals with severely deficient autobiographical memory (SDAM) [13]. Our findings align broadly with this large body of literature, with theta-gamma cross-frequency coupling predicting DMN activity (see Figure 4), further asserting that the trained classifiers identified critical EEG features that represent processes inherent to each ICN.

Interestingly, cross-frequency coupling between delta, alpha and gamma bands was found to be predictive of SN activity, with the gamma and delta bands being concentrated at the right frontal electrodes (see Figure 4i). Combined with the predominantly right lateralization of theta-gamma cross-frequency coupling indicative of DMN activity (see Figure 4h), our results point towards a broadband salience hub anchored close to the right frontal region, which is also active during DMN-linked tasks such as autobiographical memory recall and spatial navigation. This pattern of activity is consistent with activation of the right anterior insula (rAI), a major hub of the salience network [25]. In the context of the tri-network model, the SN is thought to control the switch between the CEN and DMN, and is found to co-activate with the task-appropriate network during the task used in Cohort A [44]. As such, across a wider range of tasks, this node might need to synchronize with numerous brain regions across a wide range of frequencies, as seen in the cross-frequency results of Figure 4i.

In sum, the data-driven analysis pipeline used in this study identified EEG features that captured the core oscillatory dynamics of critical CEN, DMN and SN functions. Furthermore, the classification results demonstrate the collective utility of these salient features in discriminating between CEN, DMN and SN activity.

While this study focused on the three major ICNs within the tri-network model, the discussed approach is equally applicable to identifying unique EEG signatures of other fMRI-derived ICNs, such as ventral and dorsal attention networks (VAN/DAN), somato-sensory networks and the motor network [52, 45]. This opens up the possibility of using EEG to detect the ICN dynamics of any fMRI-derived ICNs, greatly improving the accessibility of such measures. To this end, the code developed for this analysis pipeline, along with a pre-trained generalized classifier, is available for download from github.com/saurabhshaw/EEGnet.

The approach used in this study estimated ICN features using connectivity measures of electrode-level activity, rather than the traditionally used, and much more computationally complex, source-localization techniques. This permitted the exploration of a massive (4M-dimensional) feature space to find optimal features predictive of ICN activity. Unfortunately, when working with features in electrode space (rather than source space), there is the problem of volume conduction, which distorts the neural signal measured at each electrode. To mitigate this, phase-based functional connectivity measures such as phase lag index (PLI) and directed phase lag index (dPLI) were included in the feature space, given their immunity to volume conduction-based distortions [19].

Furthermore, staying in electrode-space, rather than transforming the data to source space, also reduced the computational time required to extract relevant features and predict the predominantly active ICN of a new EEG window. Further algorithmic optimizations were made to parallelize the computation of these features, allowing the use of massively parallel super-computing clusters to accelerate these computations. Future work can use this ability to make rapid predictions of ICN activity to develop novel network-based neurofeedback interventions to directly target the network dysfunction observed in numerous psychopathologies.

In conclusion, this paper takes an important step towards enabling EEG-based investigation of ICN dynamics, greatly increasing the accessibility of such measures in scenarios where fMRI-based ICN assessment might not be practical or economically feasible. This is particularly relevant for studying and characterizing

complex psychopathologies such as PTSD, with various disease subtypes that require drastically different treatment plans. For example, the DMN is erroneously recruited instead of the CEN during working memory in PTSD patients [8], and the pattern of resting state ICN activity can be predictive of the dissociative PTSD subtype [30]. Hence, identifying the pattern of disrupted ICN switching dynamics can greatly help in characterizing the patients' psychopathology and inform their treatment plans. This study makes it clinically feasible to do so by successfully applying cutting-edge machine learning techniques, such as semi-supervised learning, to create an analysis pipeline capable of detecting ICN activation using EEG data alone with high accuracy.

4. Methods

In this study, we developed a purely EEG-based model capable of classifying and tracking ICNs, traditionally investigated using fMRI. To accomplish this, we used simultaneously acquired EEG-fMRI data to identify an optimal EEG-based feature set that could predict the fMRI-derived ICN activation. To maximize the ability of our model to predict ICN activity across a wide range of cognitive tasks, we collected simultaneous EEG-fMRI data from two cohorts of participants that engaged in two different task-switching paradigms, as shown in Figure 1A and B.

4.1. Cohort A Task

The first cohort consisted of 14 participants that dynamically switched between a 2-back working memory (WM) task and an autobiographical memory (ABM) retrieval task, designed to activate the CEN, DMN respectively, and at task-switching points,to activate the SN.

Prior to scanning, each participant recorded up to 10 positive or neutral autobiographical memories in vivid detail, as well as descriptive words corresponding to each memory, that would serve as retrieval cues during the ABM task. Following this, they completed a 1 hour 20 min long memory assessment in the MRI scanner, comprised of randomly ordered 30-second blocks of either cued autobiographical memory retrieval or a 2-back working memory task. Each ABM block included 10 cues pertaining to one of the previously recorded memories, while each WM block included a sequence of 10 to-be-remembered, commonly used English words, each shown for 2 seconds. The participants were instructed to recall the cued memory shown during the ABM blocks, and to remember the word shown two words ago during WM blocks. To assess whether the participant was staying on task, at random points within each block, they were asked to perform a 4-alternative forced choice (4-AFC) trial. During ABM blocks, the 4-AFC trial required the participant to select, from four displayed words, the one representing the memory they were currently recalling from a selection of four words. During WM blocks, the 4-AFC trial required the participant to choose, from four displayed words, the one they saw two words ago. Each pending task switch was cued with 2-second trial showing either Word Memory or Autobiographical Memory, representing an upcoming WM block or ABM block respectively. Each run of sixteen blocks was followed by a 60 second rest period. Participants completed as many 16-block runs as possible within 80 minutes, up to a maximum of 4 runs (64 blocks - 32 ABM, 32 WM). The two distinct block types (ABM and WM) were predicted to activate the DMN and CEN respectively, while the cue to a pending task switch between blocks was expected to activate the SN. Shaw *et al.* [44] reports the activation dynamics of the DMN, CEN and SN observed during this study, using fMRI data alone.

4.2. Cohort B Task

The second cohort consisted of 12 participants that dynamically switched between 7 different tasks. Each participant performed a total of 18 trials of each of the following 7 tasks. This was split into three runs of 6 trials of each of the 7 different task types (42 trials/run), ensuring that trials of the same task type were not repeated more than twice in a row. Each trial began with a pre-recorded, single-word auditory cue indicating the type of the upcoming trial to the participant, followed by a 16-second imagery period. The participant was shown a fixation cross at the center of the screen during this imagery period to avoid any eye movements.

1. **Sport-related motor imagery** Participants were asked to imagine intensely performing a sport or full-body activity (*e.g.*, dancing, jumping jacks) of their choice. Additionally, they were instructed to focus on the kinesthetic and somatosensory aspects of that activity rather than on visual aspects.

2. **Navigation imagery** Participants were asked to imagine navigating around their home from room to room, paying attention to all aspects of the room (*e.g.*, placement of furniture, decor, objects in room).

3. **Music imagery** Participants were asked to imagine listening to a familiar song of their choice through

headphones, while concentrating on all aspects of the song (*e.g.* the melody, instrumentation).

4. **Mental arithmetic** Participants were asked to count backwards by threes from a random 3-digit number of their choice. They were instructed to choose a different 3-digit number for each trial.

5. **Finger tapping imagery** Participants were asked to imagine pushing a button with each of the fingers of the right hand in succession, repeatedly, focusing on the somatosensory and kinesthetic rather than visual aspects of the imagery.

6. **Running imagery** Similar to the sport imagery condition, the participants were instructed to imagine only running, while attending to the kinesthetic and somatosensory aspects of the imagery.

7. **Rest** Participants were asked to clear their mind and think of nothing in particular.

Of these seven tasks, only two tasks were expected to maximally activate the DMN, *i.e.* navigation imagery and rest. The rest of the tasks were expected to activate a combination of attention and somatosensory ICNs that would work in tandem to successfully perform the relevant imagery. Other results exploring the strength of mental imagery in this dataset have been previously reported in Harrison *et al.* [16].

4.3. Data acquisition

Data acquisition for both cohorts was performed at the same site, using the same EEG system and MRI machine.

All EEG data were acquired using a BrainProducts (Brain Products GmbH, Gilching, Germany) 64 channel MR compatible EEG cap, at a sampling rate of 5000Hz. The electrode locations followed the extended international 10-20 system of electrode placement, with the reference at FCz and the ground at AFz. The impedance of each electrode was kept below 10kΩ. The physical setup used for this acquisition is further described in Shaw [42].

All MRI data were acquired using a GE Discovery MR750 3T MRI scanner and an 8 channel RF coil (General Electric Healthcare, Milwaukee, WI). A high resolution anatomical scan was acquired for each participant using an IR prepped axial 3D FSPGR sequence (*Cohort A*: TI/TR/TE=450/7.7/2.2ms, 12° flip angle, 240mm FOV, 2mm-thick slices of size 320 x 192, interpolated to 512 x 512; *Cohort B*: TI/TR/TE=900/10.312/3.22ms, 9° flip angle, 240 x 180mm FOV, 1mm-thick slices of size 512 x 248, interpolated to 512 x 512). These individualized anatomical scans were used to prescribe the fMRI scans, acquired using a 2D GRE EPI sequence (*Cohort A*: TR/TE=2000/35ms, 90° flip angle, 240mm FOV, 3.8mm-thick slices of size 64 x 64, 39 slices/volume interleaved, 300 volumes per functional run; *Cohort B*: TR/TE=3200/35ms, 90° flip angle, 240mm FOV, 4mm-thick slices of size 64 x 64, 40 slices/volume interleaved, 214 volumes per functional run).

4.4. Pre-processing

4.4.1. fMRI pre-processing

All MRI pre-processing steps were performed using SPM12 and the CONN toolbox [58]. The fMRI scans were first realigned and unwarped, followed by motion correction, performed by adding the participant's estimated motion (6 DOF) as a first-level covariate in a denoising general linear model (GLM). This was followed by frequency-domain phase shift slice timing correction (STC), and ART-based identification of outlier scans to be scrubbed. The functional scans were then normalized to the MNI152 atlas by aligning them to each participant's MNI-aligned anatomical scan, followed by segmentation of the functional scans to remove skull, white matter and cerebral spinal fluid (CSF). Spatial smoothing was applied by convolving the BOLD signal with a 2mm Gaussian kernel. Finally, the BOLD data was band-pass filtered between 0.008Hz−0.09Hz.

4.4.2. EEG pre-processing

The gradient artefact (GA) in the EEG data, collected concurrently with the fMRI data, was filtered using a parallel optimized version of the FASTR gradient artefact removal toolbox Shaw [42], which relies on the GA template subtraction algorithm [1]. This was followed by the detection of the QRS complex using data from the ECG lead, which was then used to filter the ballistocardiogram (BCG) using optimal basis set filtering (OBS). The artefact-free EEG data were then downsampled from 5000Hz to 500Hz, followed by temporal band-pass filtering into six different frequency bands - full (1-50Hz), delta (1-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), and gamma (30-50Hz). All temporal filtering was performed twice, once in the forward, and once in the reverse direction for zero phase lag using a 12-order Butterworth IIR filter. Bad channels were then detected and removed based on their spectral characteristics, followed by an ICA decomposition using the FAST-ICA algorithm implemented within EEGLAB [9]. The EEG components representing ocular artefacts, eye blinks and muscle artefacts were detected and removed using the automated ICLabel tool.

Finally, the EEG data were referenced to average EEG channel data, after which the previously removed bad channels were interpolated using spherical interpolation. The bad channels were only interpolated if the number of bad channels was less than 5% of the total number of channels, while also ensuring no two bad channels were neighbours.

4.5. Feature Extraction

4.5.1. FMRI Independent Component Analysis (ICA)

A group independent component analysis (group-ICA) [3] was performed on the denoised fMRI voxel-level data using the iterative FastICA algorithm. This analysis identified 20 mutually independent spatio-temporal patterns of activity, known to represent ICNs. The spatial overlaps (Dice coefficients) of the group-ICA components with known ICNs[45] were used to label the ICA components, identifying the components that corresponded to CEN, DMN or SN sub-networks, as shown in Figure 1C.iii. These group-level ICA components were then back-projected to individual participants' data using GICA back-projection [11], to obtain the activity timeseries of each ICA component. The activity of the CEN, DMN and SN networks were identified by averaging the activity of their constituent sub-networks. Using a sliding-window method, the time courses were segmented into overlapping windows of length 10 seconds (5 TRs) for cohort A and 9.6 seconds (3 TRs) for cohort B. Using a time step of 1 fMRI time point, the time series were segmented into 1184 windows for each participant in cohort A and 636 windows for each participant in cohort B. Each window was labeled with the most active ICN by first thresholding the activity at 70%, followed by comparing the activity of the CEN, DMN and SN during that time period to pick the ICN with the highest activity (shown in Figure 1C.v). These labels were used as the "ground-truth" labels of the ICN activity for all subsequent analyses.

4.5.2. EEG Feature Extraction

A large battery of functional connectivity features (>40 million) were extracted from the EEG data corresponding to each fMRI window. These included single-frequency features representing connectivity between two channels within the same frequency band: coherence (COH), phase lag index (PLI) and directed phase lag index (dPLI); and cross-channel cross-frequency coupling features representing connectivity between two channels, across two different frequency bands: the synchronization index (SI) and phase-amplitude coupling (PAC). Each single-frequency feature was computed for each of the 6 frequency bands (full band, delta, theta, alpha, beta, gamma) described above, and each cross-frequency feature was computed for all possible pairs of the 6 frequency bands. Each fMRI window was further divided into 99 windows of width 200ms, which were used to estimate the features, described in more detail below.

Coherence (COH). represents the synchrony between two channels by comparing their power spectral densities [59], and is computed as follows

$$COH_{xy} = \frac{P_{xy}}{\sqrt{P_{xx}P_{yy}}} \tag{1}$$

where P_{xx} and P_{yy} are the power spectral densities (PSD) of two channels x and y, and P_{xy} is the cross-spectral density (Fourier Transform of cross-correlation) of the signals x and y. COH_{xy} values range between 0 and 1, with 0 representing no coherence between the two signals and 1 representing perfect coherence. Since the power spectral densities are heavily dependent on the signals' amplitudes, COH is sensitive to volume conduction effects.

Phase Lag Index (PLI). is another measure of functional connectivity that addresses the susceptibility of COH-based measures to volume conduction effects. PLI estimates the functional connectivity between two channels (x and y), by estimating the phase-synchronization between them. This relies on the assumption that two channels are functionally connected if there is a consistent phase delay in the signals coming from the two channels. This is defined in terms of their cross-spectrum, given by $C_{xy} = H(x) \cdot H(y)^*$, where $H(\cdot)$ represents the Hilbert transform and * represents the complex conjugate. The PLI is then defined as $PLI_{xy} = |\langle sgn(Im(C_{xy}))\rangle|$, where $Im(C_{xy})$ is the imaginary part of the cross-spectrum (C_{xy}), $sgn(\cdot)$ is the "sign" operator, and $\langle\cdot\rangle$ is the expected value operator. It is important to note that $Im(C_{xy})$ is equivalent to $\Delta\phi_{xy}$-based definition in Stam *et al.* [50]. To minimize spurious noise, we used the weighted PLI variant in this study [54], where the $sgn(Im(C_{xy}))$ is weighted by the imaginary component of the cross-spectrum ($|Im(C_{xy})|$), as given in Equation 2.

$$PLI_{xy} = \frac{|\langle |Im(C_{xy})| \cdot sgn(Im(C_{xy}))\rangle|}{\langle |Im(C_{xy})|\rangle} \tag{2}$$

PLI values lie between 0 and 1, with 0 representing no consistent phase synchrony between channels x and y, and 1 representing perfect phase-locking. To minimize spurious PLI values, every PLI estimate was compared against its surrogate, estimated by the PLI of the

signals spliced at random time points. Only those PLI values that were significantly different from their surrogates were accepted.

Directed Phase Lag Index (dPLI). is a variant of PLI that retains information on phase directionality [49]. This relies on the assumption that lagging signals occur downstream from leading signals, establishing a directional link going from the leading channel to the lagging channel. Equation 3 was used to compute the dPLI for signals from two channels, x and y, of length N.

$$dPLI_{xy} = \frac{1}{N}\sum_{t=1}^{N}\mathcal{H}(\Delta\phi_{xy}(t)) \tag{3}$$

where $\Delta\phi_{xy}(t) = \phi_x(t) - \phi_y(t)$ is the difference in phases of the two signals, and $\mathcal{H}(\cdot)$ is the Heaviside step function. The value of $dPLI_{xy}$ represents the direction of functional connectivity, with $0.5 < dPLI_{xy} \leq 1$ representing x leading y ($x \rightarrow y$), and $0 \leq dPLI_{xy} < 0.5$ representing x lagging y ($y \rightarrow x$). Given its similarity, it is not suprising that PLI and dPLI are related: $PLI_{xy} = 2|0.5 - dPLI_{xy}|$. Similar to PLI estimates, only the dPLI estimates significantly different from its surrogates were retained.

Phase Amplitude Coupling (PAC). is a measure of phase-amplitude cross-frequency coupling between two channels, specifically probing the impact of the phase of the signal in the lower frequency band on the amplitude of the signal in the higher frequency band. Among the several methods of estimating PAC, the modulation index (MI) approach was used due to its superior performance [53]. This approach estimates the PAC between two channels x and y, between two frequency bands f_A and f_B ($f_A < f_B$) using the following steps:

1. Filter the x and y channel data into the respective frequency bands, creating x_{f_A} and y_{f_B}.

2. Extract the phase time series of the lower frequency signal using: $\phi_{X_{f_A}} = phase(H(x_{f_A}))$, and the amplitude time series of the higher frequency signal using: $A_{Y_{f_B}} = amplitude(H(y_{f_B}))$, where $H(\cdot)$ is the Hilbert transform.

3. Bin the lower frequency phases in $\phi_{X_{f_A}}$ into N bins: $\phi_{X_{f_A},j}\ \forall j = 1, 2, ..., N$, and find the mean higher frequency amplitudes for each phase bin j, denoted by $\langle A_{Y_{f_B},j}\rangle_{\phi_{X_{f_A},j}}$. This amplitude distribution is normalized by its sum over all bins, as follows

$$P_{xy}(j) = \frac{\langle A_{Y_{f_B},j}\rangle_{\phi_{X_{f_A},j}}}{\sum_{k=1}^{N}\langle A_{Y_{f_B},k}\rangle_{\phi_{X_{f_A},k}}} \tag{4}$$

   In the case of no PAC, the amplitude distribution $P_{xy}(j)$ is expected to be flat (uniform distribution) since the higher frequency amplitudes will not vary with the lower frequency phase. $N = 20$ was used for estimating the amplitude distribution.

4. The phase amplitude coupling is then estimated as the Kullback-Liebler (KL) divergence ($D_{KL}(\cdot)$) between the observed amplitude distribution $P_{xy}(j)$ and the uniform distribution U, normalized by the maximal possible entropy value (occurs for the uniform distribution).

$$PAC_{xy} = \frac{D_{KL}(P_{xy}, U)}{log(N)} \tag{5}$$

This procedure was repeated for each channel-frequency band pair. Each PAC estimate was retained only if it was significantly different from its surrogate, estimated in a manner similar to the PLI and dPLI measures.

Synchronization Index (SI). is another measure of cross-frequency coupling between two channels, probing the impact of the phase of the signal in the lower frequency band on the phase of the power time series of the signal in the higher frequency band [6]. The following steps were followed to estimate the SI between two channels x and y, between two frequency bands f_A and f_B ($f_A < f_B$).

1. Filter the x and y channel data into the higher frequency band, creating y_{f_B}.

2. Extract the power time series of the signal in the higher frequency band using: $P_{Y_{f_B}} = \|H(y_{f_B})\|^2$, where $H(\cdot)$ is the Hilbert transform. To identify rhythmic fluctuations in this power time series at the lower frequency, compute the FFT of $P_{Y_{f_B}}$ within the lower frequency band (f_A). The peak in this FFT ($f_{AB} = peak(\mathcal{F}(P_{Y_{f_B}}))$) is the synchronization frequency within f_A, which is then used to fine tune the bounds of f_A. The revised lower frequency band $\tilde{f_A}$ is defined by picking a window around the empirically identified synchronization frequency (f_{AB}), using $\tilde{f_A} = [(f_{AB} - 3)\ \ (f_{AB} + 3)]$.

3. Extract the phase time series of the signal within this revised lower frequency band ($\tilde{f_A}$) using: $\phi_{X_{f_A}} = phase(H(x_{f_A}))$, and the phase time series of the higher frequency band power time series using: $\phi_{Y_{f_B}} = phase(H(P_{Y_{f_B}}))$.

4. The synchronization index is then calculated using Equation 6

$$SI_{xy} = \left| \frac{1}{N} \cdot \sum_{t=1}^{N} e^{i \cdot [\phi_{X_{f_A}}(t) \; - \; \phi_{Y_{f_B}}(t)]} \right| \qquad (6)$$

This procedure was repeated for each channel-frequency band pair. Each SI estimate was retained only if it was significantly different from its surrogate, estimated in a manner similar to the PAC, PLI and dPLI measures.

4.6. Feature selection

The feature set was reduced from > 40million features to 5000 features by using a hierarchical version of the popular minimum redundancy maximum relevance (mRMR) feature selection algorithm. mRMR minimizes the mutual information between features while maximizing the mutual information between the features and the class labels [33]. The hierarchical mRMR consisted of repeatedly applying mRMR to identify the top 1000 features at the channel level, followed by selection of the top 1000 features at the frequency band level, followed by selection of the top 1000 features at the window level. This resulted in 1000 top features for each of the five feature types described above, which were then concatenated to create the final 5000-dimensional feature set used for classification. This process is illustrated in Figure 1C.vi.

4.7. Classification

A multi-class support vector machine was trained on the three-way classification task of predicting the most active ICN during each fMRI window, using only the 5000-dimensional EEG feature set identified in section 4.6. Three classifiers were trained for each participant using the identified feature set: a generalized classifier, a semi-supervised individualized classifier and a fully-supervised individualized classifier. The generalized classifier was trained using data from all participants, except the participant being tested (leave-one-out cross validation), followed by prediction of the class labels of the test-participant. The semi-supervised individualized classifier used the highly confident label predictions as "expert" labels, while leaving the rest of the time windows unlabelled to select a participant-specific feature set, and use it to classify the unlabelled time points. Finally, the fully-supervised individualized classifier was trained on a subset of the data from the same participant being tested, ensuring that the training set and test set did not contain overlapping windows. A 75/25 train/test split was used with 20-fold cross validation for testing the performance of this classifier.

References

[1] Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., and Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *NeuroImage*, **8**(3), 229–239.

[2] Bluhm, R. L., Williamson, P. C., Osuch, E. A., Frewen, P. A., Stevens, T. K., Boksman, K., Neufeld, R. W. J., Théberge, J., and Lanius, R. A. (2009). Alterations in default network connectivity in posttraumatic stress disorder related to early-life trauma. *Journal of psychiatry & neuroscience : JPN*, **34**(3), 187–94.

[3] Calhoun, V. D., Liu, J., and Adal, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage*, **45**(1), S163–S172.

[4] Cavanagh, J. F. and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, **18**(8), 414–421.

[5] Charbonnier, S., Roy, R., Doležalová, R., Campagne, A., and Bonnet, S. (2016). Estimation of Working Memory Load using EEG Connectivity Measures. *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies*, **4**(Biostec), 122–128.

[6] Cohen, M. X. (2008). Assessing transient cross-frequency coupling in EEG data. *Journal of Neuroscience Methods*, **168**(2), 494–499.

[7] Connolly, C. G., Wu, J., Ho, T. C., Hoeft, F., Wolkowitz, O., Frank, G., Hendren, R., Max, J. E., Paulus, M. P., Susan, F., Banerjee, D., Simmons, A. N., and Yang, T. T. (2014). Resting-State Functional Connectivity of Subgenual Anterior Cingulate Cortex in Depressed Adolescents. **74**(12), 898–907.

[8] Daniels, J. K., McFarlane, A. C., Bluhm, R. L., Moores, K. A., Clark, C. R., Shaw, M. E., Williamson, P. C., Densmore, M., and Lanius, R. A. (2010). Switching between executive and default mode networks in posttraumatic stress disorder: alterations in functional connectivity. *Journal of psychiatry & neuroscience : JPN*, **35**(4), 258–66.

[9] Delorme, A. and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**(1), 9–21.

[10] Dimitriadis, S. I., Sun, Y., Kwok, K., Laskaris, N. A., Thakor, N., and Bezerianos, A. (2015). Cognitive Workload Assessment Based on the Tensorial Treatment of EEG Estimates of Cross-Frequency Phase Interactions. *Annals of Biomedical Engineering*, **43**(4), 977–989.

[11] Erhardt, E. B., Rachakonda, S., Bedrick, E. J., Allen, E. A., Adali, T., and Calhoun, V. D. (2011). Comparison of multi-subject ICA methods for analysis of fMRI data. *Human Brain Mapping*, **32**(12), 2075–2095.

[12] Foster, B. L. and Parvizi, J. (2012). Resting oscillations and cross-frequency coupling in the human posteromedial cortex. *NeuroImage*, **60**(1), 384–391.

[13] Fuentemilla, L., Palombo, D. J., and Levine, B. (2018). Gamma phase-synchrony in autobiographical memory: Evidence from magnetoencephalography and severely deficient autobiographical memory. *Neuropsychologia*, **110**(August 2017), 7–13.

[14] Garrido, M. I., Barnes, G. R., Kumaran, D., Maguire, E. A., and Dolan, R. J. (2015). Ventromedial prefrontal cortex drives hippocampal theta oscillations induced by mismatch computations. *NeuroImage*, **120**, 362–370.

[15] Gilboa, A. and Moscovitch, M. (2017). Ventromedial prefrontal cortex generates pre-stimulus theta coherence desynchronization: A schema instantiation hypothesis. *Cortex*, **87**, 16–30.

[16] Harrison, A. H., Noseworthy, M. D., Reilly, J. P., Guan, W., and Connolly, J. F. (2017). EEG and fMRI agree: Mental arithmetic is the easiest form of imagery to detect. *Consciousness and Cognition*, **48**, 104–116.

[17] Khanna, A., Pascual-Leone, A., Michel, C. M., and Farzan, F. (2015). Microstates in resting-state EEG: Current status and future directions. *Neuroscience and Biobehavioral Reviews*, **49**, 105–113.

[18] Klimesch, W., Doppelmayr, M., Stadler, W., Pöllhuber, D., Sauseng, P., and Röhm, D. (2001). Episodic retrieval is reflected by a process specific increase in human electroencephalographic theta activity. *Neuroscience Letters*, **302**(1), 49–52.

[19] Lai, M., Demuru, M., Hillebrand, A., and Fraschini, M. (2018). A comparison between scalp- and source-reconstructed EEG networks. *Scientific Reports*, **8**(1), 1–8.

[20] Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., and McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology*, **6**(1), 27313.

[21] Liu, Q., Ganzetti, M., Wenderoth, N., and Mantini, D. (2018). Detecting Large-Scale Brain Networks Using EEG: Impact of Electrode Density, Head Modeling and Source Localization. *Frontiers in Neuroinformatics*, **12**(March), 1–11.

[22] Lundberg, S. M., Allen, P. G., and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30*, pages 4765–4774.

[23] Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, **15**(10), 483–506.

[24] Menon, V. (2015). *Salience Network*, volume 2. Elsevier Inc.

[25] Menon, V. and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, pages 1–13.

[26] Michel, C. M. and Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *NeuroImage*, (November), 1–17.

[27] Mishra, A., Englitz, B., and Cohen, M. X. (2020). EEG microstates as a continuous phenomenon. *NeuroImage*, **208**, 116454.

[28] Moretti, D. V., Fracassi, C., Pievani, M., Geroldi, C., Binetti, G., Zanetti, O., Sosta, K., Rossini, P. M., and Frisoni, G. B. (2009). Increase of theta/gamma ratio is associated with memory impairment. *Clinical Neurophysiology*, **120**(2), 295–303.

[29] Nejad, A. B., Fossati, P., and Lemogne, C. (2013). Self-referential processing, rumination, and cortical midline structures in major depression. *Frontiers in Human Neuroscience*, **7**(OCT), 1–9.

[30] Nicholson, A. A., Harricharan, S., Densmore, M., Neufeld, R. W., Ros, T., McKinnon, M. C., Frewen, P. A., Théberge, J., Jetly, R., Pedlar, D., and Lanius, R. A. (2020). Classifying Heterogeneous Presentations of PTSD via the Default Mode, Central Executive, and Salience Networks with Machine Learning. *NeuroImage: Clinical*, **Submitted**, 102262.

[31] Northoff, G. (2015). Is schizophrenia a spatiotemporal disorder of the brain's resting state? *World Psychiatry*, **14**(1), 34–35.

[32] O'Neill, G. C., Tewarie, P., Vidaurre, D., Liuzzi, L., Woolrich, M. W., and Brookes, M. J. (2018). Dynamics of large-scale electrophysiological networks: A technical review. *NeuroImage*, **180**(October 2017), 559–576.

[33] Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **27**(8), 1226–1238.

[34] Rabellino, D., Tursich, M., Frewen, P. A., Daniels, J. K., Densmore, M., Théberge, J., and Lanius, R. A. (2015). Intrinsic Connectivity Networks in post-traumatic stress disorder during sub- and supraliminal processing of threat-related stimuli. *Acta Psychi-*

atrica Scandinavica, **132**(5), 365–378.

[35] Raniga, P., Paton, B., and Egan, G. (2016). Resting state functional coupling between the ascending synchronising system, limbic system and the default mode network via theta oscillations. *bioRxiv*, page 086058.

[36] Sammer, G., Blecker, C., Gebhardt, H., Bischoff, M., Stark, R., Morgen, K., and Vaitl, D. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. *Human Brain Mapping*, **28**(8), 793–803.

[37] Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, **57**(2), 97–103.

[38] Sauseng, P., Griesmayr, B., Freunberger, R., and Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, **34**(7), 1015–1022.

[39] Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., and Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International Journal of Psychophysiology*, **67**(3), 242–251.

[40] Senoussi, M., Verbeke, P., Desender, K., De Loof, E., Talsma, D., Verguts, T., and Dunantlaan, H. (2020). Theta oscillations shift towards optimal frequency for cognitive control. *bioRxiv*, page 2020.08.30.273706.

[41] Shah, N. J., Arrubla, J., Rajkumar, R., Farrher, E., Mauler, J., Kops, E. R., Tellmann, L., Scheins, J., Boers, F., Dammers, J., Sripad, P., Lerche, C., Langen, K. J., Herzog, H., and Neuner, I. (2017). Multimodal Fingerprints of Resting State Networks as assessed by Simultaneous Trimodal MR-PET-EEG Imaging. *Scientific Reports*, **7**(1), 6452.

[42] Shaw, S. (2017). Real-time filtering of gradient artifacts from simultaneous EEG-fMRI data. In *2017 International Workshop on Pattern Recognition in Neuroimaging, PRNI 2017*.

[43] Shaw, S. B., Dhindsa, K., Reilly, J. P., and Becker, S. (2019). Capturing the Forest but Missing the Trees: Microstates Inadequate for Characterizing Shorter-Scale EEG Dynamics. *Neural computation*, **31**(11), 2177–2211.

[44] Shaw, S. B., McKinnon, M. C., Heisz, J., and Becker, S. (2021). Dynamic task-linked switching between brain networks A trinetwork perspective. *Brain and Cognition*, **151**(April), 105725.

[45] Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, **22**(1), 158–165.

[46] Sockeel, S., Schwartz, D., Pélégrini-Issac, M., and Benali, H. (2016). Large-Scale Functional Networks Identified from Resting-State EEG Using Spatial ICA. *PLOS ONE*, **11**(1), e0146845.

[47] Sridharan, D., Levitin, D. J., and Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, **105**(34), 12569–12574.

[48] Sripada, R. K., King, A. P., Welsh, R. C., Garfinkel, S. N., Wang, X., Sripada, C. S., and Liberzon, I. (2012). Neural dysregulation in posttraumatic stress disorder: Evidence for disrupted equilibrium between salience and default mode brain networks. *Psychosomatic Medicine*, **74**(9), 904–911.

[49] Stam, C. J. and van Straaten, E. C. (2012). Go with the flow: Use of a directed phase lag index (dPLI) to characterize patterns of phase relations in a large-scale model of brain dynamics. *NeuroImage*, **62**(3), 1415–1428.

[50] Stam, C. J., Nolte, G., and 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.

[51] Stoddard, J., Gotts, S. J., Brotman, M. A., Lever, S., Hsu, D., Zarate, C., Ernst, M., Pine, D. S., and Leibenluft, E. (2016). Aberrant intrinsic functional connectivity within and between corticostriatal and temporal-parietal networks in adults and youth with bipolar disorder. *Psychological Medicine*, **46**(7), 1509–1522.

[52] Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., and Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, **106**(3), 1125–1165.

[53] Tort, A. B., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of Neurophysiology*, **104**(2), 1195–1210.

[54] Vinck, M., Oostenveld, R., Van Wingerden, M., Battaglia, F., and Pennartz, C. M. (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.

[55] Wang, P., Zhou, B., Yao, H., Zhan, Y., Zhang, Z., Cui, Y., Xu, K., Ma, J., Wang, L., An, N., Zhang, X., Liu, Y., and Jiang, T. (2015). Aberrant intra-and inter-network connectivity architectures in Alzheimer's disease and mild cognitive impairment. *Scientific Reports*, **5**(May), 1–12.

[56] Wang, Y., Gao, Y., Tang, S., Lu, L., Zhang, L., Bu, X., Li, H., Hu, X., Hu, X., Jiang, P., Jia, Z., Gong, Q., Sweeney, J. A., and Huang, X. (2020). Large-scale network dysfunction in the acute state compared to the remitted state of bipolar disorder: A meta-analysis of resting-state functional connectivity. *EBioMedicine*, **54**, 102742.

[57] White, T. P., Joseph, V., Francis, S. T., and Liddle, P. F. (2010). Aberrant salience network (bilateral insula and anterior cingulate cortex) connectivity during information processing in schizophrenia. *Schizophrenia Research*, **123**(2-3), 105–115.

[58] Whitfield-Gabrieli, S. and Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*.

[59] Yaesoubi, M., Allen, E. A., Miller, R. L., and Calhoun, V. D. (2015). Dynamic coherence analysis of resting fMRI data to jointly capture state-based phase, frequency, and time-domain information. *NeuroImage*, **120**, 133–142.

6.3 Discussion

This chapter develops an analysis pipeline that can track ICN activity using EEG data alone. The resultant system relies on learned spatio-temporal features that are unique to each of the three ICNs in the tri-network model.

The utilization of simultaneous EEG-fMRI data allowed the extraction of fMRI-derived network labels; these labels were used as supervisory signals to train a classifier to select relevant EEG features, which could then be applied to classify new EEG data (without the need for fMRI labels). This created a machine learning EEG-based model that can be directly interpreted in the context of fMRI-derived ICNs. This multimodal analysis also integrated ICN features with high spatial and temporal resolutions (including frequencies up to gamma frequency activity), generating a rich description of the spatio-temporal characteristics of each ICN. These spatio-temporal signatures form a generalized classifier that can be downloaded by any researcher and used to probe ICN dynamics in their EEG data.

The accuracy of this machine learning model was further boosted by utilizing semi-supervised learning to "correct" the predicted ICN activation at the time points where the generalized classifier was not very confident in its predictions. This semi-supervised classifier can also be used by any researcher with EEG data, by downloading the analysis pipeline linked in the paper provided.

One caveat to this analysis is the use of an apparently large number of features (around 5000) to characterize each ICN. While this might seem like a rather large number of features for a classification task, it was empirically found to be the optimal number of features that could characterize the connectivity between brain regions that constitute each ICN, without suffering from overfitting.

This ICN analysis tool greatly increases the accessibility of ICN-based analysis and opens the door for its use as a clinical biomarker in characterizing and treating mental health disorders. The next chapter validates this tool by using it to track changes in ICN dynamics due to lifestyle interventions.

Bibliography

Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., and McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology*, 6(1):27313.

Sitaram, R., Ros, T., Stoeckel, L. E., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M., Rana, M., Oblak, E., Birbaumer, N., and Sulzer, J. (2016). Closed-loop brain training: the science of neurofeedback. *Nature Neuroscience*.

Chapter 7

Intervention-linked changes in ICNs

7.1 Introduction

This chapter builds upon the information gathered from chapters 4, 5, and 6 and validates the developed EEG-based ICN analysis tool.

Given the ultimate use case of the developed EEG-based ICN analysis tool to detect clinically significant abnormalities in brain dynamics and track their re-normalization through a course of treatment, it needs to be sensitive to consistent changes in ICN dynamics, such as those produced by psychiatric therapies or interventions (Lanius et al., 2015).

The study presented in this chapter provides a proof-of-concept investigation of such intervention-linked changes in ICN dynamics, assessed entirely using EEG data and the analysis tool developed in chapter 6. Based on the characteristic healthy network dynamics observed in chapter 5, the activation of DMN and CEN were

compared to that of the SN to identify periods of task-linked co-activation.

The primary author (SBS) of the study conceptualized and designed the EEG section of this study, along with the last author (SB). SBS, YL, and AM recruited participants, performed assessments and administered the interventions. SBS and SB also developed the functional task used in this study. SBS and SB co-developed the novel analysis pipeline discussed in this paper (described in the previous chapter) and wrote the first draft of the manuscript. All co-authors contributed to the writing of the manuscript. MM, JH and SB contributed funding to this study.

7.2 Combing aerobic exercise and mindfulness-like Neurofeedback for better intrinsic connectivity network (ICN) synchrony

Shaw, S. B., Levy, Y., Mizzi, A., McKinnon, M. C., Heisz, J., & Becker, S. (2021). Combing aerobic exercise and mindfulness-like Neurofeedback for better for intrinsic connectivity network (ICN) synchrony. bioRxiv

Combining Aerobic Exercise and Mindfulness-like Neurofeedback (NFB) for better Intrinsic Connectivity Network (ICN) Synchrony

Saurabh Bhaskar Shaw[1,3,4], Yarden Levy[2,4], Allison Mizzi[2,4], Margaret C. McKinnon[5,6,7], Jennifer Heisz[8], and Suzanna Becker[1,2,3,4]

[1]*Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada*
[2]*Department of Psychology Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada*
[3]*Vector Institute for Artificial Intelligence, Toronto*
[4]*Centre for Advanced Research in Experimental and Applied Linguistics (ARiEAL), Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada*
[5]*Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada*
[6]*Mood Disorders Program, St. Joseph's Healthcare, Hamilton, ON, Canada*
[7]*Homewood Research Institute, Guelph, ON, Canada*
[8]*Department of Kinesiology, McMaster University, Hamilton, ON, Canada*

Abstract

Lifestyle interventions have the potential to prevent mental health disorders in vulnerable populations, and prevent dysregulation of intrinsic brain network (ICN) dynamics, seen in many psychopathologies. Multiple lifestyle interventions can even be combined to improve outcomes due to synergies between each individual intervention. While some studies have shown the mood benefits of combining lifestyle interventions, its effect on task-linked ICN dynamics is yet unknown. This study investigated the separate and combined effects of two such lifestyle interventions - aerobic exercise and mindfulness-like neurofeedback (NFB), on task-linked ICN dynamics of the default mode network (DMN), central executive network (CEN), and the salience network (SN). The recruited participants were divided into four experimental groups - Control, Running, NFB, and Combined, performing the corresponding intervention for 16 sessions distributed over 8-weeks. Intervention-linked changes in ICN dynamics were studied using EEG-based neuroimaging scans before and after the 8-week intervention, during which participants performed multiple blocks of autobiographical memory recall (AM) and working memory recall (WM), designed to activate the DMN, CEN and SN. Combining the two interventions resulted in a more healthy CEN-SN synchrony pattern in responders, and prevented a maladaptive pattern of persistent DMN-SN synchrony in non-responders, compared to each intervention alone. Furthermore, the

CEN-SN synchrony at baseline could predict NFB response with upto 80% accuracy, demonstrating the utility of such network-based biomarkers in personalizing intervention plans.

Keywords: Aerobic exercise, Mindfulness, Neurofeedback, Intrinsic Connectivity Networks (ICN), Default Mode Network (DMN), Central Executive Network (CEN), Salience Network (SN), Tri-Network Model.

1 Introduction

Mood disorders are becoming increasingly prevalent in the general population, and young adults are at a particularly elevated risk for depressive symptoms and suicidal thoughts (Public Health Agency of Canada, 2014). After accidents, suicide is one of the most frequent causes of death for Canadian youth between the ages of 15-24 years (Mental Health Commission Of Canada, 2013). Moreover, youth in particular are much less likely to seek traditional psychotherapy and pharmacological treatments for their mood disorders (Campo and Bridge, 2009), leaving them uniquely vulnerable to the increasing amount of stressors in their life. Making matters worse, such traditional therapies are often expensive, not universally accessible, and are not always effective (Wang *et al.*, 2019; Campo and Bridge, 2009). Therefore, accessible alternative intervention strategies are urgently needed to help youth at risk, ideally in the form of simple lifestyle interventions. We propose aerobic exercise and/or neurofeedback-based mindfulness as two such candidate interventions and investigate their impact on mood and brain function in a cohort of university students.

Exercise has long known to ameliorate mood and cognitive deficits associated with depression (Cooney *et al.*, 2013; Keating *et al.*, 2018), improve attention (Prakash *et al.*, 2011; Chang *et al.*, 2015), some forms of memory (Monti *et al.*, 2012; Baym *et al.*, 2014; Schwarb *et al.*, 2017) and overall executive function (Guiney and Machado, 2013). It is also hypothesized to improve adult neurogenesis within the dentate gyrus (DG), associated with reduced memory interference (Déry *et al.*, 2013), increased hippocampal volume and improved memory performance (Erickson *et al.*, 2011). Such benefits extend to the level of large-scale intrinsic connectivity networks (ICNs), with aerobic fitness explaining individual differences in functional connectivity within the central executive network (CEN), the default mode network (DMN), and dorsal/ventral attentional networks (DAN/VAN) (Talukdar *et al.*, 2018). Even a single bout of moderate intensity aerobic exercise can enhance executive control-linked event related potentials (ERPs), indicating enhanced engagement of executive networks (Chang *et al.*, 2015).

Mindfulness or meditation is another lifestyle intervention found to be similarly effective in managing stress (Chiesa and Serretti, 2010), enhancing attention (Kozasa *et al.*, 2012), enriching executive functioning (Gallant, 2016) and improving mood. Mindfulness training has also shown to increase grey matter in vital areas such as the cerebellum

and posterior cingulate cortex (Holzel *et al.*, 2011), increase telomerase activity associated with reduced oxidative stress (Epel *et al.*, 2010; Jacobs *et al.*, 2011), and increase the synchronization between cardiac and brain activity (Gao *et al.*, 2016) important for appropriate processing of emotional stimuli (Gray *et al.*, 2012). Mindfulness based therapies lead to increased connectivity of the DMN (posterior cingulate cortex - PCC) with critical SN (dorsal anterior cingulate cortex - dACC) and CEN (dorsolateral pre-frontal cortex - dlPFC) nodes (King *et al.*, 2016; Doll *et al.*, 2015), and have shown promise in ameliorating PTSD symptoms (King *et al.*, 2016). Such therapies can also modify the temporal dynamics of brain network connectivity, leading to a re-organization of EEG microstates post-therapy (Brechet *et al.*, 2021). However, despite its benefits, mindfulness-based therapies traditionally require a highly trained practitioner/clinician leading the mindfulness sessions. Much like traditional clinician-driven psychotherapies, this reduces the accessibility of such lifestyle-based interventions. One alternative is to use mindfulness-like therapies that can be self-administered, such as EEG-based neurofeedback (NFB) training, which have been found to be similarly impactful. A single NFB training session can upregulate connectivity of critical SN nodes, such as the dACC (Ros *et al.*, 2013) and right insula (Kluetsch *et al.*, 2014), in a manner similar to mindfulness training (Kilpatrick *et al.*, 2011). Consequently, it has been found to be clinically useful in improving working memory, concentration, impulsivity and dissociative symptoms in a wide range of psychopathologies, such as ADHD (Thibault *et al.*, 2016; Escolano *et al.*, 2014) and PTSD (**?**Sitaram *et al.*, 2016). While clinical NFB training therapies still rely on therapist-driven care, the advent of commercially available, low-cost EEG systems has allowed the development of easy-to-use mediation-like NFB training devices, such as the MUSE headband. When paired with a smartphone app on the user's phone, it can be used to administer EEG-based neurofeedback (NFB) training without the need for a trained clinician or practitioner. Hence, such a system could be a viable self-driven alternative to traditional mindfulness training protocols, while still providing similar cognitive and mood benefits.

Finally, there is some evidence surrounding the synergistic benefits of combining these two interventions. For example, Mental and Physical (MAP) training, a combined meditation and aerobic exercise intervention, is thought to decrease symptoms of depression and anxiety through improved neurogenesis (Shors *et al.*, 2014). While, aerobic exercise alone is well known to upregulate hippocampal neurogenesis, a sizeable proportion of new neurons are lost to programmed apoptosis in the absence of intensive mental training processes (Curlik and Shors, 2013). This synergistic effect is believed to underlie the benefits of MAP training. Given the crucial role of hippocampal circuits in large-scale brain network dynamics (Shaw *et al.*, 2021a), such a combined aerobic exercise and mindfulness training protocol can potentially augment network behaviour in individuals with dysfunctional network dynamics, such as those seen in PTSD and bipolar disorder. In fact, MAP training was found to be effective in reducing post-traumatic cognitions and ruminative thoughts in a population of PTSD patients (Shors *et al.*, 2018).

While studies using MAP training have shown that a combination of mental training

and physical activity can act synergistically compared to each intervention alone (Shors *et al.*, 2018), this research is lacking in two key areas. Firstly, group mindfulness classes, though effective, are not readily available or manageable by everyone. There is a need to investigate more accessible mindfulness alternatives such as mindfulness-like NFB training. Secondly, the effects of exercise and/or NFB training on task-related brain network dynamics is not well understood. While some studies have investigated exercise or NFB-linked brain network changes (Kluetsch *et al.*, 2014; Brechet *et al.*, 2021; Talukdar *et al.*, 2018), none have studied the combined effect of both interventions. Furthermore, most of these studies probe the resting-state dynamics of large-scale brain networks and do not investigate the intervention-linked changes in network dynamics under task loads.

The current study addresses these knowledge gaps by investigating the separate and combined effects of aerobic exercise and mindfulness-like NFB training on task-related brain network dynamics in a cohort of healthy undergraduate students.

2 Methods

A total of 140 participants were recruited for the study, after excluding participants who answered "Yes" on any of the questions in the Get Active Questionnaire (GAQ). They were randomly assigned to one of the four experimental groups - Control, Running, NFB, and Combined (Running + NFB), as shown in Figure 1 and further detailed below.

1. Those in the **Running** group partook in an 8-week aerobic exercise protocol comprising of 2 sessions of 24 minutes each week. Inspired by Galloway (2016)'s Run-Walk-Run method, each 24 minute session was further divided into 8 intervals of 3 minutes each, consisting of running and walking. The difficulty of each session was increased over time by increasing amount of time participants spent running in each 3 minute interval (Table 1).

2. On the other hand, those within the **Neurofeedback (NFB)** group performed mindfulness-like NFB training 2 times a week, for an 8-week period. Similar to the exercise protocol, the difficulty of the NFB sessions were increased over time by increasing the session lengths as the participants progressed through the weeks (Table 1).

3. Finally, participants in the **Combined** group (Running + NFB) performed both, the NFB training and the aerobic exercise protocol at each session.

4. Participants assigned to the **Control** group did not participate in any of the training protocols.

Before and after the 8-week protocol, each participant also completed mood questionnaires, a VO2-max assessment and a task-based neuroimaging scan, as shown in Figure 1.

Week	Running Protocol		Neurofeedback Protocol
	Run Interval	Walk Interval	NFB Length
1	1:00	2:00	5:00
2	1:15	1:45	7:00
3	1:30	1:30	9:00
4	1:45	1:15	11:00
5	2:00	1:00	13:00
6	2:15	0:45	15:00
7	2:30	0:30	17:00
8	2:45	0:15	19:00

Table 1: The running and NFB protocols used by participants in this study. Each running session consisted of eight 3 minute intervals, which were split into varying amounts of running and walking, as shown in the table. The difficulty of the running sessions increased every week by increasing the proportion of time spent running in each 3 minute interval. In contrast, the NFB difficulty was increased by increasing the length of time the participants spent performing mindulness-like NFB.

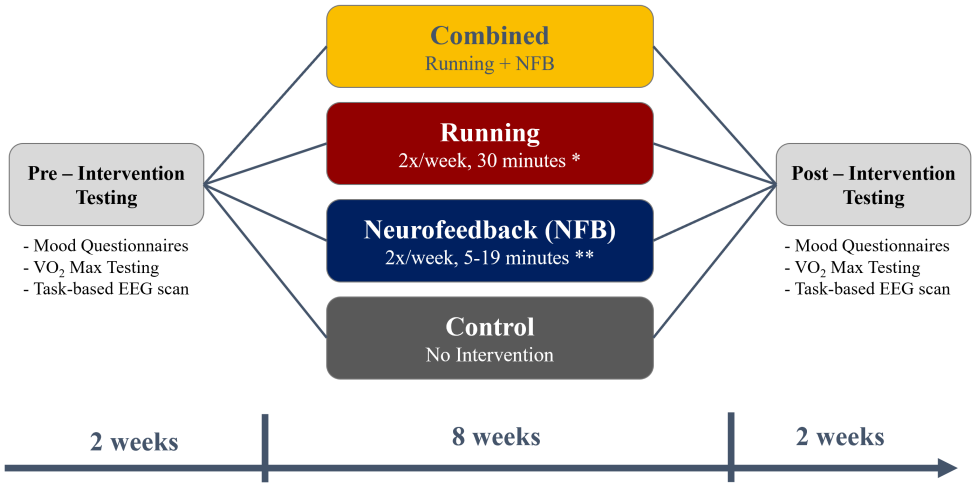

Figure 1: A flowchart depicting the timing of the pre-intervention testing (2 weeks), the intervention (8 weeks), and post-intervention testing (2 weeks). The timings were chosen to fit within the first 3 months of the 4-month long university term.

During the neuroimaging scan, high resolution (128-channel) EEG data was recorded while participants performed multiple blocks of an autobiographical memory recall (AM) and a 2-back working memory task (WM). This memory paradigm is identical to that described in Shaw *et al.* (2021a) and is designed to activate the three intrinsically connected networks (ICNs) of the tri-network model (Menon, 2011), namely, the default mode network (DMN), the central executive network (CEN), and the salience network (SN). All research protocols used in the study were scrutinized and approved by the McMaster University Research Ethics Board (MREB).

Of the 140 recruited participants, 69 participants completed the full intervention protocol and pre/post-intervention testing yielding an attrition rate of 50%. This included 15 participants in the Control group, 20 participants in the Running group, 14 participants in the NFB group, and 20 participants in the Combined group.

To study intervention-linked changes to large-scale brain networks, the task-based EEG data from the pre-intervention and post-intervention testing sessions were processed using a novel connectivity pipeline designed to detect ICN activation using EEG data alone, as described in Shaw *et al.* (2021b). This provided the probability of DMN (P_{dmn}), CEN (P_{cen}) and SN (P_{sn}) activation over 5 second windows of EEG data. The pairwise cross-entropy (PCE) between the probability distributions of each network pair was computed to estimate their similarity/synchronization over the course of the WM and AM trials. For example, equation 1 shows the process of estimating the PCE between DMN and SN over M trials of either WM or AM task. According to this definition, a lower PCE_{dmn-sn} corresponds to higher similarity between the probability distributions of DMN activation and SN activation across multiple trials of either WM or AM task, which can be interpreted as higher DMN-SN synchronization over the M trials.

$$PCE_{dmn-sn} = -\sum_{i=1}^{M} P_{dmn} \times log(P_{sn}) \tag{1}$$

This was performed across the entire length of the WM and AM trials, followed by separately estimating the PCE across the first half and second half of the WM and AM trials.

We predicted that the Running group and the NFB training group would show improvements in their VO2-max levels and their NFB scores, respectively (hypothesis 1). We further hypothesized that the combination of the two interventions would produce a synergistic effect, and lead to significant enhancements in the Combined group. These enhancements are expected to be significantly greater than those seen in the separate exercise protocol, and the separate NFB training protocol. Furthermore, the observed pattern of inter-network activity between the three core networks discussed, was expected to change over the course of the study (hypothesis 2).

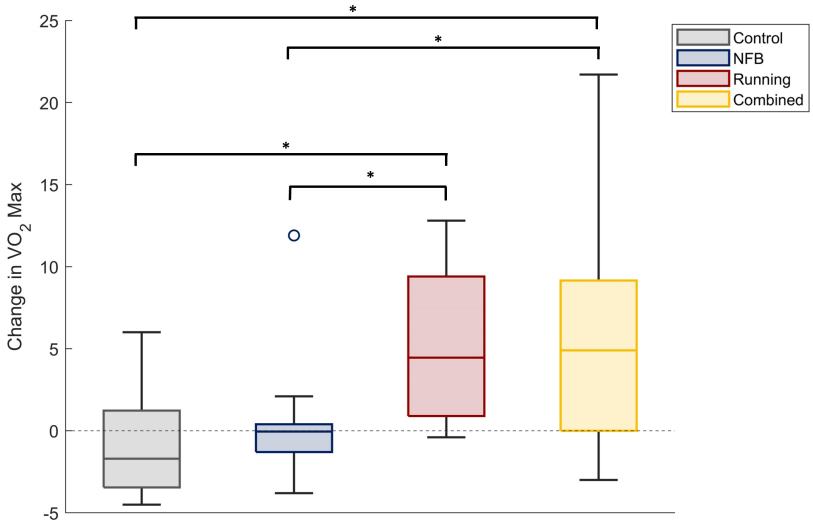

Figure 2: The change in VO2-max scores of participants within each experimental group. The Running and Combined groups show a significant increase in VO2Max compared to the Control and NFB groups.

3 Results

3.1 Intervention effectiveness

We first assessed the effectiveness of the aerobic exercise and mindfulness-like NFB interventions by conducting one-way ANOVAs with participant group as the independent factor (IV), and various outcome variables as the dependent variable, as discussed below.

3.1.1 Aerobic exercise

The change in each particpants' VO2-max level from baseline (Pre-timepoint) was used to assess the effectiveness of the aerobic exercise protocol in increasing their aerobic fitness. A significant effect of group was observed ($F=3.43$, $p=0.028$), with participants' in the Running and the Combined groups showing a significant increase in their VO2-max levels, compared to the Control and NFB groups (shown in Figure 2). The exercise protocol employed in this study was thereby effective in increasing the aerobic capacity of the participants.

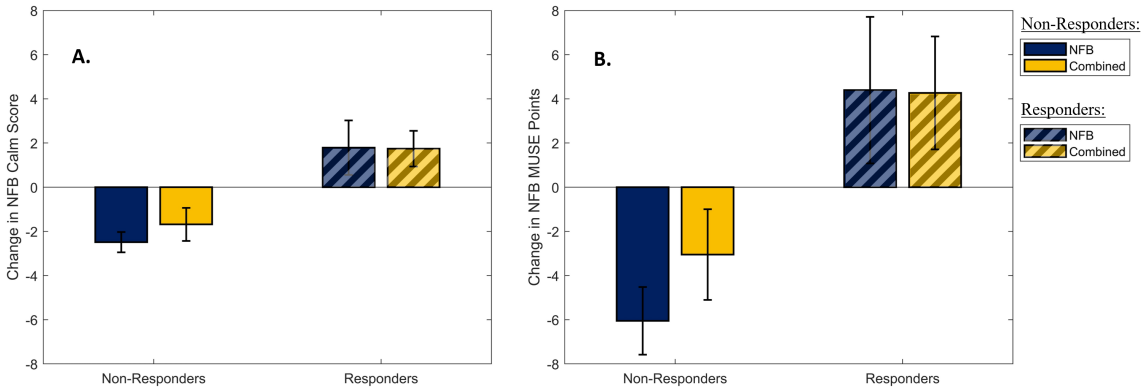

Figure 3: The distribution of the two metrics used to determine NFB "Responders" and "Non-Responders". **A.** The mean slope of the amount of time participants spent in the "Calm" state during their mindfulness-like NFB training sessions. **B.** The mean change in the MUSE points during each mindufulness-like NFB training sessions.

3.1.2 Mindfulness-like NFB

Two different outcome measures were used to assess the efficacy of the NFB intervention due to its bi-factorial nature, merging mindfulness-like aspects with NFB training. The first dependent variable was the slope of the percentage of time spent in the "calm zone" over the course of the protocol. This provided an estimate of the participants' meditative ability. The second dependent variable used was the slope of the points awarded by MUSE through the course of each NFB session. Although linked to the time spent in the calm zone, this measure was more directly linked to the NFB performance of the participants.

While there was no statistically significant effect of the NFB protocol on the two outcome measures described above, some participants were successful in achieving higher scores (positive slope) while others were unable to do so (negative slope). These measures were used to split the participants into responders and non-responders on the basis of their success in increasing their NFB score over the 8-week period, as shown in Figure 3.

3.2 Network Dynamics

To study the impact of the interventions on task-dependent network dynamics, the co-activation of the CEN and DMN with the SN were studied by assessing the pairwise cross-entropy (PCE) between the task-linked probability distribution of CEN/DMN activation, and that of SN activation during corresponding task intervals. The PCE between each pair of networks was expected to be lower for higher synchronization/co-activation of the networks with the SN. The interventions were expected to shift the participants' network dynamics towards a healthy pattern of network co-activation, where the SN was expected

to co-activate with the task-appropriate network, *i.e.* the DMN during the AM trials, and the CEN during the WM trials (Shaw *et al.*, 2021a). Hence, the interventions were expected to decrease the PCE between DMN and SN during the AM trials, and between the CEN and SN during the WM trials. Conversely, the interventions were expected to increase the PCE between the DMN and SN during the WM trials, and between the CEN and SN during the AM trials.

To test this, two-way repeated measures ANOVA was performed with each of the PCE as the outcome variable, Group and Responder as the two between-group factors, and Session as the within-group factor. This was performed for each of the two outcome variables of NFB response discussed in the previous section - Responder (calm state), and Responder (NFB points). These findings are discussed below.

3.2.1 NFB non-responders show increased task-opposite DMN synchronization with SN

A significant Session x Group x Responder (NFB points) interaction was observed for the PCE between DMN and SN during the WM trials (F=6.859,p=0.002). The DMN-SN PCE was found to be significantly lower in the non-responders within the NFB group at the post-intervention session, compared to the responders within the same group (Figure 4A). This implies that the DMN was more synchronized with SN during the WM trials in the non-responders within the NFB group, indicating that NFB non-responders could have shifted their task-linked network dynamics towards a more dysregulated state.

Furthermore, this pattern was observed in both halfs of the WM trials (First half: F=3.398,p=0.041; Second half: F=5.968,p=0.018), as seen in Figure 4B and Figure 4C respectively. This is in contrast to the healthy pattern of DMN deactivation observed during the second half of the WM trials (Shaw *et al.*, 2021a). This finding further suggests that the non-responders within the NFB group might have trouble deactivating the DMN according to changing task demands.

3.2.2 Addition of aerobic exercise to NFB training rescues task-opposite DMN-SN synchronization within NFB non-responders

Interestingly, the pattern of lower DMN-SN PCE observed within the non-responders of the NFB group, was not observed in the non-responders of the Combined group (Figure 4 A,B and C). This suggests a protective role of aerobic exercise in preventing the development of maladaptive network dynamics.

3.2.3 CEN more synchronized with SN during WM trials in responders within the combined group

A significant Group x Responder (calm state) interaction was observed for the PCE between CEN and SN during the WM trials (F=3.266,p=0.045). As seen in Figure 5A and D, the

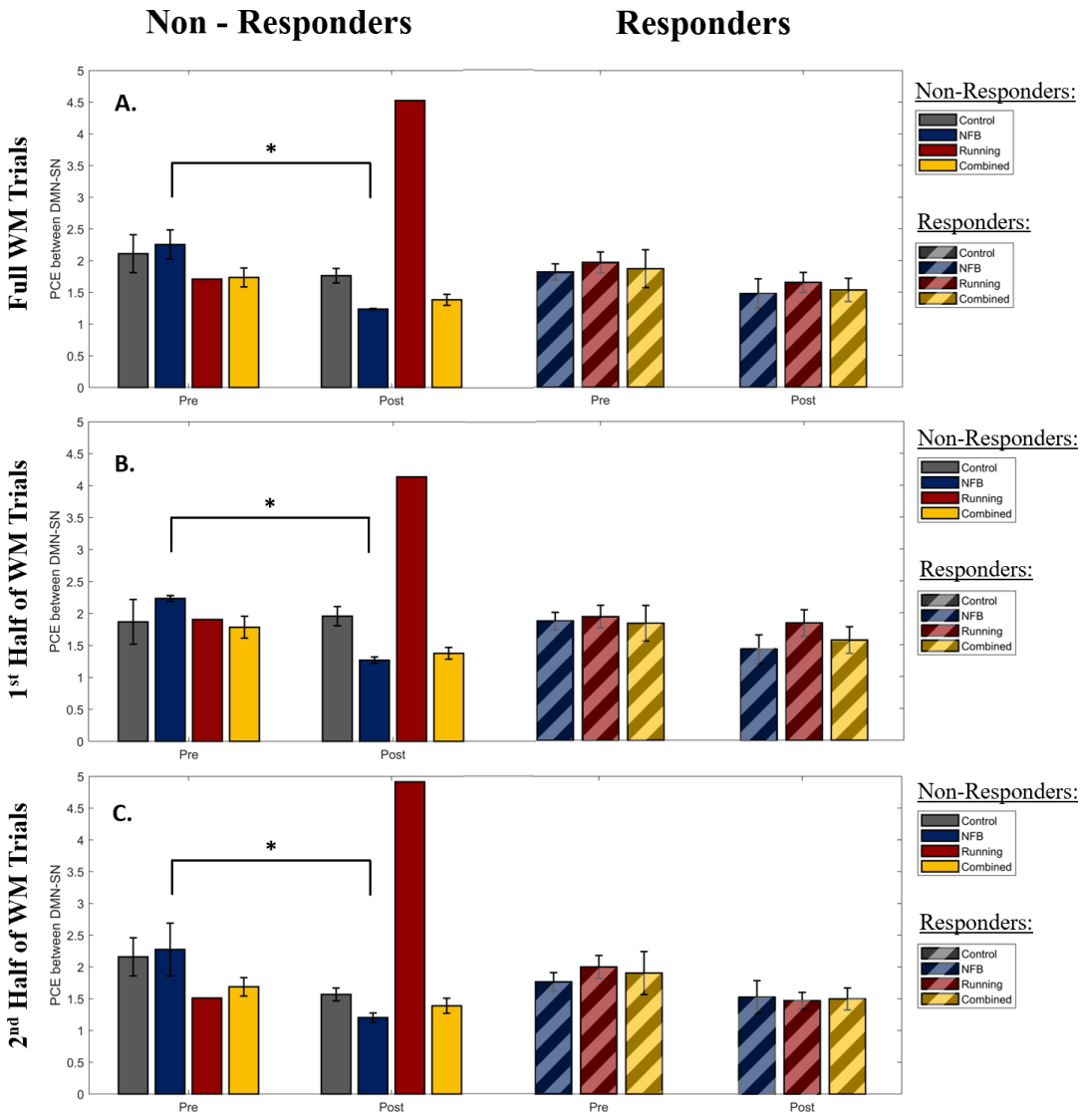

Figure 4: The pairwise cross-entropy (PCE) between the DMN and SN networks, estimated over the **A.** full WM trials, **B.** first half of the WM trials, and the **C.** second half of the WM trials, for the Non-Responders (left column) and Responders (right column). Kindly note the absence of error bars on Non-Responders within the Running group is due to the inclusion of only a single qualifying participant within this subgroup.

responders within the combined group show a lower CEN-SN PCE compared to that of the non-responders, as well as responders in the other groups. Given that there was no main or interaction effect of Session (Pre/Post) on the CEN-SN PCE, this difference was present at the start of the intervention and was not a result of the intervention. This implies that the participants with higher CEN-SN synchronization during executive tasks at baseline, and that participated in aerobic exercise were able to perform better on the mindfulness-like NFB task of staying within the "calm state". Hence aerobic exercise modulated the impact of CEN-SN synchrony on NFB training.

When this pattern was examined across the two halves of the WM trials, the CEN-SN PCE did not show a Group x Responder (calm state) interaction effect during the first half (F=2.819, p=0.067), as seen in Figure 5 B and E, while it did show an even stronger Group x Responder (calm state) interaction effect during the second half of the WM trials (F=3.556,p=0.035), seen in Figure 5 C and F. Therefore, the higher CEN-SN synchronization observed within the responders of the combined group during the WM trials is due to an increase in CEN-SN synchronization during the second half of the WM trials, similar to the healthy pattern of network dynamics observed during WM (Shaw *et al.*, 2021a). This implies that the network dynamics within this group of participants better follows task demands, and could be the reason they responded to the NFB task.

3.2.4 CEN-SN synchronization at baseline predicts NFB response

To further test the link between CEN-SN synchronization and NFB response suggested in section 3.2.3, the PCE between CEN and SN at the pre-intervention time point was used as the predictor variable to classify NFB response (calm state). The resultant 10-fold cross-validated classification accuracy using a linear support vector machine (SVM) was 77.5 ± 12.4%. This implies that a participants' response to future NFB training could be predicted solely based on their present CEN-SN synchronization during a working memory task.

Interestingly, performing this classification task within the NFB group and the combined group separately did not show ($p = 0.72$) better classification accuracy for the combined group (80.0 ± 25.8%), compared to the NFB only group (75.0 ± 35.3%).

4 Discussion

The promising mood and neurobiological benefits of lifestyle interventions such as aerobic exercise and mindfulness training (Shors *et al.*, 2014, 2018) could potentially make them an effective prophylactic measure to prevent mental health deterioration in vulnerable and at-risk populations. However, its widespread adoption is limited due to the requirement for trained meditation and exercise coaches, making them prohibitively expensive for at-risk populations with limited resources such as undergraduate university students. Furthermore, despite evidence on the benefits of combining multiple lifestyle interventions (Shors

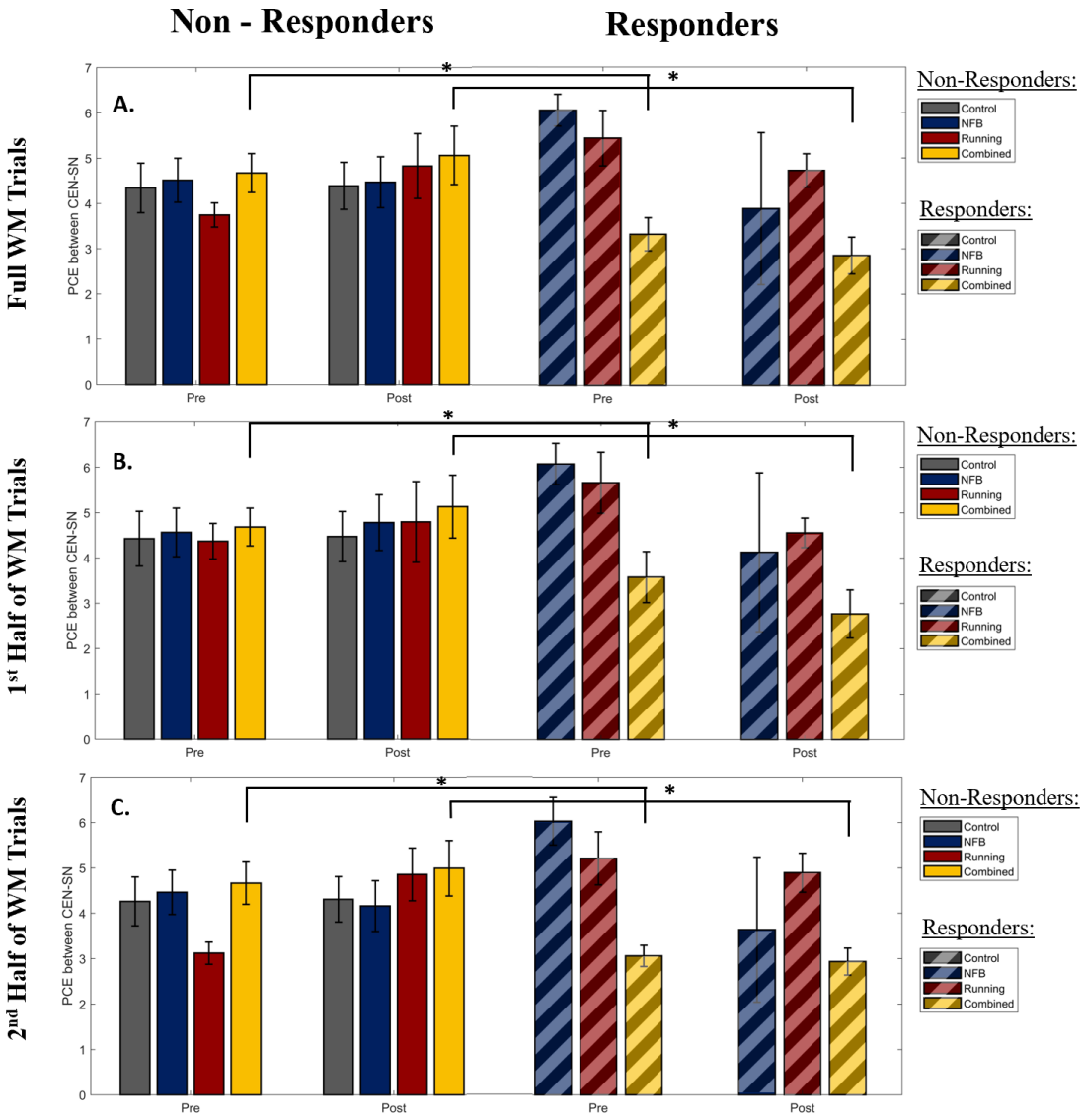

Figure 5: The pairwise cross-entropy (PCE) between the CEN and SN networks, estimated over the **A.** full WM trials, **B.** first half of the WM trials, and the **C.** second half of the WM trials, for the pre-intervention (left column) and post-intervention (right column) time points.

et al., 2014), the impact of combining such lifestyle interventions on task-based dynamics of large-scale brain networks is poorly understood. This study is a first step towards understanding task-related changes in large-scale brain networks as a result of either a running-based aerobic exercise protocol, or a self administered mindfulness-like NFB protocol, or both in combination. This well controlled approach of studying each lifestyle intervention separately, and then in combination allowed us to identify intervention-linked changes in large-scale brain network dynamics that are unique to one intervention, and further showed the added benefit of combining these interventions. Finally, the self-administered nature of the NFB protocol increases the accessibility of such an intervention for populations with limited resources, addressing one of the major barriers for the adoption of such interventions.

Interestingly, non-responders within the NFB group showed task-opposite DMN-SN synchronization during the WM trials. The findings of Shaw *et al.* (2021a) suggest that this could represent an abnormal DMN recruitment during the WM trials, implying that non-responders within the mindfulness-like NFB training group adopted the maladaptive strategy of persistently activating the DMN, even during executive tasks. Further evidence of DMN-SN synchronization during the second half of the WM trials, during which the DMN is expected to be deactivated (Shaw *et al.*, 2021a), suggest that the NFB non-responders might have difficulty suppressing the DMN when needed. Such maladaptive strategies can be particularly problematic since they can result in poor cognitive performance (Anticevic *et al.*, 2012), higher mind-wandering (Xiao *et al.*, 2019) and higher levels of rumination, as seen in psychopathologies such as MDD (Goncalves *et al.*, 2017) and OCD (Hamilton *et al.*, 2015).

However, the results further indicate that such maladaptive strategies linked to NFB non-response can be mitigated by engaging in aerobic exercise in addition to mindfulness-like NFB training. While the exact mechanism of such a synergistic action is yet unknown, the aerobic exercise might act through hippocampal circuits which are known to feed into posterior DMN nodes, such as the posterior cingulate cortex (PCC). These nodes are the primary foci of DMN-linked changes after a short mindfulness training protocol (Xiao *et al.*, 2019), and could be modulated by the hippocampal changes brought about through aerobic exercise. Another mechanism of action could be improvements in CEN dynamics and executive processing often linked with aerobic exercise (Chang *et al.*, 2015). CEN activation can directly result in deactivating the DMN (Chen *et al.*, 2013), resulting in the suppression of any maladaptive mind-wandering or ruminative behaviours.

In fact, the CEN was found to be more synchronized with SN during the task-appropriate WM trials, within responders of the Combined group. This difference was even seen at the pre-intervention timepoint and could predict whether a participant would be able to successfully perform the NFB training with upto 80% accuracy.

Predicting NFB response is a particularly important, albeit challenging task. If possible, such predictions could enable personalized intervention plans for each participant, maximizing the impact of such protocols. While recent literature-wide surveys find that

current predictors are incapable of predicting NFB response in healthy and clinical populations (Weber *et al.*, 2020), this study shows that NFB response can be predicted with the participants' task-based CEN-SN synchrony. These results highlight the importance of novel EEG-based biomarkers (Shaw *et al.*, 2021b), that have great clinical potential in informing personalized therapy plans and improving overall treatment efficacy. The novel analysis pipeline used in this study also allowed the use of the cheaper and more accessible EEG modality to assess ICN synchrony, which would otherwise require more expensive functional imaging modalities such as fMRI. This greatly increases the clinical accessibility of such biomarkers, reducing the barriers to adopt such biomarker-based predictive strategies to treatment planning and personalization.

In sum, this study concludes that the combined effect of aerobic exercise and mindfulness-like NFB is more beneficial for task-linked ICN synchrony than the individual effect of each intervention. Combining these interventions results in more healthy CEN functioning in responders, and prevents maladaptive persistent DMN activation in non-responders.

References

Anticevic, A., Cole, M. W., Murray, J. D., Corlett, P. R., Wang, X. J., and Krystal, J. H. (2012). The role of default network deactivation in cognition and disease. *Trends in Cognitive Sciences*, **16**(12), 584–592.

Baym, C. L., Khan, N. A., Pence, A., Raine, L. B., Hillman, C. H., and Cohen, N. J. (2014). Aerobic Fitness Predicts Relational Memory but Not Item Memory Performance in Healthy Young Adults. *Journal of Cognitive Neuroscience*, **26**(11), 2645–2652.

Brechet, L., Ziegler, D. A., Simon, A. J., Brunet, D., Gazzaley, A., and Michel, C. (2021). Reconfiguration of EEG microstate networks after breath-focused digital meditation training. *Brain Connectivity*.

Campo, J. V. and Bridge, J. A. (2009). Treatment of Youth Depression. *American Journal of Psychiatry*, **166**(9), 958–960.

Chang, Y.-K., Pesce, C., Chiang, Y.-T., Kuo, C.-Y., and Fong, D.-Y. (2015). Antecedent acute cycling exercise affects attention control: an ERP study using attention network test. *Frontiers in Human Neuroscience*, **9**(APRIL), 1–13.

Chen, A. C., Oathes, D. J., Chang, C., Bradley, T., Zhou, Z. W., Williamsa, L. M., Glover, G. H., Deisseroth, K., and Etkin, A. (2013). Causal interactions between fronto-parietal central executive and default-mode networks in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(49), 19944–19949.

Chiesa, A. and Serretti, A. (2010). A systematic review of neurobiological and clinical features of mindfulness meditations. *Psychological Medicine*, **40**(8), 1239–1252.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., McMurdo, M., and Mead, G. E. (2013). Exercise for depression. *Cochrane Database of Systematic Reviews*, **4**(9).

Curlik, D. M. and Shors, T. J. (2013). Training your brain: Do mental and physical (MAP) training enhance cognition through the process of neurogenesis in the hippocampus? *Neuropharmacology*, **64**, 506–514.

Déry, N., Pilgrim, M., Gibala, M., Gillen, J., Martin Wojtowicz, J., MacQueen, G., and Becker, S. (2013). Adult hippocampal neurogenesis reduces memory interference in humans: Opposing effects of aerobic exercise and depression. *Frontiers in Neuroscience*, **7**(7 APR), 1–15.

Doll, A., Hölzel, B. K., Boucard, C. C., Wohlschläger, A. M., and Sorg, C. (2015). Mindfulness is associated with intrinsic functional connectivity between default mode and salience networks. *Frontiers in Human Neuroscience*, **9**, 461.

Epel, E. S., Lin, J., Dhabhar, F. S., Wolkowitz, O. M., Puterman, E., Karan, L., and Blackburn, E. H. (2010). Dynamics of telomerase activity in response to acute psychological stress. *Brain, Behavior, and Immunity*, **24**(4), 531–539.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., Kim, J. S., Heo, S., Alves, H., White, S. M., Wojcicki, T. R., Mailey, E., Vieira, V. J., Martin, S. A., Pence, B. D., Woods, J. A., McAuley, E., and Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences*, **108**(7), 3017–3022.

Escolano, C., Navarro-gil, M., Garcia-campayo, J., Minguez, J., Escolano, C., Navarro-gil, M., Garcia-campayo, J., Congedo, M., and Minguez, J. (2014). The Effects of Individual Upper Alpha Neurofeedback in ADHD : An Open-Label Pilot Study. pages 1–10.

Gallant, S. N. (2016). Mindfulness meditation practice and executive functioning: Breaking down the benefit. *Consciousness and Cognition*, **40**, 116–130.

Galloway, J. (2016). *The Run Walk Run® Method*.

Gao, J., Fan, J., Wu, B. W. Y., Zhang, Z., Chang, C., Hung, Y. S., Fung, P. C. W., and Sik, H. h. (2016). Entrainment of chaotic activities in brain and heart during MBSR mindfulness training. *Neuroscience Letters*, **616**, 218–223.

Goncalves, F., Soares, J. M., Carvalho, S., Leite, J., Ganho-Ávila, A., Fernandes-Goncalves, A., Pocinho, F., Carracedo, A., and Sampaio, A. (2017). Patterns of Default Mode Network Deactivation in Obsessive Compulsive Disorder. *Scientific Reports*, **7**(February), 1–7.

Gray, M. A., Beacher, F. D., Minati, L., Nagai, Y., Kemp, A. H., Harrison, N. A., and Critchley, H. D. (2012). Emotional appraisal is influenced by cardiac afferent information. *Emotion*, **12**(1), 180–191.

Guiney, H. and Machado, L. (2013). Benefits of regular aerobic exercise for executive functioning in healthy populations. *Psychonomic Bulletin & Review*, **20**(1), 73–86.

Hamilton, J. P., Farmer, M., Fogelman, P., and Gotlib, I. H. (2015). Depressive Rumination, the Default-Mode Network, and the Dark Matter of Clinical Neuroscience. *Biological Psychiatry*, **78**(4), 224–230.

Holzel, B., Carmody, J., Vangel, M., Congleton, C., Yerramsetti, S., Gard, T., and Lazar, S. (2011). Meditation-Gray Matter. *Psychiatry Res.*, **191**(1), 36–43.

Jacobs, T. L., Epel, E. S., Lin, J., Blackburn, E. H., Wolkowitz, O. M., Bridwell, D. A., Zanesco, A. P., Aichele, S. R., Sahdra, B. K., MacLean, K. A., King, B. G., Shaver,

P. R., Rosenberg, E. L., Ferrer, E., Wallace, B. A., and Saron, C. D. (2011). Intensive meditation training, immune cell telomerase activity, and psychological mediators. *Psychoneuroendocrinology*, **36**(5), 664–681.

Keating, L. E., Becker, S., McCabe, K., Whattam, J., Garrick, L., Sassi, R. B., Frey, B. N., and McKinnon, M. C. (2018). Effects of a 12-week running programme in youth and adults with complex mood disorders. *BMJ Open Sport and Exercise Medicine*, **4**(1), 1–7.

Kilpatrick, L. A., Suyenobu, B. Y., Smith, S. R., Bueller, J. A., Goodman, T., Creswell, J. D., Tillisch, K., Mayer, E. A., and Naliboff, B. D. (2011). Impact of mindfulness-based stress reduction training on intrinsic brain connectivity. *NeuroImage*, **56**(1), 290–298.

King, A. P., Block, S. R., Sripada, R. K., Rauch, S., Giardino, N., Favorite, T., Angstadt, M., Kessler, D., Welsh, R., and Liberzon, I. (2016). Altered default mode network (DMN) resting state functional connectivity following a mindfulness-based exposure therapy for posttraumatic stress disorder (PTSD) in combat veterans of Afghanistan and Iraq. *Depression and Anxiety*, **33**(4), 289–299.

Kluetsch, R. C., Ros, T., Théberge, J., Frewen, P. A., Calhoun, V. D., Schmahl, C., Jetly, R., and Lanius, R. A. (2014). Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback. *Acta Psychiatrica Scandinavica*, **130**(2), 123–136.

Kozasa, E. H., Sato, J. R., Lacerda, S. S., Barreiros, M. A., Radvany, J., Russell, T. A., Sanches, L. G., Mello, L. E., and Amaro, E. (2012). Meditation training increases brain efficiency in an attention task. *NeuroImage*, **59**(1), 745–749.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences*, **15**(10), 483–506.

Mental Health Commission Of Canada (2013). Making the case for investing in mental health in Canada. Technical report.

Monti, J. M., Hillman, C. H., and Cohen, N. J. (2012). Aerobic fitness enhances relational memory in preadolescent children: The FITKids randomized control trial. *Hippocampus*, **22**(9), 1876–1882.

Prakash, R. S., Voss, M. W., Erickson, K. I., Lewis, J. M., Chaddock, L., Malkowski, E., Alves, H., Kim, J., Szabo, A., White, S. M., Wójcicki, T. R., Klamm, E. L., McAuley, E., and Kramer, A. F. (2011). Cardiorespiratory Fitness and Attentional Control in the Aging Brain. *Frontiers in Human Neuroscience*, **4**(January), 1–12.

Ros, T., Théberge, J., Frewen, P. A., Kluetsch, R., Densmore, M., Calhoun, V. D., and Lanius, R. A. (2013). Mind over chatter: Plastic up-regulation of the fMRI salience network directly after EEG neurofeedback. *NeuroImage*, **65**, 324–335.

Schwarb, H., Johnson, C. L., Daugherty, A. M., Hillman, C. H., Kramer, A. F., Cohen, N. J., and Barbey, A. K. (2017). Aerobic fitness, hippocampal viscoelasticity, and relational memory performance. *NeuroImage*, **153**(March), 179–188.

Shaw, S. B., McKinnon, M. C., Heisz, J., and Becker, S. (2021a). Dynamic task-linked switching between brain networks – A tri-network perspective. *Brain and Cognition*, **151**(April), 105725.

Shaw, S. B., Mckinnon, M. C., Heisz, J. J., Harrison, A. H., Connolly, J. F., and Becker, S. (2021b). Tracking the Brain's Intrinsic Connectivity Networks in EEG. *bioRxiv*.

Shors, T. J., Olson, R. L., Bates, M. E., Selby, E. A., and Alderman, B. L. (2014). Mental and Physical (MAP) Training: A neurogenesis-inspired intervention that enhances health in humans. *Neurobiology of Learning and Memory*, **115**, 3–9.

Shors, T. J., Chang, H. Y., and Millon, E. M. (2018). MAP Training My Brain™: Meditation plus aerobic exercise lessens trauma of sexual violence more than either activity alone. *Frontiers in Neuroscience*, **12**(APR), 1–11.

Sitaram, R., Ros, T., Stoeckel, L. E., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M., Rana, M., Oblak, E., Birbaumer, N., and Sulzer, J. (2016). Closed-loop brain training: the science of neurofeedback. *Nature Neuroscience*.

Talukdar, T., Nikolaidis, A., Zwilling, C. E., Paul, E. J., Hillman, C. H., Cohen, N. J., Kramer, A. F., and Barbey, A. K. (2018). Aerobic fitness explains individual differences in the functional brain connectome of healthy young adults. *Cerebral Cortex*, **28**(10), 3600–3609.

Thibault, R. T., Lifshitz, M., and Raz, A. (2016). The self-regulating brain and neurofeedback: Experimental science and clinical promise. *Cortex*, **74**, 247–261.

Wang, Y., Bernanke, J., Peterson, B. S., McGrath, P., Stewart, J., Chen, Y., Lee, S., Wall, M., Bastidas, V., Hong, S., Rutherford, B. R., Hellerstein, D. J., and Posner, J. (2019). The association between antidepressant treatment and brain connectivity in two double-blind, placebo-controlled clinical trials: a treatment mechanism study. *The Lancet Psychiatry*, **6**(8), 667–674.

Weber, L. A., Ethofer, T., and Ehlis, A. C. (2020). Predictors of neurofeedback training outcome: A systematic review. *NeuroImage: Clinical*, **27**(April), 102301.

Xiao, Q., Zhao, X., Bi, G., Wu, L., Zhang, H., Liu, R., Zhong, J., Wu, S., Zeng, Y., Cui, L., Chen, Y., Wu, K., and Chen, Z. (2019). Alterations of Regional Homogeneity and Functional Connectivity Following Short-Term Mindfulness Meditation in Healthy Volunteers. *Frontiers in Human Neuroscience*, **13**(October), 1–12.

7.3 Discussion

The results of the study presented in this chapter represent a first step toward the validation of this novel EEG-based ICN tool, showing its potential significance in tracking changes in ICN dynamics through the course of a clinical treatment.

Although the study presented here was performed in healthy undergraduate students, it should be noted that about a third of our participants had extremely high stress and depression scores. Thus, our results represent a promising proof of concept, and suggest that a similar analysis protocol could be applicable in tracking the progress of clinical patients undergoing therapy.

The results in this study also highlight the utility of examining the dynamic interactions within and between nodes of the three core ICNs in predicting treatment response. More specifically, improved co-activation of the CEN and SN was observed in participants who had a higher success rate of performing the neurofeedback task. This finding further validates the utility of such ICN biomarkers in predicting the participants that might respond to a given therapy, allowing the clinician to better customize each patient's treatment plan.

Bibliography

Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., and McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology*, 6(1):27313.

Chapter 8

General Discussion & Conclusion

8.1 Contribution to detecting ICN activation using EEG

The primary goal of this thesis was to develop and assess an EEG-based method of detecting ICN activity and dynamics. Chapter 4 found that the currently popular method of doing this using EEG microstates made assumptions that were incompatible with known ICN dynamics, requiring the development of a novel method. Most notably, EEG microstates required only one microstate pattern to be active at any given time. However, this is in stark contrast to the co-activation of ICNs observed in many scenarios, and even in chapter 5 of this thesis.

Hence, an alternate methodology was developed in chapter 6 by directly learning EEG signatures of the three major ICNs in the tri-network model, using simultaneously collected EEG-fMRI data. This unique process developed a machine learning model that can predict ICN activation using EEG alone, accomplishing the primary

goal of this thesis. This model is freely available for download online, along with the analysis pipeline used to develop it.

Chapter 7 tested the methodology developed in chapter 6 to identify DMN, CEN and SN dynamics using EEG data alone, assessing intervention-linked changes in these networks. Given the success of this methodology in detecting such changes in ICN dynamics, we can conclude that this thesis has created and validated a novel methodology to detect ICN activation using EEG alone.

This tangible outcome has immense implications for anyone that needs to study ICNs, but cannot use an fMRI system. In particular, this will reduce the barriers for the clinical adoption of such ICN-based biomarkers, especially in planning and tracking treatment response of patients with mental health disorders.

8.2 Contribution to understanding mechanisms of network dynamics

Starting from connected neurons performing synchronous tasks, to interconnected large-scale brain regions forming intrinsically connected brain networks (ICNs), the human brain shows a hierachical architecture of interconnected functional modules. ICN-based analysis represents one of the highest levels of this hierarchy and can therefore be readily linked to complex cognitive processes and behaviours. On the other hand, the mechanisms underlying such ICN behaviour are much more complex, and consequently harder to understand in the context of the behaviour and cognition.

This thesis undertook this difficult task of elucidating the mechanisms underlying task-linked ICN dynamics within the tri-network model in a healthy population.

Chapter 5 provided the first direct evidence of a major prediction of the tri-network model, suggesting that the SN gates the activation of the DMN and CEN based on task demands by co-activating with the task-relevant network. Such coherent activation for the purpose of selecting specific neural populations has been previously observed in attentional control of incoming stimuli (Fries, 2015). Chapter 5 suggests a similar "connection through coherence" mechanism of SN's gating behaviour.

The results of chapter 7 add to this newfound understanding by discovering that increased CEN-SN synchrony during cognitive tasks is predictive of future neurofeedback (NFB) performance. Chapter 7 further shows that participants can develop maladaptive ICN dynamics if they were unable to effectively perform the NFB training, and that additionally engaging in aerobic exercise mitigated the development of such dysfunctional ICN dynamics.

These insights are critical if and when translating this research into clinical practice, leading to the development of ICN dynamics informed treatment plans.

8.3 Contribution to reframing our current understanding of EEG microstates

Given the wide-spread clinical adoption of EEG microstates in characterizing changes in network dynamics in patients with numerous psychopathologies, and its use as a neurofeedback target, it is critical to understand the validity of all assumptions underlying this analysis method.

Chapter 4 of the thesis first showed that all three assumptions of EEG microstates does not hold up to scrutiny, suggesting that the discontinuous, "winner-takes-all"

framework of EEG microstates is not an appropriate descriptor of neural dynamics. The results of chapter 3 suggested that spatiotemporal patterns of EEG are much more continuous in nature, leading to a revised view of EEG microstates.

Follow-up work based on this paper has shown that this continuous behaviour is dominant in low amplitude periods of the global field potential (GFP), representing periods of asynchronous neural activity, with brief periods of synchronized neural activity (high GFP) (Mishra et al., 2020). Further work citing the research in chapter 4 showed that the periodic patterns of EEG scalp topology previously thought to represent attractor-like dynamics are in fact a by product of rotating phase patterns of resting-state alpha oscillations (von Wegner et al., 2020).

Interestingly, this revised understanding of EEG microstates reinforces the origin of EEG microstates-based analysis, which was developed to understand alpha patterns of resting-state EEG during a period where source-localization methods were not fully developed. EEG microstates was developed as a workaround to interpret underlying dipoles using rotating patterns of EEG scalp topology (Lehmann, 1971; Lehmann and Skrandies, 1980; Lehmann, 1990).

8.4 Future work

An obvious next step is to clinically validate the developed model by testing it in a cohort of patients with mental health disorders. An ideal study design would include an intervention, similar to that described in chapter 7, to assess any changes in ICN dynamics post-intervention, that are associated with an improvement in symptoms.

These changes in network dynamics post-intervention can also help create a novel

neurofeedback target that can directly re-normalize the three networks in the tri-network model, allowing for the development of novel individualized treatments to neuropathologies. The EEG-based tool modeled from the identified EEG signatures will also provide an objective treatment response prediction tool that can greatly improve treatment planning and treatment response.

Patients with post traumatic stress disorder (PTSD) would be an ideal target for this future work, given the wide range of tri-network ICN nodes implicated in the cognitive dysfunctions observed in this psychopathology.

PTSD patients are found to show abnormally low connectivity within DMN nodes at rest (Bluhm et al., 2009), associated with poor autobiographical memory recall and emotion processing (Nicholson et al., 2018). Additionally, the DMN is abnormally recruited, instead of the CEN, during executive task performance in PTSD patients (Daniels et al., 2010), potentially explaining the cognitive and emotional dysregulation seen in PTSD symptomatology (Lanius et al., 2015). This abnormal recruitment can be reduced by neuromodulatory interventions such as neurofeedback (Nicholson et al., 2018) that decrease the level of activity in the amygdala. The DMN is also found to be fractionated in PTSD, with a relatively spared posterior DMN subnetwork (also called community) (Akiki et al., 2018). Neuromodulation of the intact posterior DMN subnetwork could be used to access the other dysfunctional ICNs in the brain, and could be the reason NFB down-regulation of alpha power (oscillations within the 8-12 Hz range) at the Pz electrode, located directly above the precuneus is found to be beneficial in PTSD patients (Nicholson et al., 2020), and can modulate connectivity of the SN (Ros et al., 2013; Kluetsch et al., 2014).

Finally, the utility of the developed model extends beyond characterizing and

tracking PTSD, and has much broader potential to investigate ICN dynamics in a wide range of disorders affecting mental and brain health.

Bibliography

Akiki, T. J., Averill, C. L., Wrocklage, K. M., Scott, J. C., Averill, L. A., Schweinsburg, B., Alexander-Bloch, A., Martini, B., Southwick, S. M., Krystal, J. H., and Abdallah, C. G. (2018). Default mode network abnormalities in posttraumatic stress disorder: A novel network-restricted topology approach. *NeuroImage*, 176(February):489–498.

Bluhm, R. L., Williamson, P. C., Osuch, E. A., Frewen, P. A., Stevens, T. K., Boksman, K., Neufeld, R. W. J., Théberge, J., and Lanius, R. A. (2009). Alterations in default network connectivity in posttraumatic stress disorder related to early-life trauma. *Journal of psychiatry & neuroscience : JPN*, 34(3):187–94.

Daniels, J. K., McFarlane, A. C., Bluhm, R. L., Moores, K. A., Clark, C. R., Shaw, M. E., Williamson, P. C., Densmore, M., and Lanius, R. A. (2010). Switching between executive and default mode networks in posttraumatic stress disorder: alterations in functional connectivity. *Journal of psychiatry & neuroscience : JPN*, 35(4):258–66.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88(1):220–235.

Kluetsch, R. C., Ros, T., Théberge, J., Frewen, P. A., Calhoun, V. D., Schmahl,

C., Jetly, R., and Lanius, R. A. (2014). Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback. *Acta Psychiatrica Scandinavica*, 130(2):123–136.

Lanius, R. A., Frewen, P. A., Tursich, M., Jetly, R., and McKinnon, M. C. (2015). Restoring large-scale brain networks in PTSD and related disorders: a proposal for neuroscientifically-informed treatment interventions. *European Journal of Psychotraumatology*, 6(1):27313.

Lehmann, D. (1971). Multichannel topography of human alpha EEG fields. *Electroencephalography and Clinical Neurophysiology*, 31(5):439–449.

Lehmann, D. (1990). Brain Electric Microstates and Cognition: The Atoms of Thought. In *Machinery of the Mind*, pages 209–224. Birkhäuser Boston, Boston, MA.

Lehmann, D. and Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, 48(6):609–621.

Mishra, A., Englitz, B., and Cohen, M. X. (2020). EEG microstates as a continuous phenomenon. *NeuroImage*, 208:116454.

Nicholson, A. A., Rabellino, D., Densmore, M., Frewen, P. A., Paret, C., Kluetsch, R., Schmahl, C., Théberge, J., Ros, T., Neufeld, R. W. J., McKinnon, M. C., Reiss, J. P., Jetly, R., and Lanius, R. A. (2018). Intrinsic connectivity network dynamics in PTSD during amygdala downregulation. *Human Brain Mapping*, (July).

Nicholson, A. A., Ros, T., Densmore, M., Frewen, P. A., Neufeld, R. W., Théberge, J., Jetly, R., and Lanius, R. A. (2020). A randomized, controlled trial of alpha-rhythm EEG neurofeedback in posttraumatic stress disorder: A preliminary investigation showing evidence of decreased PTSD symptoms and restored default mode and salience network connectivity using fMRI. *NeuroImage: Clinical*, 28.

Ros, T., Théberge, J., Frewen, P. A., Kluetsch, R., Densmore, M., Calhoun, V. D., and Lanius, R. A. (2013). Mind over chatter: Plastic up-regulation of the fMRI salience network directly after EEG neurofeedback. *NeuroImage*, 65:324–335.

von Wegner, F., Bauer, S., Rosenow, F., Triesch, J., and Laufs, H. (2020). EEG microstate periodicity explained by rotating phase patterns of resting-state alpha oscillations. *NeuroImage*, page 117372.