Stepping Beyond Behaviour: Explainable Machine Learning for Clinical Neurophysiological Assessment of Concussion Progression

STEPPING BEYOND BEHAVIOUR: EXPLAINABLE MACHINE LEARNING FOR CLINICAL NEUROPHYSIOLOGICAL ASSESSMENT OF CONCUSSION PROGRESSION

BY

ROBER BOSHRA, M.Sc., B.Sc.

A THESIS

SUBMITTED TO THE SCHOOL OF BIOMEDICAL ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2019) McMaster University

(School of Biomedical Engineering) Hamilton, Ontario, Canada

TITLE:	Stepping Beyond Behaviour: Explainable Machine Learning for Clinical Neurophysiological Assessment of Concussion Progression
AUTHOR:	Rober Boshra M.Sc. (Neuroscience), McMaster University, Canada B.Sc. (Computer Science), Dalhousie University, Canada
SUPERVISOR:	Dr. John F. Connolly & Dr. James P. Reilly
NUMBER OF PAGES:	xix, 169

To Mom & Dad

Abstract

The present dissertation details a sequence of studies in mild traumatic brain injury, the progression of its effects on the human brain as recorded by event-related electroencephalography, and potential applications of machine learning algorithms in detecting such effects. The work investigated data collected from two populations (in addition to healthy controls): 1) a recently-concussed adolescent group, and 2) a group of retired football athletes who sustained head trauma a number of years prior to testing. Neurophysiological effects of concussion were assessed across both groups with the same experimental design using a multi-deviant auditory oddball paradigm designed to elicit the P300 and other earlier components. Explainable machine learning models were trained to classify concussed individuals from healthy controls. Cross-validation performance accuracies on the recently-concussed (chapter 4) and retired athletes (chapter 3) were ≈80% and ≈85%, respectively. Features showed to be most useful in the two studies were different, motivating a study of potential differences between the different injury-stage/age groups (chapter 5). Results showed event-related functional connectivity to modulate differentially between the two groups compared to healthy controls. Leveraging results from the presented work a theoretical model of mild traumatic brain injury progression was proposed to form a framework for synthesizing hypotheses in future research.

Acknowledgements

First and foremost, I would like to thank my two phenomenal supervisors: Dr. John Connolly and Dr. James Reilly. Their never-ending encouragement, advice, and mentorship made this thesis possible. This PhD journey has had its fair share of ups and downs; they were always present to stabilize it when direly needed and to celebrate when a drink was in order. For their generosity and guidance, I am forever indebted. I would also like to thank Dr. Ranil Sonnadara for generously accepting to be on my supervisory committee and for his valuable input and guidance throughout my degree. While not directly part of the work presented in this thesis, I would also like to extend my gratitude to three mentors whose passion for science and knowledge has helped shape the researcher I am today: Drs. Suzanna Becker, Nauzer Kalyaniwalla, and Thomas Trappenberg.

I consider myself most fortunate to have been in a nurturing environment with thriving multidisciplinary collaborators. The manner by which these individuals approached teamwork and collaboration is one I have taken every chance to learn and adapt for my own. I count our collaborations amongst my dearest of achievements. I would like to acknowledge and thank this brilliant group of supportive colleagues and collaborators for each and every glimpse they gave me of their fascinating disciplines and unique perspectives. For that and much more, I extend my thanks to Kyle Ruiter, Kiret Dhindsa, Richard Kolesar, Adianes Herrera-Diaz, Gaisha Oralova, Daniel Schmidtke, Stefanie Blain-Moraes, Richard Mah, Victor Kuperman, Aki Kyröläinen, Amanda Ho, Nathalee Ewers, Nitish Dhingra, and Omar Boursalie.

Throughout my years at McMaster University, I have been lucky to get to know a number of great minds and wonderful friends in the Language Brain and Memory Lab, the Linguistics and Languages department, and the the centre for Advanced Research in Experimental and Applied Linguistics (ARiEAL). Thank you to all the fellow students/fellows and the great faculty, namely: Chelsea Whitwell, Narcisse Torshizi, Richard Mah, Natalia Lapinskaya, Constance Imbault, Kelly Nisbet, Amanda Ho, Richard Kolesar, Adianes Herrera-Diaz, Nathalee Ewers, Kyle Ruiter, Jitka Bartosova, Gaisha Oralova, Fareeha Rana, Cassandra Chapman, Laura Beaudin, Aki Kyröläinen, Zoe Walchli, Daniel Schmidtke, Melda Coskun, Davide Gentile, Victor Kuperman, Elizabet Service, and Daniel Pape. I would also like to especially thank all of the VoxNeuro team for a truly unique experience and for their incredible support. As this chapter draws to a close and another begins, I thank you all for an enriching, warm, and fulfilling time.

I would like to dearly thank Chia-Yu Lin for being a phenomenal manager at the ARiEAL research centre. Thank you for always having the students' best interests at heart and for genuinely cultivating an enriching environment of research and discovery. I would also like to recognize Jane Mah for her help, patience, and support despite all the numbing questions I sent her way.

I have been privileged to receive financial support that made this PhD process substantially less strenuous than it could have been. For that I thank the Vector Institute, the MacData Institute, the Ontario Centres of Excellence, the Ontario

Ministry of Research and Innovation, the Natural Sciences and Engineering Research Council, and my two supervisors. This work was partially supported by an operating grant funded by the Canadian Institutes of Health Research (CIHR) and by the Hamilton Spectator. I would like to dearly thank Steve Buist for his instrumental role in making the collaboration with the Hamilton Spectator a reality.

The years leading to this PhD were anything but straightforward. Surviving the turbulent times was made possible only by the support and encouragement of a large group of amazing individuals that I don't have enough space or the eloquence to adequately thank. My gratitude goes to Jitka, Kyle, Chelsea, Amanda, Chia-Yu, Narcisse, and Roksana. I would also like to extend my gratitude to the amazing friends and family from my first home, Egypt. Even after 9 years away, it is always the occasional trip home that, time and time again, has kept me grounded when I needed it the most. I can't thank my wife, Mira, enough for her help and support. Your patience, love, and understanding even during my brushes with insanity have been humbling. I would like to thank my sister Carol for her continuous support and the never-ending banter. To my parents, whom I dedicate this dissertation to, I owe so much. Their sacrifices, humility, and perseverance have been a source of inspiration and a goal that I will always vie to achieve.

Notation and abbreviations

CNN: Convolutional Neural Network

DAI: Diffuse Axonal Injury

DL: Deep Learning

EEG: Electroencephalography

ERP: Event-Related Potentials

fMRI: Functional Magnetic Resonance Imaging

ICA: Independent Component Analysis

ML: Machine Learning

mTBI: Mild Traumatic Brain Injury

MRI: Magnetic Resonance Imaging

NN: Neural Network

RS: Resting State

SHAP: SHapley Additive exPlanations

SVM: Support Vector Machine

TBI: Traumatic Brain Injury

Contents

List of Figures

Declaration of Academic Achievements

The present document constitutes a "sandwich" thesis as defined by the School of Graduate Studies, McMaster University. As such, in addition to an introduction and conclusion, it includes three self-contained chapters with reformatted articles with significant contributions made by the author that have been either accepted, submitted, or in preparation for submission to peer-reviewed journals. Moreover, a pilot study is presented in chapter 2 motivating several key goals of the dissertation. The following outlines each of the three main studies forming chapters 3, 4, and 5.

Chapter 3 is a reprint of an article accepted at the IEEE Transactions on Neural Systems and Rehabilitation Engineering as "**Boshra, R.**, Dhindsa K. Boursallie, O., Ruiter, K. I., Sonnadara, R., Doyle, T., Samavi, R., Reilly, J. P., & Connolly, J. F. From Group-Level Statistics to Single-Subject Prediction: Machine Learning Detection of Concussion in Retired Athletes."

RB, JFC, and JPR conceived of the machine learning study. **RB** and KIR collected the EEG/ERP data. **RB**, JPR, KD, and OB developed and refined the machine learning design, analysis, and validation. **RB**, KD, and OB prepared the first draft. All authors edited the manuscript.

Chapter 4 has been submitted to Scientific Reports as "**Boshra, R.**, Ruiter, K. I., DeMatteo, C., Reilly, J. P., & Connolly, J. F. Post-acute Neurophysiological Correlates of Concussion: Deep Learning for Clinical Assessment."

RB conceived of the neural network design. **RB**, JFC, and JPR formulated the machine learning study design and question. **RB** and KIR collected the EEG/ERP data. **RB** implemented, processed, and analyzed the study's models and their performance results. **RB** prepared the first draft. All authors edited the manuscript.

Chapter 5 is a manuscript prepared for submission in the Nature Neuroscience as "**Boshra, R.**, Ruiter, K. I., Dhindsa K., Sonnadara R., Reilly, J. P., & Connolly, J. F. On the Time-Course of Functional Brain Connectivity: Theory of the Dynamic Progression of Concussion Effects."

RB conceived of the functional connectivity hypotheses based on findings from work detailed in the earlier chapters and other work detailed in the *Additional Achievements* section below. **RB**, JFC, KIR, and KD formulated the brain injury model presented in the chapter. **RB** and KIR collected the EEG/ERP data. **RB** implemented, processed, and statistically analyzed the connectivity data. **RB** and KD refined the statistical analysis procedure. **RB** and KD wrote the first draft. All authors edited the manuscript.

Additional Achievements

In addition to what is presented in the chapters, the author was a primary contributor on three other studies closely associated with the theme of this thesis. The first of these studies has been published as:

- Ruiter, K. I., **Boshra, R.**, Doughty, M., Noseworthy, M., & Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. Clinical Neurophysiology, 130(1), 111–121. https://doi.org/10.1016/j.clinph.2018.10.013

The second has been submitted on 2019-05-01 to Brain Research as:

- Ruiter, K. I., **Boshra, R.**, DeMatteo, C., Noseworthy, M., & Connolly, J. F. Neurophysiological markers of cognitive deficits and recovery in concussed adolescents.

The third is in preparation for submission to Clinical Neurophysiology as:

- Ruiter, K. I., **Boshra, R.**, DeMatteo, C., Noseworthy, M., & Connolly, J. F. Tracking concussion recovery in adolescents using neurophysiological markers: An ERP Study.

Three additional manuscripts have been published, but were not included in the dissertation for thematic reasons. These are listed below:

- Blain-Moraes, S., **Boshra, R.**, Ma, H. K., Mah, R., Ruiter, K., Avidan, M., & Mashour, G. A. (2016). Normal Brain Response to Propofol in Advance of Recovery from Unresponsive Wakefulness Syndrome. Frontiers in Human Neuroscience, 10(June), 1–6. https://doi.org/10.3389/fnhum.2016.00248

- Ho, A., **Boshra, R.**, Schmidtke, D., Oralova, G., Moro, A. L., Service, E., & Connolly, J. F. (2019). Electrophysiological Evidence for the Integral Nature of Tone in Mandarin Spoken Word Recognition. Neuropsychologia.

- Connolly, J. F., Reilly, J. P., Fox-Robichaud A., Britz, P., Blain-Moraes, S., Sonnadara, R., Hamielec, C., Herrera-Díaz, A., & **Boshra, R.** (2019). Development of a Point of Care System for Automated Coma Prognosis: A Prospective Cohort Study Protocol. BMJ Open.

CHAPTER 1 Introduction and Background

The brain is the most complex organ in the human body. Its intertwined networks of neurons give rise to the spectacular spectrum of human abilities from fine motor control to logical reasoning and creative thought. Neuroscience, an entire domain of science dedicated to the study of the nervous system and its main hub, continues to explore the brain and its ailments at varying sizes of granularity. While the brain's anatomy and structure are extremely complex, it is of the functioning brain that our understanding begins to pale in comparison with the breadth of the unknown. These gaps in the current knowledge form massive barriers to developing clinical interventions to resolve the direly expansive brain-related disorders and conditions.

The present dissertation is an investigation of the phenomenon termed *mild traumatic brain injury* (mTBI) in the literature. Particularly, the work builds on a substantial body of work describing an unrealized utility of electrophysiology in clinical standards of mTBI identification, management, and tracking. Utilizing a comprehensive tool-set of machine learning techniques, the presented work extends previously identified neurophysiological markers of mTBI to prototype a single-subject tool for mTBI identification. In doing so, multiple facets of machine learning are explored in the context of mTBI and electrophysiology. In turn, that gave rise to a realized theoretical model supported by a multitude of modalities that is intended to act as a framework for formulating future lines of research that aim to better our understanding of the phenomenon.

This chapter acts as the necessary introduction and background of topics and themes discussed in the dissertation. A total of five sections follow, containing, in order, a high-level modern view of brain injury and mTBI as one of its subcategories; the current clinical standards of mTBI management as well as gaps in the literature; a brief introduction to electrophysiology, event-related potentials, and their utility in clinical practice; an introduction to machine learning and its potential in EEG and other clinical applications; and lastly, a detailed description of the present dissertation's primary hypotheses and an outline of the following chapters.

1.1 Severe Implications of the Mild

The brain's exposure to trauma is one ailment that upsets normally functioning mechanisms, causing a wide variety of sequelae. Contact forces and biomechanics of traumatic impact have been seen to cause both focal and more diffuse effects (Raghupathi, 2004). Diffuse axonal injury (DAI) is the prime mechanism of brain-wide affliction after trauma and is seen in 40% to 50% of injuries requiring hospital admissions (Meythaler et al., 2001). DAI describes a brain injury understood to be caused by rotational forces rather than direct contact to the head (Raghupathi, 2004). DAI-related loss of membrane integrity and axonal degeneration in the white-matter is on the microscopic level and is commonly unobservable using computed tomography (CT) or magnetic resonance imaging (MRI). Severe DAI has been associated with coma, postulating the injury to impact tracts to the hypothalamus and the pituitary stalk (Meythaler et al., 2001). In fact, any form of traumatic brain injury (TBI) causing unconsciousness is believed to co-occur with DAI (Meythaler et al., 2001). TBI is a wide-spread affliction with 2.8 million annually reported cases in the United States, as documented by the Centres for Disease Control and Prevention (CDC; Taylor et al., 2017). Despite its impact and the continuous research into its mechanisms, TBI remains a burden on both the patient and the clinical system. In the United states alone, an estimated 26,000 deaths occur annually due to DAI, with surviving patients suffering a wide range of functional and neurobehavioural losses causing upwards of 25$ billion annually (Meythaler et al., 2001).

The term TBI typically encompasses all forms of brain trauma covering both severe and milder impacts. Mild TBI (mTBI) has been gaining significant traction and exposure in the last few decades. This subset of TBI cases are identified as having closed-head injuries with no detectable brain lesioning or hemorrhaging. Despite the afflicted suffering from a multitude of symptoms and a common deterioration in cognitive ability, mTBI's effects were often seen as transient, with symptoms commonly subsiding less than a month after insult (McCrory et al., 2013, 2017). However, recent work on post-mortem professional athletes in high-impact sports showed the critical link between mTBI and what is termed Chronic Traumatic Encephalopathy (CTE; Omalu et al., 2005). Although the work remains polarizing, it has since been explored in a larger population with replication linking repeated head injuries to the neurodegeneration observable in CTE (McKee et al., 2009; Gavett et al., 2011). Primarily, work into direct links of CTE is restricted to post-mortem autopsies of the brain, inhibiting both the intervention to help the afflicted and the ability to monitor CTE's progression. Consequently, CTE and the primary drivers into its development remain an open question; however, CTE presents perhaps the most ubiquitous example of mTBI's non-mild consequences.

Anatomical aberrations due to mTBI convey the severity of the condition. However, it is the functional consequences that mark the condition's uniqueness. Following concussion (henceforth used interchangeably with mTBI; see McCrory et al., 2013), several reports presented a spectrum of neuropsychological and cognitive abnormalities observable through many modalities of assessments (c.f., McAllister et al., 2001; De Beaumont et al., 2007; Heitger et al., 2009; Baillargeon et al., 2012). Work has shown concussion to cause a variety of symptoms directly after insult, including headaches, fogginess, irritability, and insomnia (McCrory et al., 2013). Additionally, cognitive impairments are typically observable. Studies have reported individuals with concussion to have afflicted reaction times, memory, executive functioning, attention, and inhibition (McCrory et al., 2013, 2017; Broglio et al., 2017; De Beaumont et al., 2013). Notably, the literature reports 80%-90% of concussion symptoms to resolve briefly after injury – typically 7-10 days, longer for children and adolescents – where persisting symptoms are identified as developing Post Concussion Syndrome (PCS; McCrory et al., 2013; Gaetz et al., 2000). Despite the apparent transient nature of behaviourally manifesting sequelae, many studies have showed long-lasting alterations in neurophysiological responses in asymptomatic patients (e.g., Broglio et al., 2011; Ruiter et al., 2019; De Beaumont et al., 2007). These results were significantly emphasized by a resurgence of symptoms in previously concussed individuals later in life (Tremblay et al., 2013; Ruiter et al., 2019).

A common theory has been proposed to try and explain the return of symptoms in individuals with concussion history, as well as the discrepancy often found between behavioural assessments and neurological measurements such as electrophysiological measures (Ruiter et al., 2019; De Beaumont et al., 2012, 2009) and functional hemodynamics (McAllister et al., 2001; Hocke et al., 2018). It is now viewed that the brain is able to allocate additional resources after injury to try and accomplish complex tasks with performance not dissimilar to healthy controls. Further, the notion of a cognitive reserve (see Stern, 2009; Kesler et al., 2003) has been hypothesized to explain the overall neurocognitive decline of previously concussed individuals through aging (De Beaumont et al., 2009, 2012). According to cognitive reserve, a previously-injured brain loses the ability to sustain its compensatory mechanisms with age, resulting in an abnormal aging process with a resurgence of symptoms and other cognitive deficits (De Beaumont et al., 2009, 2012). While there is broad consensus that brain function is altered following concussion, there has been little work clarifying the progression of post-concussive effects throughout aging, the relationship between those effects to observable symptomatology, and the linkage between results from different imaging methods.

1.2 Current Management Guidelines and Remaining Gaps

With the growing prominence of concussion in the public and scientific scenes alike, numerous efforts and resources have been targeted at evolving the condition's management and clinical care standards (McCrory et al., 2009, 2013, 2017). Due to the integral nature of impact in several popular sports, sports related concussions (SRC) specifically have been the target of several key protocol refinements – which tended to inform management of concussion from other causes such as motor vehicle accidents (MVA). Prior to these changes, sideline evaluation of athletes after a hit was constrained to simple questions such as: “how many fingers are these?”, “where are you now?”, or “what is your name?.” Other tests examined a player’s balance by either closing their eyes or lifting one foot off the ground. Particularly, these assessments have been described as rudimentary, subjective to the person administering them, and were associated with mediocre performance (Broglio et al., 2017, 2007; Maddocks et al., 1995).

As of the date of writing, the latest consensus on concussion in sports defines sidelines evaluation of cognitive function as an essential component in concussion identification (McCrory et al., 2017). Neuropsychological batteries and tests are recommended to assess injury-related deficits after a hit, with the most well-known being the Sports Concussion Assessment Tool version 5 (SCAT5). It has been acknowledged, however, that the SCAT5’s utility diminishes significantly 3-5 days after injury (McCrory et al., 2017). Further, the consensus stated that the utility of any single tool is only adequate when used in conjunction with a variety of other neurological assessments that target mental cognition, oculomotor function, gait, coordination, balance, and others. Moreover, while there are several advancements in computerized neuropsychological assessment and a tendency to prescribe baseline tests that serve as a subject-specific comparison point, the consensus viewed them as not required. In regards to residual effects and sequelae, most prominently targeted in the present dissertation, the consensus stated that the literature is inconsistent, but that “Clinicians need to be mindful of the potential for long-term problems such as cognitive impairment, depression, etc. ...” (McCrory et al., 2017). While many studies have argued for the linkage between repeated concussions and CTE, the consensus remains that the causality of the two is not confirmed.

Based on the latest consensus, it is clear that despite a more formal approach towards the injury and its management strategies, there remains a multitude of gaps in the knowledge scientifically and, subsequently, clinically. First, while the dependency on a multitude of clinical assessment modalities has been argued as more accurate in concussion, the specifics of such assessments can vary widely across institutions, and/or clinical providers. That variance gives rise to inconsistent management standards and may cause unsound clinical decisions. Moreover, the nature of persistent symptoms remains largely unknown, particularly due to the complexity of disentangling pre-existing, co-morbid, and trauma-specific clinical symptoms (McCrory et al., 2017). Lastly, cognitive recovery has been shown to occasionally not overlap with clinical symptom resolution, an added layer of complexity when considering clinical decisions such as return-to-play and return-to-work (McCrory et al., 2017; Broglio et al., 2009; Cao et al., 2008). More refined methods are evidently needed to address all the previous points, providing more objective clinical measures of concussion detection, severity, and recovery.

1.3 Neurophysiological Indexing of Brain Function

Electroencephalography (EEG) is a method by which brain electrophysiological signals, mostly postsynaptic potentials in cortical layers, can be monitored non-invasively from the scalp. EEG is characterized by its high temporal resolution that is able to capture brain oscillations and cognitive brain responses as they unfold in time (see below; Luck, 2014). In contrast, spatial resolution of EEG is poor, with limited information on the neural origin of electrophysiological signals captured from the scalp. Since its inception, EEG has been purposed to cover several utilities that range from clinical applications to basic science research (Ebersole and Pedley, 2003). A routine, clinical, EEG typically describes what is referred to in the literature as resting-state (RS) EEG – the monitoring of EEG signals as a continuous signal over time, as opposed to event-related EEG (see below). In modern-day hospitals, EEG is rather underutilized. Epilepsy and seizure foci localization form the dominant portion of clinical EEG application. Routine EEG are also commonly utilized in assessing severe cases of brain injury, where some events can be diagnostically valuable (e.g., flat-line or isoelectric EEG, slow-waves, and wake-sleep cycles; see Ebersole and Pedley, 2003).

Beyond RS, EEG has also demonstrated both clinical and research viability using event-related methods – describing the time-locking of EEG signals to particular events of interest and in essence indexing the brain's response to those events (Luck, 2014). Of particular importance in the clinical setting, somatosensory evoked potentials (SEP) utilized in spinal cord surgery and involves continuous stimulation of the medial nerve while actively monitoring the responses on the brain. Auditory brainstem responses (ABR) and auditory evoked potentials (EP) are two other clinical applications of EEG that leverage continuous stimulation of the brain to investigate auditory cortex functioning and diagnose cases of brain death in severe brain injury and anoxia (Lütschg et al., 1983).

The above event-related responses index exogenous responses by the brain – coupled with direct stimulus processing and are affected by stimulus characteristics.

These responses are evoked at small latencies, ranging 0-100 ms after stimulus onset (Lütschg et al., 1983). EEG research has expanded on these responses using intricate presentations of stimulus patterns. Longer-latency responses, termed event-related potentials (ERP), are typically endogenous, elicited to high level cognitive function that is downstream from simple stimulus processing (Näätänen et al., 2007). The most common ERPs fall in the 100-500 ms range after stimulus onset and are often named after their polarity (positive or negative; P and N) and latency – for instance an N400, a negative response maximal 400 ms after stimulus onset. ERPs are commonly identified by five characteristics: polarity, latency after stimulus, waveform morphology, scalp topography, and context for elicitation (e.g., stimulus type, pattern, and frequency). Particularly, ERPs have been utilized in cognitive neuroscience research to investigate an expansive list of topics in cognition including language, memory, and emotion (e.g., Connolly and Phillips, 1994; Harker and Connolly, 2007; Cahn and Polich, 2006). ERPs have also seen extensive utility in research of consciousness and its disorders (Duncan et al., 2009). A large body of work showed the mismatch negativity (MMN) to be a sensitive marker for emergence from coma, surpassing all current clinical tools (Fischer et al., 2006, 1999; Qin et al., 2008; Duncan et al., 2009). The extensive work strengthens the view that ERPs are a powerful and reliable tool in both research and clinical endeavours; however, it remains true that their utility today is typically restricted to research studies.

The ERP literature is expansive, offering a wide range of common responses and their utilization in a variety of research applications as briefly overviewed above. Despite the dynamic nature of ERPs and their modulation by a variety of nuanced changes to experimental designs, there exist a number of canonical ERPs that are considered field standards. For the purposes of the present thesis and its topic of concussion, only the N1/P2 complex, the N2b, the MMN, and the P300 are elaborated on here.

- The *N1/P2 complex* is the earliest of the components listed above and is composed of two consecutive peaks: the N1 (N100) and the P2 (P200). The two are closely linked with low-level processing of stimuli and are evoked to both visual and auditory stimulation – although only the auditory response is considered here for relevance to the current dissertation (Näätänen and Picton, 1987). The N1 is predominantly a frontocentral negative response peaking at 100 ms after the presentation of a sound. The P2 is a positive response, topographically similar to the N1, and peaks after the N1 at 200 ms post-stimulus onset. The complex is modulated directly by stimulus characteristics such as frequency and loudness, categorizing it as an exogenous response (Näätänen and Picton, 1987). Moreover, the complex can be subjected to habituation effects in experimental designs with repeating stimuli, where the amplitudes tend to attenuate the longer the experiment duration (Näätänen and Picton, 1987). While the N1 is a consistent response, seen even under sedation in some cases (Blain-Moraes et al., 2016), the P2 overlaps with negative components after the N1 in some ERP designs and is often difficult to observe (see MMN and N2b below).

- The *MMN* is a heavily studied endogenous brain response, dating back to the late 1970s, and is typically elicited to stimulus pattern deviations (Näätänen et al., 1978; Näätänen et al., 1982; Näätänen et al., 2007). Of most importance, the MMN is elicited without the requirement of actively attending to the stimuli; however, the MMN can also be observed in attentiveness, albeit overlapping with the N2b (see below; Folstein and Van Petten, 2008). The MMN is a frontocentral negative response peaking 175-250 ms after stimulus onset. Particularly, the MMN is observed only in trials when a constant repeating pattern is interrupted by an unexpected (deviant) sound. While the MMN has seen numerous debates on its mechanism and its implications on consciousness and awareness, that is not addressed here (Näätänen et al., 2007); however, it is widely agreed that the ERP is a tool of great importance in clinical research (Todd et al., 2008; Fischer et al., 2006, 1999).

- The N2b is the temporal successor to the MMN (historically was called the N2a). The N2b signifies a negative peak, sharper and later than the MMN, and has been reported to require attention for elicitation (Folstein and Van Petten, 2008). The N2b is reported to have a frontocentral distribution in the auditory modality. While there are several reports of an additional component termed the N2c, its temporal and topographical overlap makes it difficult to disambiguate from the N2b (and the MMN) in the auditory modality. Thus, for the purposes of the current dissertation, the N2b denotes the response to infrequent stimuli in active oddball tasks, indexing aspects of cognitive control and executive functioning (Folstein and Van Petten, 2008). Moreover, the N2b will be used synonymously to indicate the N2 (N200) response seen in visual flanker and switch tasks (Folstein and Van Petten, 2008; Moore et al., 2014).

- The P300 is the most studied ERP with a span of applications from clinical research to brain-computer interfacing (Polich, 2007; Farwell and Donchin, 1988;

Fischer et al., 2006; Sellers et al., 2006). The ERP is characterized as a positive deflection peaking around 300 ms after stimulus onset. The P300 is often divided into two subcomponents: the P3a and the P3b. The P3a is a fronto-central component that peaks in the 250-300 ms window and is often linked to orienting and early attention processing (Polich, 2007). Conversely, the P3b peaks later, in the 300-500 ms window, and is predominantly centro-parietal. The P3b corresponds to late processing and reflects activities involving working memory (Polich, 2007). While the two components have been dissociated in the literature, common paradigms typically elicit a combination of the two. The most common ERP design to elicit the P300 is the oddball paradigm, showing a P300 when a target stimulus is observed by the participant. Further, the P3a is elicited to non-target infrequent stimuli in oddball designs (Polich, 2007). Lastly, the P300 is typically regarded as a response requiring an active task; however, passive designs were also found to elicit P300s, albeit with typically smaller amplitudes (Polich, 2007).

Specific to mTBI, ERPs have been extensively explored in the literature to extract potential markers for the condition's identification as well as to examine underlying cognitive deficits it leaves the afflicted. Early work by Segalowitz et al. (1997) demonstrated the neurophysiological implications of concussion as seen on ERPs. This line of ERP work has since expanded and numerously replicated, showing concussion-related effects in a multitude of ERPs tied to multiple facets of cognitive dysfunction known to follow concussion. In brief, the P300 (or P3b) has been reported as delayed and/or attenuated following concussion (Ruiter et al., 2019; De Beaumont et al., 2009; Gaetz et al., 2000; Baillargeon et al., 2012; Gosselin et al., 2012; Broglio et al., 2009).

The N2b has shown more varying patterns, but has been typically altered after concussion (Ledwidge and Molfese, 2016; De Beaumont et al., 2009; Brush et al., 2018; discussed at length in chapter 5). Recent reports have found significant changes in the MMN and the N1/P2 complex (Ruiter et al., 2019; Boshra et al., 2019). Further, a study by Fickling et al. (2019) found an attenuation of the N400 after concussion, representing the only reported effect of language-related electrophysiological change.

1.4 Machine Learning and Brain Signal Decoding

Machine learning (ML) defines a thriving field of computer science that focuses on the design of general tools to simulate artificial intelligence (AI). While AI is a general term, ML primarily denotes the design and application of the algorithms to facilitate, improve, and automate a broad range of problems (Obermeyer and Emanuel, 2016; Rajkomar et al., 2019). Supervised ML denotes the most common form in the field. It defines leveraging pre-labeled data to build a system that is able to classify (or regress) unseen data points. Although unsupervised learning and reinforcement learning are on the rise as established methodologies in the current scene, pure supervised learning remains the main component of the unprecedented utilization of ML in commercial and clinical applications (LeCun et al., 2015).

A typical (supervised) ML pipeline is dependent on labeled data. Observations, each with values known as features, are collected and labeled either at the point of collection or post-hoc using domain experts. The labeled data are split into three sets: training, validation, and testing. It is important to note that these names change depending on discipline and author; thus, we define them here in details. The training set contains a number of observations that comprise a subset of the original dataset used strictly to train a ML model. The validation set is used to optimize model hyperparameters and is used to maintain a divide between development and generalization testing in order to mitigate potential knowledge leakage, which inflates model performance. The testing set refers to data held out from all training and hyperparameter tuning and is only used to test the generalization performance of a fully trained model.

In practice, and especially with smaller datasets, a single split can be susceptible to statistical instability or idiosyncratic results based on random split configurations. This is partially mitigated by enforcing multiple splits (termed folds) of the data where the generalization performance is estimated as the mean performance of all individual folds (Combrisson and Jerbi, 2015). Briefly, data are split k unique ways where the split defines training tr_i and testing te_i sets where $i \in 1 \dots k$. Moreover, in cases when hyperparameter tuning or feature selection (see below) is required, a nested split is performed in every fold i to generate further training and validation splits.

Typical machine learning tools are able to automatically derive functions to best fit the data provided, with support vector machines (SVM) a most common traditional tool of ML (Cortes and Vapnik, 1995). For complex data, "shallow" ML tools require a degree of feature engineering, where a field expert is tasked with synthesizing and optimizing features that best describe the data. This process is commonly termed feature extraction or calculation. In cases where the number of observations outnumber the features, an additional feature selection procedure is executed, where a number of features are identified as most useful for the present provided dataset (Peng et al., 2005). Note that a valid execution of ML limits feature selection to the training and/or validation sets.

With the advent of deep learning tools, feature extraction (and selection) can be regarded as an inelegant solution to a given ML problem. Particularly, while expert knowledge and feature engineering are often a necessity, a purely data-driven extraction of features can outperform those practices as well as identify features not easily characterized by a human expert (LeCun et al., 2015); however, deep learning's need for large datasets and architecture tuning necessitates the use of shallow ML tools in many cases. Deep learning is defined as using multiple processing layers to automatically learn abstract hierarchical representations of the data. With minimal expert feature-engineering, deep learning currently holds the benchmark for many state-of-the-art applications (see LeCun et al., 2015 for review). Particularly, combinations of simple feed-forward (FF) and convolutional layers have been revolutional in convolutional neural networks' (CNN) extensive utility in speech recognition, image processing, and object recognition.

In recent years, ML has gained significant traction in the clinical field, with potential venues in offering a cost-efficient way of replicating expert judgements and decisions in a setting overloaded with data (Obermeyer and Emanuel, 2016; Rajkomar et al., 2019). ML enables expert-systems that are able to process high-dimensional clinical data and learn complex patterns that might also be difficult to detect or visualize for a human expert (Obermeyer and Emanuel, 2016; Rajkomar et al., 2019). A commonly cited downside of ML in healthcare and diagnostic applications is that most ML models are difficult to interpret Miotto et al. (2017). The tendency of viewing trained models as black boxes poses difficulties in a clinical setting since ML decisions cannot be checked against the body of available clinical knowledge. Several solutions have been proposed to provide insight on how a trained model makes its judgments (Ribeiro et al., 2016; Lundberg and Lee, 2017). Despite some scrutiny due to black-box solutions and susceptibility to bias in misapplication, machine learning remains a great tool for exploiting resources to improve clinical standards (Miotto et al., 2017; Chen and Asch, 2017; Miotto et al., 2017; Obermeyer and Emanuel, 2016; Lundberg et al., 2018).

EEG/ERP data are characterized by their rich high-dimensionality that requires certain degrees of aggregation to simplify to a human observer. That complexity was often exploited by the application of ML tools. Primarily, ML is the engine for all EEG-based brain computer interfacing (BCI), serving to train real-time, automated decoders that create a communications link with a computer strictly guided by the brain (Lotte et al., 2018; Blankertz et al., 2016; Dal Seno et al., 2010). Beyond BCI applications, ML has also demonstrated its clinical utility as a valuable method in EEG analysis (Tzovara et al., 2013; Cao et al., 2008; Lawhern et al., 2018; Schirrmeister et al., 2017; Cecotti and Gräser, 2011; Opałka et al., 2018; Sturm et al., 2016).

1.5 Dissertation Overview

Several studies have been conducted (a few with shared datasets) as part of the present PhD degree, four of which are included as separate chapters here. Each chapter targets a set of different hypotheses; however, all studies share three main goals that form the critical contribution of the studies. This dissertation's work was conducted to:

- Qualify and support the hypothesis that EEG/ERP effects can be leveraged for single-subject identification of mTBI, be that in acute or chronic stages of injury.

- Address whether a single model is capable of capturing the heterogeneity of different injuries and their respective variations. This is juxtaposed against current EEG and EEG/ERP analysis standards.
- Enforce the need for explainable, transparent models in targeting clinical ML applications in mTBI, as well as formulate falsifiable and clear theoretical framework for sources of detected variance.

The next four chapters detail independent studies investigating the dissertation's different hypotheses and research questions. An overview of each chapter is presented below:

- Chapter 2: The second chapter includes a pilot study on the feasibility of using significant group-level effects to visually assess single-subject averages for concussion. Two EEG/ERP experts were presented with subject-averages that were either from a chronically concussed group, or an age-matched control. Results showed that even for domain experts, identification of effects observed on groups is susceptible to confusion with subject variability. The chapter provides motivation for using systematic, data-driven, machine learning methods for EEG/ERP concussion assessment.
- Chapter 3: The chapter details a thorough investigation of machine learning's application in detecting long-lasting effects of concussion in retired athletes with a history of repeated concussions. The findings of the chapter support the utilization of ML to infer the state of a single individual's brain responses. Despite promising results, the chapter indicates that making clinical judgements

on the individual level requires more insight from data beyond what is typically investigated in group-level studies.

- Chapter 4: The fourth chapter expands on the previous one, exploring a fully realized deep learning architecture for a finer-grained analysis of concussion effects. The data utilized in this chapter were recorded from a younger population in a more acute stage of injury. Results demonstrated the effectiveness of the presented architecture in the classification of multi-deviant oddball data in detecting neurophysiological ERP changes after concussion. The chapter also takes a deeper look at the progression of a subset of concussed subjects, as they recover from their injury, to investigate whether the trained model is capable of directly capturing that progression.

- Chapter 5: As a conclusion from the previous two chapters, it was apparent that chronic and acute effects of concussions are speciously treated alike. This chapter attempts to clarify that difference by investigating event-related functional brain connectivity and its change from acute to chronic stages of concussion. The chapter reports a significant hyperconnectivity in the acute stages of concussion, a unique finding in the EEG/ERP literature that replicates a large body of functional MRI work. Strikingly, the effect was found to be reversed later in age as retired athletes demonstrated a significant decrease in connectivity. Results indicate a complex interaction of injury and compensatory mechanisms that require further investigation in longitudinal studies as well as single-subject exploration for clinical utilization.

- Chapter 6: The last chapter discusses the overall conclusions from the different studies, their impact on the literature, and their clinical implications. The work introduces several valuable tools for potential adaptation in concussion management and assessment; however, there are a number of limitations that stand barrier to direct implementation. These are discussed in detail. Lastly, a non-exhaustive list of future directions that maximize the impact of this dissertation is presented.

Bibliography

Baillargeon, A., Lassonde, M., Leclerc, S., and Ellemberg, D. (2012). Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Injury*, 26(3):211–220.

Blain-Moraes, S., Boshra, R., Ma, H. K., Mah, R., Ruiter, K., Avidan, M., Connolly, J. F., and Mashour, G. A. (2016). Normal Brain Response to Propofol in Advance of Recovery from Unresponsive Wakefulness Syndrome. *Frontiers in Human Neuroscience*, 10(June):1–6.

Blankertz, B., Acqualagna, L., Dähne, S., Haufe, S., Schultze-Kraft, M., Sturm, I., Uščumlic, M., Wenzel, M. A., Curio, G., and Müller, K. R. (2016). The Berlin brain-computer interface: Progress beyond communication and control. *Frontiers in Neuroscience*, 10(NOV).

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Macciocchi, S. N., and Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, 60(6):1050–7; discussion 1057–8.

Broglio, S. P., Moore, R. D., and Hillman, C. H. (2011). A history of sport-related concussion on event-related brain potential correlates of cognition. *International Journal of Psychophysiology*, 82(1):16–23.

Broglio, S. P., Pontifex, M. B., O'Connor, P., and Hillman, C. H. (2009). The Persistent Effects of Concussion on Neuroelectric Indices of Attention. *Journal of Neurotrauma*, 26(9):1463–1470.

Brush, C. J., Ehmann, P. J., Olson, R. L., Bixby, W. R., and Alderman, B. L. (2018). Do sport-related concussions result in long-term cognitive impairment? A review of event-related potential research. *International Journal of Psychophysiology*, 132(March 2017):124–134.

Cahn, B. R. and Polich, J. (2006). Meditation states and traits: Eeg, erp, and neuroimaging studies. *Psychological bulletin*, 132(2):180.

Cao, C., Tutwiler, R. L., and Slobounov, S. (2008). Automatic classification of athletes with residual functional deficits following concussion by means of EEG signal using support vector machine. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(4):327–335.

Cecotti, H. and Gräser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):433–445.

Chen, J. H. and Asch, S. M. (2017). Machine learning and prediction in medicine—beyond the peak of inflated expectations. *The New England journal of medicine*, 376(26):2507.

Combrisson, E. and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, 250:126–136.

Connolly, J. F. and Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of cognitive neuroscience*, 6(3):256–266.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*.

Dal Seno, B., Matteucci, M., and Mainardi, L. (2010). Online detection of P300 and error potentials in a BCI speller. *Computational Intelligence and Neuroscience*, 2010.

De Beaumont, L., Beauchemin, M., Beaulieu, C., and Jolicoeur, P. (2013). Long-term attenuated electrophysiological response to errors following multiple sports concussions. *Journal of Clinical and Experimental Neuropsychology*, 35(6):596–607.

De Beaumont, L., Brisson, B., Lassonde, M., and Jolicoeur, P. (2007). Long-term electrophysiological changes in athletes with a history of multiple concussions. *Brain Injury*, 21(6):631–644.

De Beaumont, L., Henry, L. C., and Gosselin, N. (2012). Long-term functional alterations in sports concussion. *Neurosurgical Focus*, 33(6):E8.

De Beaumont, L., Thoret, H., Mongeon, D., Messier, J., Leclerc, S., Tremblay, S., Ellemberg, D., and Lassonde, M. (2009). Brain function decline in healthy retired athletes who sustained their last sports concussion in early adulthood. *Brain*, 132(3):695–708.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., and Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, 120(11):1883–1908.

Ebersole, J. S. and Pedley, T. A. (2003). *Current practice of clinical electroencephalography*. Lippincott Williams & Wilkins.

Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510–523.

Fickling, S. D., Smith, A. M., Pawlowski, G., Ghosh Hajra, S., Liu, C. C., Farrell, K., Jorgensen, J., Song, X., Stuart, M. J., and D'Arcy, R. C. (2019). Brain vital signs detect concussion-related neurophysiological impairments in ice hockey. *Brain*, 142(2):255–262.

Fischer, C., Luauté, J., Némoz, C., Morlet, D., Kirkorian, G., and Mauguière, F. (2006). Improved prediction of awakening or nonawakening from severe anoxic coma using tree-based classification analysis. *Critical care medicine*, 34(5):1520–1524.

Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jourdan, C., and Salord, F. (1999). Mismatch negativity and late auditory evoked potentials in comatose patients. *Clinical Neurophysiology*, 110(9):1601–1610.

Folstein, J. R. and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1):152–170.

Gaetz, M., Goodman, D., and Weinberg, H. (2000). Electrophysiological evidence for the cumulative effects of concussion. *Brain Injury*, 14(12):1077–1088.

Gavett, B. E., Stern, R. A., and McKee, A. C. (2011). Chronic Traumatic Encephalopathy: A Potential Late Effect of Sport-Related Concussive and Subconcussive Head Trauma. *Clinics in Sports Medicine*, 30(1):179–188.

Gosselin, N., Bottari, C., Chen, J.-K., Huntgeburth, S. C., De Beaumont, L., Petrides, M., Cheung, B., and Ptito, A. (2012). Evaluating the cognitive consequences of mild traumatic brain injury and concussion by using electrophysiology. *Neurosurgical Focus*, 33(6):E7.

Harker, K. T. and Connolly, J. F. (2007). Assessment of visual working memory using event-related potentials. *Clinical neurophysiology*, 118(11):2479–2488.

Heitger, M. H., Jones, R. D., Macleod, A. D., Snell, D. L., Frampton, C. M., and Anderson, T. J. (2009). Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*, 132(10):2850–2870.

Hocke, L. M., Duszynski, C. C., Debert, C. T., Dleikan, D., and Dunn, J. F. (2018). Reduced Functional Connectivity in Adults with Persistent Post-Concussion Symptoms: A Functional Near-Infrared Spectroscopy Study. *Journal of Neurotrauma*, 35(11):1224–1232.

Kesler, S. R., Adams, H. F., Blasey, C. M., and Bigler, E. D. (2003). Premorbid Intellectual Functioning, Education, and Brain Size in Traumatic Brain Injury:

An Investigation of the Cognitive Reserve Hypothesis. *Applied Neuropsychology*, 10(3):153–162.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 15(5):1–30.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

Ledwidge, P. S. and Molfese, D. L. (2016). Long-Term Effects of Concussion on Electrophysiological Indices of Attention in Varsity College Athletes: An Event-Related Potential and Standardized Low-Resolution Brain Electromagnetic Tomography Approach. *Journal of Neurotrauma*, 33(23):2081–2090.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F. (2018). A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces: A 10-year Update. *Journal of Neural Engineering*, pages 0–20.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Lundberg, S. and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems*, 16(3):426–430.

Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., Kim, J., and Lee, S.-I. (2018). Explainable machine learning predictions to help anesthesiologists prevent hypoxemia during surgery. *Nature Biomedical Engineering*.

Lütschg, J., Pfenninger, J., Ludin, H. P., and Vassella, F. (1983). Brain-stem auditory evoked potentials and early somatosensory evoked potentials in neurointensively treated comatose children. *American Journal of Diseases of Children*, 137(5):421–426.

Maddocks, D. L., Dicker, G. D., and Saling, M. M. (1995). The assessment of orientation following concussion in athletes. *Clinical journal of sport medicine: official journal of the Canadian Academy of Sport Medicine*, 5(1):32–35.

McAllister, T. W., Sparling, M. B., Flashman, L. A., and Saykin, A. J. (2001). Neuroimaging Findings in Mild Traumatic Brain Injury *. *Journal of Clinical and Experimental Neuropsychology*, 23(6):775–791.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R. C., Cassidy, D., Echemendia, R. J., Castellani, R. J., Davis, G. A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C. C., Guskiewicz, K. M., Herring, S., Iverson, G. L., Johnston, K. M., Kissick, J., Kutcher, J., Leddy, J. J., Maddocks, D., Makdissi, M., Manley, G. T., McCrea, M., Meehan, W. P., Nagahiro, S., Patricios, J., Putukian, M., Schneider, K. J., Sills, A., Tator, C. H., Turner, M., and Vos, P. E. (2017). Consensus statement on concussion in sport—the 5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, (October 2016):bjsports–2017–097699.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., and

Cantu, R. (2009). Consensus Statement on Concussion in Sport – The Third International Conference on Concussion in Sport Held in Zurich, November 2008. *The Physician and Sportsmedicine*, 37(2):141–159.

McCrory, P., Meeuwisse, W. H., Aubry, M., Cantu, B., Dvořák, J., Echemendia, R. J., Engebretsen, L., Johnston, K., Kutcher, J. S., Raftery, M., Sills, A., Benson, B. W., Davis, G. A., Ellenbogen, R. G., Guskiewicz, K., Herring, S. A., Iverson, G. L., Jordan, B. D., Kissick, J., McCrea, M., McIntosh, A. S., Maddocks, D., Makdissi, M., Purcell, L., Putukian, M., Schneider, K., Tator, C. H., and Turner, M. (2013). Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, 47(5):250–258.

McKee, A. C., Cantu, R. C., Nowinski, C. J., Hedley-Whyte, E. T., Gavett, B. E., Budson, A. E., Santini, V. E., Lee, H. S., Kubilus, C. A., and Stern, R. A. (2009). Chronic Traumatic Encephalopathy in Athletes: Progressive Tauopathy After Repetitive Head Injury. *Journal of neuropathology and experimental neurology*, 68(7):709–735.

Meythaler, J. M., Peduzzi, J. D., Eleftheriou, E., and Novack, T. A. (2001). Current concepts: Diffuse axonal injury-associated traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 82(10):1461–1471.

Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, (February):1–11.

Moore, R. D., Broglio, S. P., and Hillman, C. H. (2014). Sport-related concussion and sensory function in young adults. *Journal of Athletic Training*, 49(1):36–41.

Näätänen, R., Gaillard, A. W., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, 42(4):313–329.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (mmn) in basic research of central auditory processing: a review. *Clinical neurophysiology*, 118(12):2544–2590.

Näätänen, R. and Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*.

Näätänen, R., Simpson, M., and Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*.

Obermeyer, Z. and Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, 375(13):1216.

Omalu, B. I., DeKosky, S. T., Minster, R. L., Kamboh, M. I., Hamilton, R. L., and Wecht, C. H. (2005). Chronic Traumatic Encephalopathy in a National Football League Player. *Neurosurgery*, 57(1):128–134.

Opałka, S., Stasiak, B., Szajerman, D., and Wojciechowski, A. (2018). Multi-channel convolutional neural networks architecture feeding for effective EEG mental tasks classification. *Sensors (Switzerland)*, 18(10):1–21.

Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b.

Qin, P., Di, H., Yan, X., Yu, S., Yu, D., Laureys, S., and Weng, X. (2008). Mismatch negativity to the patient's own name in chronic disorders of consciousness. *Neuroscience Letters*, 448(1):24–28.

Raghupathi, R. (2004). Cell death mechanisms following traumatic brain injury. *Brain Pathology*, 14(5):215–222.

Rajkomar, A., Dean, J., and Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14):1347–1358.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135—-1144.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38(11):5391–5420.

Segalowitz, S. J., Dywan, J., and Unsal, A. (1997). Attentional factors in response time variability after traumatic brain injury: an ERP study. *J Int Neuropsychol Soc*, 3(2):95–107.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2006). A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance. *Biological Psychology*, 73(3):242–252.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, 47(10):2015–2028.

Sturm, I., Lapuschkin, S., Samek, W., and Müller, K. R. (2016). Interpretable deep neural networks for single-trial EEG classification. *Journal of Neuroscience Methods*, 274:141–145.

Taylor, C. A., Bell, J. M., Breiding, M. J., and Xu, L. (2017). Traumatic Brain Injury–Related Emergency Department Visits, Hospitalizations, and Deaths — United States, 2007 and 2013. *MMWR. Surveillance Summaries*, 66(9):1–16.

Todd, J., Michie, P. T., Schall, U., Karayanidis, F., Yabe, H., and Näätänen, R. (2008). Deviant Matters: Duration, Frequency, and Intensity Deviants Reveal Different Patterns of Mismatch Negativity Reduction in Early and Late Schizophrenia. *Biological Psychiatry*, 63(1):58–64.

Tremblay, S., De Beaumont, L., Henry, L. C., Boulanger, Y., Evans, A. C., Bourgouin, P., Poirier, J., Théoret, H., and Lassonde, M. (2013). Sports concussions and aging: A neuroimaging investigation. *Cerebral Cortex*, 23(5):1159–1166.

Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., and De Lucia, M. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, 136(1):81–89.

CHAPTER 2 Visual Inspection: Insight on Effectiveness in Modulated ERP Responses

Clinically, visual inspection of EEG waveforms is considered the standard of analysis (Ebersole and Pedley, 2003). Several studies have argued the subjectivity of such practice, citing a multitude of variables introducing variance between clinicians' opinion both within- and across-clinician Little and Raffel (1962); Nuwer et al. (2005). Analogously in EEG/ERP work, the literature has employed a number of statistical methods to both validate and standardize reported results (Tzovara et al., 2013; Naccache et al., 2016; Gabriel et al., 2016). Despite such efforts, visual inspection of waveforms remains the de-facto standard in the field, often providing better results in comparison to more structured analysis methods (Fischer et al., 1999; Gabriel et al., 2016). While advancements in signal-processing techniques continues to improve the available tools for ERP identification, reports show visual-inspection to be adequate for identification and diagnosis in patients with DOC (Naccache et al., 2016; Gabriel et al., 2016). Provided the nature of ERP application in severe cases, a clinician's goal is to identify whether an ERP was elicited or not. Less severe cases, however, show more nuanced differences in their ERPs and require a more precise identification of subtle modulations in ERP topography, morphology, or both.

The present chapter is a preliminary investigation into the effectiveness of visual inspection for the identification of non-catastrophic brain injury. We collected judgements from EEG/ERP experts on the identification of post-concussive effects as exhibited in subjects' waveforms. Our hypothesis predicted the ability of an expert's ability to detect between typical and aberrant responses. However, we expected that visual-inspection performance to be lower in comparison to DOC applications. We argued this reduction to be caused by the inherent complexity of identifying ERP modulations due to an overlap of inter-subject variability and inter-group differences.

2.1 Methods

2.1.1 EEG Data

Two EEG/ERP experts were asked to rate responses to an active multi-deviant oddball protocol visually presented in a survey format. The EEG/ERP protocol was split to four conditions: Standard tones (Std), Frequency Deviants (FDev), Duration Deviants (DDev), and Intensity Deviants (IDev). The details of the protocol stimuli, presentation parameters, and procedure can be found in more detail in an earlier publication by our group (Ruiter et al., 2019). EEG data were collected from 39 consenting male adults. 19 of the participants were retired Canadian football league athletes with a history of self-reported concussions. The other 20 participants formed a control group with no reported history of head trauma. Statistical analysis on the data yielded a range of effects. Particularly, main effects were found in the amplitudes of the N1, N2b, P3a, and P3b responses, indicating an attenuated peak for concussed individuals (Ruiter et al., 2019).

2.1.2 Survey Stimuli

The survey contained a total of 117 question groups each targeting a single-subject's averaged response. A plot of a subject's averaged response to a single deviant type was overlaid on the response to Std as recorded from the Fz, Cz, and Pz electrodes. Grand-averages for both experimental groups (controls and concussed athletes) were displayed on both sides of the single-subject's responses (see figure 2.1). Note that the grand-averages displayed for each question were from to same deviant-type as the single-subject plot.

Each question group was comprised of two multiple-choice questions regarding the single-subject's averaged responses. Question 1 asked the expert to select one of five options to describe the single-subject average: "very similar to controls," "somewhat similar to controls," "undecided," "somewhat similar to concussed population," or "very similar to concussed population." The second question asked about the expert's confidence in his/her answer on the previous question where the answer was on a 5-point Likert scale spanning from "not confident at all" to "very confident."

2.1.3 Expert Visual Inspection

Four volunteer EEG/ERP experts were asked to complete the survey. Two of the experts did not complete the survey in full and were discarded from further analyses. All experts considered for the present experiment had a doctoral degree with demonstrable history of peer-reviewed publications using the EEG/ERP methodology. The survey was hosted and completed online using the LimeSurvey platform (Schmitz, 2012). Experts were provided unlimited time to complete the survey.

2.1.4 Statistical Analysis

Responses were exported in full from LimeSurvey and imported for analysis using R statistical software (version 3.5.3). Multiple logistic regression was conducted to examine predictability of Group (control vs. concussed) provided the effect of Deviant (3 levels: FDev, DDev, and IDev), in addition to judgement from expert 1, judgment from expert 2, and their interaction. Additionally, an overall estimate of sensitivity and specificity for each expert (with respect to concussion identification) was reported such that all answers of "undecided" were discarded, and answers suggesting either side were reduced to a binary decision. Sensitivity was defined for the concussed class as $\frac{TP}{TP+FN}$ where TP is true positive predictions as concussed and FN is falsely predicting a control where the true label was concussed. Specificity was defined as $\frac{TN}{TN+FP}$ such that TN (true negative) was the number of controls correctly identified as such and FP (false positive) was the number of controls misidentified as concussed. Lastly, a correlation analysis was conducted to examine the agreement between the two experts using spearman's correlation.

2.2 Results

There was a positive correlation between the two experts' judgements ($\rho = 0.71, N = 114, p < 0.01$). Logistic regression results showed a significant effect of judgements from both expert 1 ($\beta = -0.85, z = -1.98, p < 0.05$) and expert 2 ($\beta = -1.32, z = -2.31, p < 0.05$). These findings were in accordance with the hypothesis that visual

	Accuracy	**Sensitivity**	**Specificity**
Expert 1			
DDev	57.9	52.6	63.2
FDev	63.2	68.4	57.9
IDev	63.2	63.2	63.2
Total	61.1	61.4	61.4
Expert 2			
DDev	60.6	43.8	76.5
FDev	65.5	61.5	68.8
IDev	60.0	38.5	76.5
Total	62.0	47.6	74.0
Combined			
DDev	59.2	48.6	69.4
FDev	64.2	65.6	62.9
IDev	61.8	53.1	69.4
Total	61.7	55.6	67.3

Table 2.1: Performance (as measured by accuracy, sensitivity, and specificity) of the experts identifying a history of concussion in single-subject averages using each of the deviant types. Aggregate performance across the two experts is also shown.

inspection was able to predict a single-subject's group, where an increase on the scale – indicating similarity to the concussed grand-average – corresponded with an increased prediction of belonging to the concussed group. No reliable effects of condition or the interaction between the two judgments were found.

Computation of binary classification metrics showed slightly varying results between the two experts (see table 2.1). Expert 1 had an overall higher sensitivity (61.4%), whereas expert 2 had a higher specificity (74.0%). Both had comparable accuracies with 61.1% and 62.0% for expert 1 and 2, respectively. Insight on individual deviant types could be drawn from table 2.1, showing a capability for high sensitivity in FDev responses and high specificity in the DDev and IDev. Note that these observations were not supported by statistical results, which we argue are attributable to the low number of observations and insufficient statistical power.

2.3 Discussion

This pilot study provided a clear demonstration of apparent difficulty in assessing complexly modulated responses in an application such as concussion identification. While EEG/ERP experts provided judgments that were statistically predictive of a subject's experimental group (control vs. concussed), metrics of more clinical utility conveyed a need for more elaborate analytics (table 2.1.

Particularly, it is noted that one expert was primarily focused on minimizing type II errors (higher sensitivity), whereas the other had more inclination towards more conservative judgements that prioritized high specificity. Critically, while this promotes a layer of subjectivity based on the expert's view of the correct approach, identifying a correct strategy is not trivial. For instance, an expert that prioritizes the reduction of false negatives may do so in an effort to maximize patient access to clinical help; in exchange, this strategy is more susceptible to false-positives which may put an extra load on the clinical system or the patient, if treatment was not required. In contrast, the expert that did not skew their judgement towards higher sensitivity had a marginally higher overall accuracy; however, there were more missed positives, which, in a real clinical scenario, would be detrimental to the well-being of a patient. In essence, an EEG/ERP expert forced to make a clinical decision with what could be argued as limited information does not necessarily have an optimal strategy to follow when dealing with a complex problem with a large number of unknown variables. Notably, while results conveyed relatively poor performance, visual inspection remained a better discriminatory tool than a number of behavioural measurements used by clinicians to assess concussions today (Broglio et al., 2007, 2017).

Despite a simple design and a limited number of expert judgements, we argue that findings from this pilot study correspond with and support the primary motivation and hypothesis question of the investigation: modulatory effects on ERPs due to mTBI are drastically harder to quantify and predict for a human expert than typical cases of clinical ERP applications in the literature. Perhaps the most established clinical application of ERPs to date is using the MMN to predict coma emergence, with some studies reporting a 100% positive predictive value for awakening (Fischer et al., 2006; Duncan et al., 2009). Several reports have attempted to automate the ERP detection procedures in this application either statistically or using methods of ML (Fischer et al., 2006; Naccache et al., 2016; Tzovara et al., 2013; Qin et al., 2008; Armanfard et al., 2018); however, the standard largely remains visual inspection by a human expert (Gabriel et al., 2016; Naccache et al., 2016). The current results suggest more complex requirements for clinically utilizing ERPs when the pathological effect involves a gradation effect, as opposed to an all-or-nothing effect in MMN detection. Specific to concussion, ERPs have been reported to be smaller, larger, or delayed (Moore et al., 2014; De Beaumont et al., 2009; Ruiter et al., 2019), which can be difficult to observe using typical ERP visualization (see figure 2.1). Thus, our findings point towards the applicability of fine-grained machine learning methods that can assist a human expert in processing datapoints with dimensions difficult to visualize by traditional tools (Shrikumar et al., 2017; Obermeyer and Emanuel, 2016; Rajkomar et al., 2019).

The current pilot exhibits several limitations that require further work to formalize and support the presented findings. Primarily, the number of judgements was limited to only two due to limited access to EEG/ERP experts available to complete the expansive survey. While this limits our ability to draw strong conclusions from our results, the consistency between the two experts' opinions in terms of performance suggests a trend. Secondly, in an effort to simplify the experiment, only static waveforms were provided as extracted from a small subset of the total recorded channels (three out of 64 total). Provided that visual inspection in the literature is traditionally based on the midline, we argue that our approach minimally affected our results.

2.4 Conclusions

Findings from the presented pilot study motivate stepping beyond visual inspection in application of ERPs for concussion assessment. Results convey the relative variability and subjectivity of the practice when information provided, and full understanding of the pathology and its effects, are limited. In the next two chapters, the applicability of ML to ERP data for concussion assessment is thoroughly investigated.

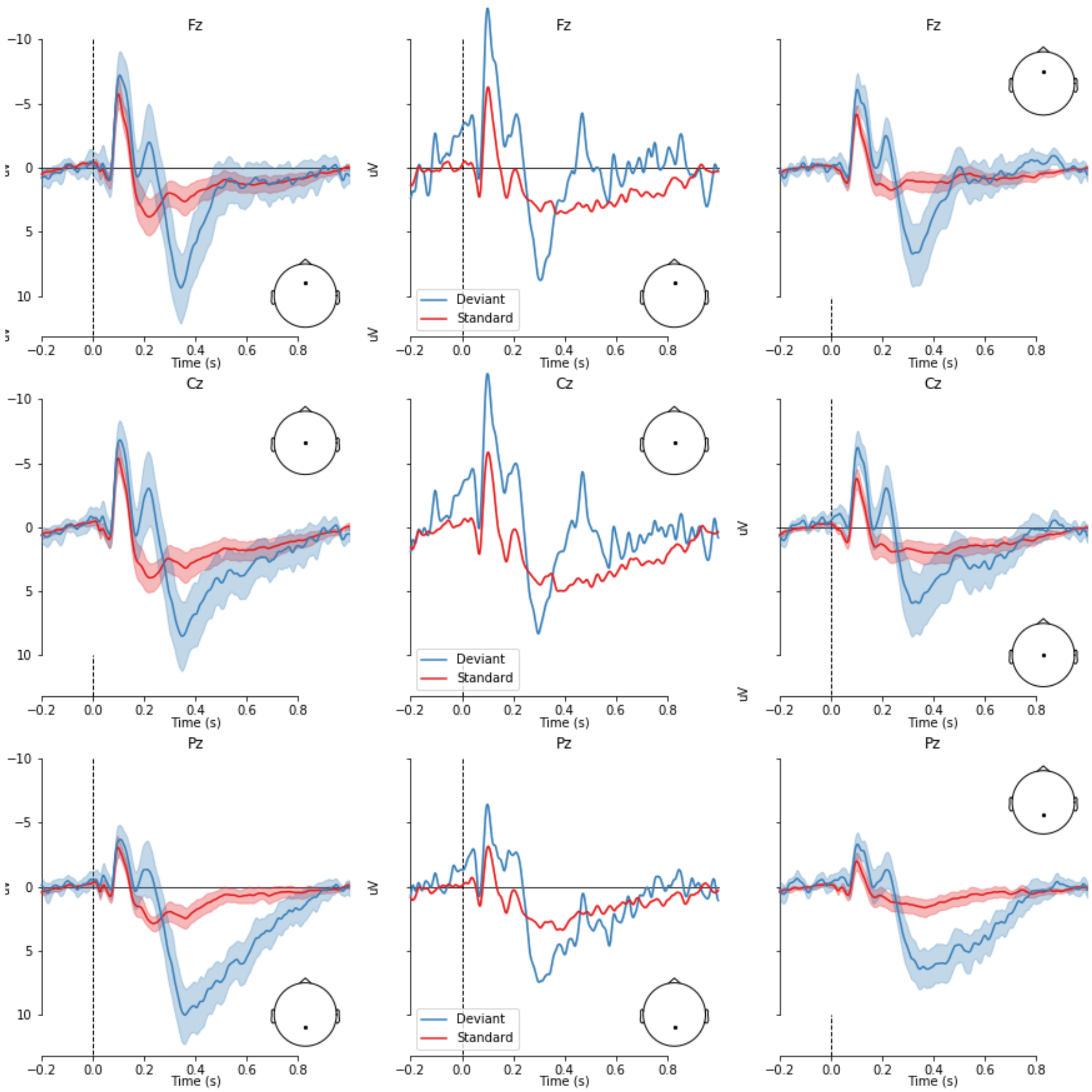

Figure 2.1: A sample plot as seen by an expert during survey completion. Rows from top to bottom indicate responses as recorded from the the Fz, Cz, and Pz electrodes, respectively. The columns from left to right are responses of: control group, single-subject, and concussed group, respectively. Shaded regions represent standard deviation across the respective group's responses. Ordinate and absicca for each plot correspond to amplitude (in μVolt) and time (in seconds) where 0 indicates stimulus onset. Red waveform indicate the response to standard tones whereas the red waveforms indicate the responses to one of the three deviants (FDev, DDev, or IDev).

Bibliography

Armanfard, N., Komeili, M., Reilly, J. P., and Connoly, J. (2018). A machine learning framework for automatic and continuous mmn detection with preliminary results for coma outcome prediction. *IEEE journal of biomedical and health informatics.*

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Macciocchi, S. N., and Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, 60(6):1050–7; discussion 1057–8.

De Beaumont, L., Thoret, H., Mongeon, D., Messier, J., Leclerc, S., Tremblay, S., Ellemberg, D., and Lassonde, M. (2009). Brain function decline in healthy retired athletes who sustained their last sports concussion in early adulthood. *Brain*, 132(3):695–708.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., and Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, 120(11):1883–1908.

Ebersole, J. S. and Pedley, T. A. (2003). *Current practice of clinical electroencephalography*. Lippincott Williams & Wilkins.

Fischer, C., Luauté, J., Némoz, C., Morlet, D., Kirkorian, G., and Mauguière, F. (2006). Improved prediction of awakening or nonawakening from severe anoxic coma using tree-based classification analysis. *Critical care medicine*, 34(5):1520–1524.

Fischer, C., Morlet, D., Bouchet, P., Luaute, J., Jourdan, C., and Salord, F. (1999). Mismatch negativity and late auditory evoked potentials in comatose patients. *Clinical Neurophysiology*, 110(9):1601–1610.

Gabriel, D., Muzard, E., Henriques, J., Mignot, C., Pazart, L., André-Obadia, N., Ortega, J. P., and Moulin, T. (2016). Replicability and impact of statistics in the detection of neural responses of consciousness. *Brain*, 139(6):e30.

Little, S. C. and Raffel, S. C. (1962). Intra-rater reliability of eeg interpretations. *The Journal of nervous and mental disease*, 135(1):77.

Moore, R. D., Broglio, S. P., and Hillman, C. H. (2014). Sport-related concussion and sensory function in young adults. *Journal of Athletic Training*, 49(1):36–41.

Naccache, L., Sitt, J., King, J.-R., Rohaut, B., Faugeras, F., Chennu, S., Strauss, M., Valente, M., Engemann, D., Raimondo, F., Demertzi, A., Bekinschtein, T., and Dehaene, S. (2016). Reply: Replicability and impact of statistics in the detection of neural responses of consciousness. *Brain*, 139(6):e31–e31.

Nuwer, M. R., Hovda, D. A., Schrader, L. M., and Vespa, P. M. (2005). Routine

and quantitative EEG in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025.

Obermeyer, Z. and Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, 375(13):1216.

Qin, P., Di, H., Yan, X., Yu, S., Yu, D., Laureys, S., and Weng, X. (2008). Mismatch negativity to the patient's own name in chronic disorders of consciousness. *Neuroscience Letters*, 448(1):24–28.

Rajkomar, A., Dean, J., and Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14):1347–1358.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Schmitz, C. (2012). LimeSurvey: An open source survey tool. *LimeSurvey Project Hamburg, Germany. URL http://www. limesurvey. org.*

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning important features through propagating activation differences. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3145–3153. JMLR. org.

Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., and De Lucia, M. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, 136(1):81–89.

CHAPTER 3

From Group-Level Statistics to Single-Subject Prediction: Machine Learning Detection of Concussion in Retired Athletes

Preface

This study utilizes data that have been traditionally analyzed, supporting a persistent effect of concussions decades after injury in retired athletes (Ruiter et al., 2019). The study serves as the first of the two ML studies presented in the dissertation to attempt and demonstrate the transfer of group effects in EEG/ERP to a clinically applicable single-subject tool. The study focused on validating the ML design by rigorous validation to counteract a limited dataset size and reported promising findings. The chapter is a reprint from an article submitted to IEEE Transactions on Neural Systems and Rehabilitation Engineering in 2019 with the following authors:

Boshra, R., Dhindsa K. Boursallie, O., Ruiter, K. I., Sonnadara, R., Doyle, T., Samavi, R., Reilly, J. P., & Connolly, J. F.

Abstract

There has been increased effort to understand the neurophysiological effects of concussion aimed to move diagnosis and identification beyond current subjective behavioural assessments that suffer from poor sensitivity. Recent evidence suggests that event-related potentials (ERPs) measured with electroencephalography (EEG) are persistent neurophysiological markers of past concussions. However, as such evidence is limited to group-level analyses, the extent to which they enable concussion detection at the individual-level is unclear. One promising avenue of research is the use of machine learning to create quantitative predictive models that can detect prior concussions in individuals. In this study we translate the recent group-level findings from ERP studies of concussed individuals into a machine learning framework for performing single-subject prediction of past concussion. We found that a combination of statistics of single-subject ERPs and wavelet features yielded a classification accuracy of 81% with a sensitivity of 82% and a specificity of 80%, improving on current practice. Notably, the model was able to detect concussion effects in individuals who sustained their last injury as much as 30 years earlier. However, failure to detect past concussions in a subset of individuals suggests that the clear effects found in group-level analyses may not provide us with a full picture of the neurophysiological effects of concussion.

3.1 Introduction and Related Work

Mild traumatic brain injury (mTBI), commonly referred to as concussion, has recently received increased attention in the scientific community and the public alike owing to recent studies demonstrating a link between mTBI and various detrimental neurological conditions, such as chronic traumatic encephalopathy (see (Henry et al., 2017) for a review). The subject of sports-related concussion has in particular undergone significant advancement as our improved understanding of its cognitive and neurophysiological impact has led to new clinical guidelines and management strategies (McCrory et al., 2009, 2013, 2017). One key driver of this rise in interest is the growing body of evidence suggesting that the effects of mTBI have long-term and cumulative effects on cognitive and neurological health, potentially persisting over the entire lifespan of an individual (Ruiter et al., 2019; Broglio et al., 2009; De Beaumont et al., 2007b). Despite the ongoing efforts, however, as many as 50% of concussions may still go undetected (Harmon et al., 2013).

Although recent evidence indicates that mTBI is more detrimental to health than previously thought, there has been little advancement in our ability to detect concussion in individual patients. Currently, concussion detection relies solely on the observation of symptoms and behaviourally-manifesting cognitive deficits in memory, attention and others (McCrory et al., 2017). Studies based on such measures suggest that the symptoms of concussion are primarily short-term, typically resolving within only one week, though an estimated 15% of patients may remain symptomatic for months or years post-injury (McCrory et al., 2017). Behaviour-based approaches suffer from inherent subjectivity and dependence on a multitude of factors unrelated to injury (Broglio et al., 2017). Further, behaviour alone has been shown to be unable to identify all concussed individuals, nor track injury recovery (Broglio et al., 2007). This has prompted researchers to search for neurophysiological markers of concussion directly reflective of cognitive processes upstream of behavioural manifestations.

Early investigations into the presence of neurophysiological markers of concussion were focused on the spectral characteristics of eyes-closed resting-state (RS) electroencephalogram (EEG). While a decrease in alpha-band activity was found to follow concussion in some cases, this marker was not sufficiently reliable for clinical adoption (Nuwer et al., 2005). Particularly, Nuwer *et al.* argued that mTBI markers in RS EEG were not consistently replicated across different studies. Additionally, there is converging evidence that such markers in RS EEG are not observable six months after injury (see (Nuwer et al., 2005) for review).

More recent investigations of potential neurophysiological markers of mTBI have focused on event-related potentials (ERPs), which are canonical brain responses induced by specific kinds of stimuli. ERPs are commonly identified by four characteristics: polarity, latency after stimulus, scalp topography, and context for elicitation (e.g., stimulus type, pattern, and frequency). Neurophysiological effects of concussion in ERPs extended beyond acute injuries and were observable decades after injury (Broglio et al., 2007; De Beaumont et al., 2007b,a; Fickling et al., 2019; Ruiter et al., 2019). These findings persisted long after behaviourally-identified symptoms had resolved, further demonstrating the inadequacy of assessing concussion using behavioural measures alone. Various ERPs were found to be affected by mTBI (Broglio et al., 2009; Brush et al., 2018; Ruiter et al., 2019). In this study we investigated how reliably the P300, a heavily studied ERP, and other precursor ERPs can be used as concussion identification markers for individuals who sustained their last injury years prior.

The P300 is a positive component that typically peaks at around 300 ms post-stimulus onset (Polich, 2007). This ERP has been shown to be attenuated and delayed after mTBI, with the effect persisting after symptom resolution (De Beaumont et al., 2012). The P300 can be subclassified into the P3a and the P3b, which were shown to be both altered after concussion (Broglio et al., 2009; Brush et al., 2018; Ruiter et al., 2019). The P3a is a fronto-central component associated with attention orienting and is followed by the P3b, a centro-parietal component often linked to processes related to memory and attention allocation (Polich, 2007). Moreover, two other ERPs commonly elicited with the P300 were found to be affected by mTBI: the N2b (associated with cognitive inhibition) and the N100 (associated with auditory processing) (Brush et al., 2018). Provided these ERPs link to processes commonly affected by concussion, they offer a valuable tool for directly tapping into mTBI-related deficits.

The neuroscientific literature has provided a valuable understanding of how mTBI affects brain activity. However, this knowledge has been derived on the basis of group-level analysis. While neurophysiological markers of mTBI have been discovered in group-averaged ERPs and are seen in many of the individuals who form part of the group, these methods do not capture the full range of individual variability. As a result, the specific degree to which such knowledge can be applied at the individual level is uncertain. Notwithstanding the limitations in applying knowledge gained on the basis of group-level statistics to individual patients, detection, diagnosis, and assessment of mTBI in clinical settings require methods applicable at the individual level. It is therefore necessary to investigate the degree to which the group-level neurophysiological markers of mTBI can be used to detect past concussion at the single-subject level.

Machine learning (ML) provides us with a set of tools specialized for learning patterns from samples of data to then classify individual cases. As such, ML provides us with a methodological foundation for translating what concussion research has gleaned from group-level analysis into single-subject detection of mTBI. The application of ML to EEG has been spearheaded within the field of brain-computer interfacing (Blankertz et al., 2016; Dhindsa et al., 2017; Lotte et al., 2018), and has recently been used to identify predictive features of the EEG in a variety of clinical and non-clinical settings (Lin et al., 2010; Chan et al., 2011; Tzovara et al., 2013; Dhindsa and Becker, 2017; Parvar et al., 2015; Sculthorpe-Petley et al., 2015; Ravan et al., 2015; Armanfard et al., 2018).

In mTBI specifically, ML was explored for the identification of functional deficits after symptom resolution (Cao et al., 2008). Support Vector Machines (SVMs) (Cortes and Vapnik, 1995) were trained on band power features extracted from RS data from 31 athletes at baseline, in addition to from 30 athletes after they had sustained a concussion (Cao et al., 2008). Linear SVMs were reported to have a 77.1% accuracy of correctly identifying the concussed (30) vs non-concussed recordings (31) in a leave-one-out (LOO) cross-validation study. Notably, the 30 concussed athletes reported complete symptom recovery at the point of testing 30 days after injury, further supporting a subclinical effect of mTBI on brain activity. A later study targeted classification between varying severities of TBI and reported high accuracies (Prichep et al., 2012). However, the study's results primarily support RS EEG's utility in detecting more severe TBI, with anatomical damage observable using computerized tomography (CT). Reported results for RS showed dramatically reduced accuracies for classifying subjects exhibiting strictly functional deficits. Several studies have explored novel feature-extraction methods based on RS data with promising results for acute injury detection (Munia et al., 2017; Cao and Slobounov, 2011, 2010). However, no direct single-subject characteristics of the models were derived.

In this paper, we investigate the extent to which our present understanding of how mTBI affects brain activity, based on group-level studies, can be used to detect past concussions at the individual level. In addition, we investigate if we can retain the same interpretation of how mTBI affects neural processing gained through traditional neuroscientific studies when transitioning from group-level analysis to single-subject classification of mTBI. The present study aims to expand on previous group-level work employing ML to perform single-subject detection of mTBI. Specifically, we use the findings and data from our group's earlier study (Ruiter et al., 2019) on the analysis of the effects of mTBI on ERPs as a basis for transitioning to single-subject ML analysis. We do so by using features representative of the mTBI-induced ERP changes to classify mTBI. We reported a classification accuracy of 81% with a sensitivity of 82% and a specificity of 80%. Notably, this work illustrates that a history of mTBI can be detected from an individual's brain activity decades after injury. Additionally, ML analysis with the added layer of interpretability provided novel insights into how the brain is affected by concussion, revealing discriminative features not previously reported in the neuroscientific literature.

3.2 Data Collection and Preprocessing

3.2.1 Participants

Data were collected from thirty-nine consenting male adults. Nineteen of the subjects were retired Canadian Football League (rCFL) athletes with mean age of 57.6. The remaining 20 participants (mean age = 53.7) formed an age-, education-level-, and sex-matched control group reporting no history of head trauma. All recruited participants reported no auditory problems and were not on psychoactive medications. The study was approved by the local research ethics board at McMaster University.

3.2.2 Behavioural Assessments

Participants completed a battery of self-reports including: 1) Post-Concussion Symptom Scale (PCSS), 2) Beck Depression Inventory II (BDI-II), and 3) General Health Questionnaire (SF-36). Additionally, demographic data were collected from rCFL participants including: 1) date since last reported concussion, 2) number of reported concussions, and 3) years of education. Notably, demographics were dependent on self-reported values as no direct access to clinical history was attainable.

3.2.3 EEG Stimuli

Data were collected as part of a three-paradigm experiment (Ruiter et al., 2019). The present study uses data from the paradigm designed to elicit a P300 response, which revealed significant differences between individuals with prior mTBI and those without. The P300 paradigm consisted of a modified auditory oddball task adapted from (Todd et al., 2008). The paradigm was composed of a total of 600 tone presentations including a standard tone (Std) and three deviants. The standard tone, a pure 1000 Hz at 80 dB sound-pressure level (SPL) for 50 ms, was presented 82% of the time (492 instances). Each deviant tone was presented 36 times (18% of the total stimulus presentations altogether) and differed in a single way from the standard tone. Duration deviant tones (DDev) were presented for 100 ms, frequency deviant tones (FDev) were 1200Hz, and intensity deviant tones (IDev) were presented at 90 dB SPL. In the sequel we refer to four "conditions"; these are Std, DDev, FDev, and IDev. Participants were tasked to respond to every standard tone by pressing one button and another button for all deviant tones.

3.2.4 Procedure

After completion of self-report questionnaires and the demographics questionnaire, participants were seated 90-cm away from a computer monitor in a sound-attenuated room. All stimuli were delivered binaurally through sound-isolating earphones (Etymotic II). After a practice trial, participants were asked to respond to stimuli as explained above as accurately as possible.

3.2.5 EEG Recording and Preprocessing

Continuous EEG was collected from 64 Ag/AgCl electrodes placed according to the extended 10/20 system using an elastic cap. Signals were digitized at 512 Hz using the ActiveTwo BioSemi system with an online bandpass filter of 0.01-100Hz. Data were referenced online to the driven right leg. Electrooculography (EOG) was recorded using two external electrodes above and over the outer canthus of the left eye. Three additional external electrodes were placed on the two mastoid processes and on the tip of the nose. All external electrodes were recorded with the settings mentioned above. Identifying markers were automatically placed in the EEG signal at the onset of each stimulus presentation.

Data were preprocessed offline using Brain Vision Analyzer 2.1 (Brain Products Inc.). Non-ocular artifacts were manually removed followed by the application of a 0.1-30 Hz bandpass filter (24 dB/oct) and a 60 Hz notch filter. Data were then re-referenced to the averaged mastoids and segmented according to the type of stimulus. All segments spanned a duration of 1200 ms starting from 200 ms before stimulus onset. Segments followed by an incorrect behavioural response (incorrect or absent button-press) were discarded from further analyses. Grand averages across the two experimental groups for each of the conditions can be observed in Fig. 3.1 showing an attenuated P300 response in the rCFL compared to the age-matched controls. Detailed description of data acquisition and a comprehensive investigation using traditional statistical analysis can be found in an earlier publication (Ruiter et al., 2019).

3.3 The Machine Learning Process

To evaluate how well group-level effects of mTBI on ERPs enable single-subject classification of concussed individuals, we begin our ML process by extracting a list of candidate features from the single-subject averaged ERPs. The candidate features are described in Section 3.3.1. In Section 3.3.2, we describe our process of selecting an optimal feature set from the list of candidate features and training a classification model with those features. We describe our process of evaluating our model in Section 3.3.3. Finally, we discuss our approach to interpreting the trained model in Section

3.3.5. All analyses were done in Python using the MNE package for EEG signal processing (Gramfort et al., 2013) and *scikit-learn* for machine learning (Pedregosa et al., 2011).

3.3.1 Feature Extraction

To perform single-subject detection of past mTBI, we computed a list of features from the single-subject averaged ERPs corresponding to our four stimulus conditions. We investigated four types of features that could be used to characterize these single-subject ERPs (a total of 472 features):

ERP Component Statistics

In traditional EEG/ERP analysis, one common way of characterizing an ERP is by measuring the peak amplitude and latency after stimulus onset at which the ERP occurs for each component. Here we followed our earlier work where we defined time windows and EEG electrode sites to identify the peak amplitude and latency for the following components (see (Ruiter et al., 2019)): N100 (negative peak at the Fz electrode between 75 and 125 ms after stimulus onset), N2b (negative peak at the Fz electrode between 170 and 270 ms), P3a (positive peak at the Cz electrode between 275 and 375 ms), and P3b (positive peak at the Pz electrode between 400 and 700 ms). A total of 32 candidate features (amplitude and latency for each of 4 ERP components $\times$ 4 stimulus conditions) per subject were extracted using this method.

Time-Domain features

In contrast to the ERP component statistics described above, we extracted time-domain statistics that are not based on traditionally-defined ERP components. Instead, we segmented each subject's averaged responses into ten equally-spaced, non-overlapping time bins. With reference to Fig. 3.1, in order to capture only the interval containing significant electrical activity, the first time bin started 50 ms after stimulus onset, the last time bin ended at 600 ms, and each time bin was 55 ms long. In addition, we sought to capture information on the spatial distribution of the ERPs, which we can see from Fig. 3.1 was altered in our concussed sample. Since the variation of the waveforms between closely spaced electrodes is low, we reduced the number of candidate features by defining clusters of electrodes proximate in location as *regions of interest* (ROIs; (Frishkoff et al., 2011)). We defined five ROIs by averaging the signals from the electrodes over the following locations: Midline-Central (M-C: Cz, C1, C2, FCz, FC1, and FC2), Midline-Parietal (M-P: Pz, P1, P2, CPz, CP1, and CP2), Midline-Occipital (M-O: Oz, POz, PO1, and PO2), Right-Parietal (R-P: P4, P6, CP4, and CP6), and Left-Parietal (L-P: P3, P5, CP3, and CP5). We denote the data from the nth subject within the jth window for a specific ROI r as the vector $\mathbf{x}_r(n,j) \in \mathbb{R}^W, n = 1,\ldots,N, \quad j = 1,\ldots,J$. The respective time domain features were then formed by evaluating the mean of the data in each $\mathbf{x}_r(n,j)$. A total of 10 (J) non-overlapping bins of size W were evaluated starting at 50 ms post stimulus onset. W was selected to represent 55 ms of data (i.e., 28 samples). A total of 200 candidate features (5 ROIs $\times$ 10 windows $\times$ 4 stimulus conditions) were extracted for each subject using this method.

Global Field Power

Since ERPs are generated by highly synchronous brain activation that propagates widely through the brain, one way to characterize its global activation pattern is through its global field power (GFP). GFP(t) is the time series of the power averaged over all channels, and is calculated as follows. Let the signal received from the ith electrode at time t be $x_i(t)$. Then the GFP at time t is the spatially averaged root mean square of $x_i(t)$. It is defined as (Lehmann and Skrandies, 1980)

$$GFP(t) = \sqrt{\frac{\sum_{i=1}^{C} \left(x_i(t) - \bar{x}(t)\right)^2}{C}} \tag{3.1}$$

where C is the total number of electrodes and $\bar{x}(t)$ is the average value over the C electrodes at time t. GFP was averaged for each of the 10 time bins as described above. Previous work showed that the GFP can be used to characterize an ERP's general strength and latency, and that a high GFP can be interpreted as a large evoked response (Lehmann and Skrandies, 1980). Moreover, since the GFP is calculated from the spatial distribution of the EEG, it provided us with a relatively low dimensional feature space for model training. A total of 40 candidate features per subject were extracted using this method.

Wavelet Transform

The continuous wavelet transform (WT) has been used in previous work to characterize the shape of ERPs (Demiralp et al., 2001; Quiroga et al., 2001). We used a Morlet wavelet with a 3 Hz central frequency (to capture the overall ERP morphology) and one cycle per window (allowing for a broader spread of frequency components) to create a time-frequency representation of the ERPs. We then extracted averaged wavelet power coefficients corresponding to the same ROIs and time bins as described in Section 3.3.1. A total of 200 candidate features per subject were extracted using this method.

3.3.2 Feature Selection, Classification, and Validation

We performed a two-stage feature selection process in a nested cross-validation loop as presented in Fig. 3.2. The dataset of 39 subjects was randomly split into outer training and test subsets which we denote as Tr_o and Tt_o, respectively. Tr_o contained samples from 31 of the subjects, and Tt_o the remaining 8 subjects (drawn without replacement). In each outer iteration, we performed 100 inner loop iterations in which Tr_o were randomly partitioned in the same way to obtain a training set of 23 subjects Tr_i and a validation set of eight subjects Tt_i. We ensured that no two random partitions of the data are the same across cross-validation loops. On each inner loop iteration, we used the F-score univariate feature selection method (as implemented in scikit-learn; (Pedregosa et al., 2011)) to select the top 75 features. We used those 75 features to train a linear SVM classifier on Tr_i. We aggregated a list of the selected features for iterations which yielded a classification accuracy statistically above chance on Tt_i (64.1% threshold based on a binomial test; described below). After the 100 inner loop iterations were complete, we selected the top 50% of the aggregated features and pass these to the outer loop. These features are used to train a linear SVM on Tr_o. The last step of the outer loop involves testing the resulting trained model using only Tt_o. This process was repeated for 100 outer loop iterations (see Fig. 3.2). The linear SVM hyperparameter was set to its default value ($C = 1$)

for all iterations. Note that this two-stage feature selection is valid as both stages of feature selection were performed independently of the set Tt_o. The aggregated feature selection process described here improves the statistical stability and reduces the effect of overfitting with regard to the feature selection procedure.

Accuracy for the concussed class is assessed by evaluating the number of predicted concussed subjects to the total number of concussed subjects in the training set. Accuracy for the control class is evaluated in a corresponding manner.

With respect to the derivation of the 64.1% threshold used above, consider a binary classification experiment where n_1 and n_2 are the numbers of training samples in classes 1 and 2 respectively. Then if the labels corresponding to each subject are randomly permuted, the classification accuracy becomes a random variable, with an expected value for class 1 of $n_1/(n_1+n_2) \times 100\%$, and correspondingly $n_2/(n_1+n_2) \times 100\ \%$ for class 2. In the present case where $n = 39$, the upper limit of the respective 95% interval is 64.1% (Combrisson and Jerbi, 2015).

3.3.3 Secondary Model Validation

In order to confirm that the nested cross-validation process we utilized was robust to overfitting, we conducted a simulation using random data. Performance was evaluated on a 39-sample dataset of 1000 randomly generated features using the procedure in Fig. 3.2. Our procedure on the random dataset yielded a classification accuracy of 51.9% (SD = 16.9). If overfitting was present, the accuracy would be higher than this 51.9% figure.

Due to the relatively small number of subjects in the present study, there may be some concern over the statistical stability of the results. Thus as an additional verification measure, a secondary cross-validation procedure was performed independently of the main performance evaluation cross-validation where class labels were randomly permuted across all subjects, and accuracy was assessed. The probability that the result from the evaluation cross-validation was drawn from the probability density function of the secondary run where the labels were permuted, was assessed. Since the distributions are approximately Gaussian given a large number of trials, this test was performed using a two-tailed t-test with unequal variances. A low probability indicated that the results produced by the main run were unlikely to be due to chance alone. For a statistical validation of the permutation-test for validating classification accuracies in low-sample settings, see (Ojala and Garriga, 2010). Results are reported in section IV.

3.3.4 Subject Misclassification

During cross-validation, a model's estimated probability of a subject belonging to the concussed class is extracted for all subjects in the test set. Probabilities were tallied when the model achieved above-chance accuracy (64.1%) where a probability of 1 signified certainty of an identified concussed subject and a 0 signified the certainty of a control subject. A frequently misclassified concussed subject was defined as having an aggregate probability below 0.5. Similarly, a misclassified control subject scored an aggregate probability above 0.5. We performed post-hoc examination of the demographics and behavioural data for frequently misclassified subjects to check for the presence of any factors that may explain the model's error.

3.3.5 Model Interpretation

A commonly cited downside with the application of ML to healthcare and diagnostic applications is that most ML models are difficult to interpret (Miotto et al., 2017). The tendency of viewing trained models as black boxes poses difficulties in a clinical setting since ML decisions cannot be checked against the body of available clinical knowledge. Several solutions have been proposed to provide insight on how a trained model makes its judgments (Ribeiro et al., 2016; Lundberg and Lee, 2017). SHapley Additive exPlanations (SHAP) have been particularly successful, theoretically unifying a number of previously established tools with demonstrated clarity and consistency (Lundberg and Lee, 2017). The SHAP value of a particular feature indicates the effect on the model's prediction when that feature is omitted. SHAP has recently been used in the clinical setting to provide online predictions on whether a patient was at risk of a surgical complication (Lundberg et al., 2018). Particularly, the application presented a detailed explanation of why it made its predictions by relating them to input features.

To evaluate our SVM models, we utilized the kernel SHAP implementation. Kernel SHAP is model-agnostic, and constructs simple, localized models using weighted linear regression to evaluate local feature attribution of a trained model's classification decision (Lundberg and Lee, 2017). The outcome of the simple SHAP model on a trained model is termed an *explainer*. Only the 25 most-selected features (Section 3.3.2) were considered in the SHAP evaluation. To achieve the best approximation across our entire dataset, we trained 39 explainers in a LOO procedure where data from 38 subjects were used for training to explain the remaining subject. Rank and *directionality* were extracted from the averaged SHAP values over all explainers.

Rank was determined by ordering the mean absolute SHAP values from each feature. In addition, the sign (positive or negative) of a SHAP value was used to determine the *directionality* of the respective feature. Directionality of a feature indicates which class becomes more likely when that feature's value is increased. Furthermore, to investigate subjects from the commonly misclassified group, a subject's explainer (based on all but this subject's own data) was used to infer the features' influence in the model's incorrect classification. Note that SHAP was used here only as a tool for interpreting the model's behaviour after feature selection and model training were complete.

3.4 Results

Our final model achieved an average accuracy of 81% (SD=11.5) with a sensitivity of 82% and specificity of 80%. Accuracy was significantly higher than chance based on the permutation test ($p < 0.001$).

We varied the number of features selected by the F-score method in the inner cross-validation loop to experimentally choose an optimal setting. Peak performance was achieved when 75 features were selected on the first stage of feature selection. Model performance with respect to number of features chosen in the inner cross-validation loop is shown in Fig. 3.3.

The 25 most selected features are summarized using SHAP in Fig. 3.4. Most of the top features were extracted from ERPs that occurred in response to all experimental conditions (Std, FDev, DDev, and IDev), suggesting that each of these responses are altered in concussed subjects. Wavelet features were most prominent forming 16 of the 25 features used for explanation. Low feature values tended to correlated with higher association with the concussed group (e.g., GFP 215-270 DDev and WT 380-435 from FDev at L-P). This finding is consistent with previous reports showing attenuated ERP responses to deviants after concussion (Broglio et al., 2011; Ruiter et al., 2019). Interestingly, several features had an opposite correlation. For instance, larger feature values in the WT of the Std at 105-160 ms as seen in the M-P ROI and the WT from FDev at 435-490 ms in the M-O ROI directed the model towards a concussed group classification. This finding is unreported in the literature and is of particular interest both due to its direction and the involvement of the responses to Std as discriminatory features.

3.4.1 Misclassified Subjects

Five subjects were most commonly misclassified: three concussed (rCFL 1, 4, and 11) and two controls (Control 5 and 20). Post-hoc inspection of the demographics and symptomatology yielded no observable relationship indicative of the machine learning results.

The subject averages for the misclassified subjects at the Pz electrode are shown in Fig. 3.5. Aided by the grand–averaged waveforms (Fig. 3.1) and *a priori* knowledge from the EEG/ERP literature, incorrect classification of rCFL 4 and Control 5 may be attributable to responses atypical to their respective groups. However, results for the other three subjects (rCFL 1, 11 and Control 20) were not visually explainable.

Next, we examined the SHAP feature estimates for all misclassified subjects. Note that we trained the explainers using the 25 most common features (Fig. 3.4) only instead of all selected features. Therefore, the explainers only represent estimates of the main model's results. The model decision for each misclassified patient was described by SHAP as follows:

rCFL 1 Subject's SHAP explainer highlighted two main features for the incorrect classification: the wavelet features in the Std condition between 160-215 ms for the M-C and M-P regions. SHAP indicated that the features were larger than what is commonly associated with the concussed group.

rCFL 4 SHAP highlighted the mean amplitudes from Std between 215-270 ms from the M-O region as larger than expected in the concussed group. Additionally, the mean response for the DDev condition between 380-435 at M-P was larger than the concussed group's.

rCFL 11 SHAP indicated that this subject's wavelet feature from the Std condition at 325-380 ms from the M-O region was larger than the group's expected value.

Control 5 Incorrect classification was attributable to the mean feature from FDev between 380-435 m at the L-P region and the wavelet feature from Std between 215-270 ms at the L-P region being smaller than expected in the control group.

Control 20 Two features influenced the explainer most towards an incorrect classification. Both the GFP of Std from 215-270 ms and wavelet feature of Std from 160-215 ms at the M-C region were smaller than the models' definition of the control group.

3.5 Discussion

In this paper we demonstrated the use of ML as a methodological approach to utilizing neurophysiological markers found in group-level studies for single-subject detection of mTBI. Using our exploratory approach, we identified a set of features that enabled accuracy in detecting past concussion of up to 81%. Interestingly, we were able to correctly classify past concussion with high accuracy even though our concussed individuals received their last injury up to 45 years ago (Mean = 28), far beyond the suggested time-frame of a few months for symptom resolution (McCrory et al., 2017). Additionally, we conducted an interpretability analysis using SHAP to gain insight into our model's decision-making, which can be used by health professionals if machine learning tools like the one described here are adopted as diagnostic aids in the future.

The present study is the first report of ML-based EEG/ERP analysis for the assessment and identification of concussion history decades after injury. We reported a higher accuracy than previous studies classifying mTBI using RS EEG (Cao et al., 2008; Prichep et al., 2012). Notably, the two prior studies investigated acute and post-acute effects of injury on RS EEG, which were previously argued to normalize within six months after insult (Nuwer et al., 2005). A quantitative comparison with clinical tools typically used in mTBI assessment is difficult when investigating chronic effects due to a lack of consensus and standardization in the clinical literature (Broglio et al., 2017; McCrory et al., 2017); however, in comparison to identification in acute injury, the presented methods report a significant improvement on the sensitivity of individual clinical tools such as self-reported symptoms (68.0%), postural control evaluation (61.9%), and a brief pen-and-paper assessment (43.5%) (Broglio et al., 2007).

Our feature extraction methods were primarily guided by domain knowledge, which confirmed that group–level effects are transferable and indicative of underlying single–subject effects. However, our inspection of the most important features using SHAP, as summarized in Fig. 3.4, provided additional insights that may not be obvious to a domain expert. Interestingly, our results indicate that the responses to the standard condition, particularly in the N100-P200 time span, carried substantial discriminative information, making up 12 of the 25 top features. These responses are commonly discarded or used strictly to calculate ERP difference waves (i.e., by subtracting the standard condition from the averaged ERP generated by a deviant condition; see (Ruiter et al., 2019; Broglio et al., 2009)).

Of particular interest, this study is the first to report an attenuated P200 response in concussion to the standard tone. This is especially seen in the GFP at 215-270 ms and in the grand-averages (see Fig. 3.4 and Fig. 3.1 respectively). This effect was unobserved in the traditional group-level analysis previously conducted on the same data (Ruiter et al., 2019). In addition, features from responses to all four experimental conditions were selected by the trained models, demonstrating the importance of a multi-deviant design. Finally, the absence of traditionally-defined ERP amplitude and latency values from the list of top selected features is noteworthy. This finding suggests that traditional ERP analysis is insufficient to capture the individual variability required for single-subject applications.

The coupling between neurophysiological deficits and results from behavioural assays has continuously been questioned in the literature (Broglio et al., 2017, 2007).

Some studies show that neurophysiological changes were present despite the lack of observable symptoms (Cao et al., 2008; De Beaumont et al., 2012), while other studies show a co–occurrence between the two (Ruiter et al., 2019; Gosselin et al., 2012). A current prominent explanation for the persistence of anomalies in brain activity despite symptom recovery is termed the "cognitive reserve" theory, stating that individuals with a history of concussive impacts recruit additional brain resources to accomplish cognitive tasks (De Beaumont et al., 2012; McAllister et al., 2001). As the individual ages, the brain becomes less able to sustain the re-allocation of resources to compensate for the cognitive effects of mTBI, resulting in the decline often observed in subjects with a history of concussion (De Beaumont et al., 2012). Our results fit within a "cognitive reserve" explanation, although a lack of emerging patterns between the commonly misclassified subjects and their demographics, symptomatology and neurophysiological signs suggest a complex relation that varies between subjects.

Our study exhibits some limitations. First, the small sample size in our study (39 subjects) limits our ability to generalize to the broader population of concussed individual. Our sample size is representative of the studies published in the field, including for machine learning research (Cao and Slobounov, 2011); however, it is particularly relevant in a machine learning context, since classification in a high dimensional space with only a few examples can reduce the generalizability of a model. We minimized the risk of overfitting with significant two-stage dimensionality reduction prior to fitting the classification model, and by using robust statistical methods to ensure that our classification accuracy was well above chance. In addition, we can qualitatively assess that our model leveraged features that align with what is known from the neuroscientific literature, lending confidence to its generalizability. Another limitation is the lack of a concrete ground truth, as the current understanding of mTBI remains limited. As such, our labels were based on the very identification techniques argued to be flawed (Broglio et al., 2017, 2007). Lastly, most of our subjects were not clinically diagnosed at the time of injury and thus, we were dependent on self-reporting – lack of diagnosis was expected given the date of injury for most subjects was before concussion was regarded as a serious concern (Broglio et al., 2017). This introduced a likelihood of incorrect labels and consequently a reduced performance for our model. We argue that this confound puts a ceiling on attainable accuracy in the present study.

3.6 Conclusion and Future Work

In the present study, we investigated how mTBI–induced ERP changes discovered in group-level analyses translate to single-subject identification tools using machine learning methodology. We demonstrated that features reflecting changes in the N100, P200, N2b, and P300 ERPs following mTBI yield a classification accuracy of 81% with a sensitivity of 82% and a specificity of 80% using a ML model. The features found to be most discriminative for diagnosis of mTBI strongly resemble the long-lasting changes in P300s typically found after mTBI (Ruiter et al., 2019; Broglio et al., 2009, 2007; De Beaumont et al., 2007b,a). We thus demonstrate that what has been discovered through group-level ERP analysis in the neuroscientific literature provides some predictive power in the context of mTBI. However, we could not correctly identify past mTBI in a particular subset of individuals on the basis of *a priori* neuroscientific knowledge. Our results demonstrate the need go beyond what has been discovered in group-level analyses in order to identify a complete set of features required to capture the heterogeneity observed in the mTBI population. Finally, we emphasize the utility of explainable models in clinical applications, highlighting a novel ERP finding that was unobservable using previous analyses. In sum, although extensive future work is required to both validate and refine our results to enable a realized application in the clinical setting, we have made a significant step in the identification of chronic effects of concussion in single-subjects.

3.7 Acknowledgements

The authors would like to thank Mr. Steve Buist and the Hamilton Spectator for their support and help in the recruitment process. This study was supported by The Hamilton Spectator, Canada Foundation for Innovation (JFC), Senator William McMaster Chair in Cognitive Neuroscience of Language (JFC), Natural Sciences & Engineering Research Council of Canada (NSERC; OB, JPR, RS, and TD), Southern Ontario Smart Computing Innovation Platform (SOSCIP; OB, RS, and TD), and the Ontario Ministry of Research and Innovation (RB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

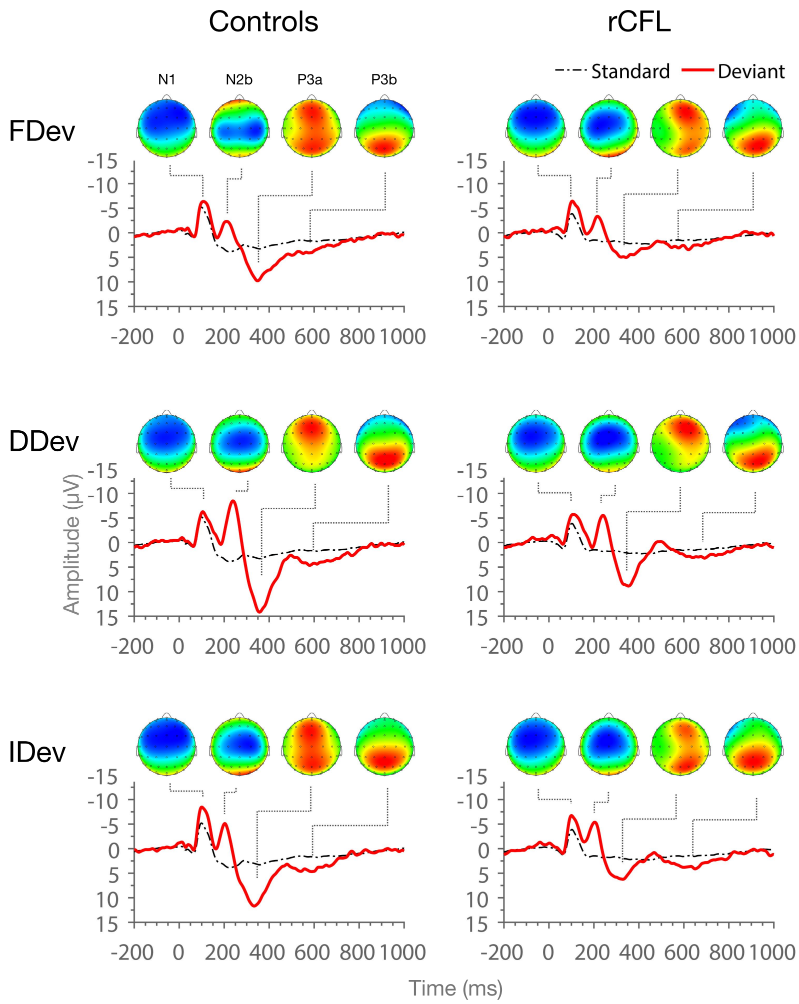

Figure 3.1: Grand averages as recorded from the Cz electrode and relevant ERP topographies across the two groups for the Frequency Deviant (FDev), Duration Deviant (DDev), and Intensity Deviant (IDev). Dotted waveforms represent group responses to standard tones (Std). Adapted with permission from (Ruiter et al., 2019).

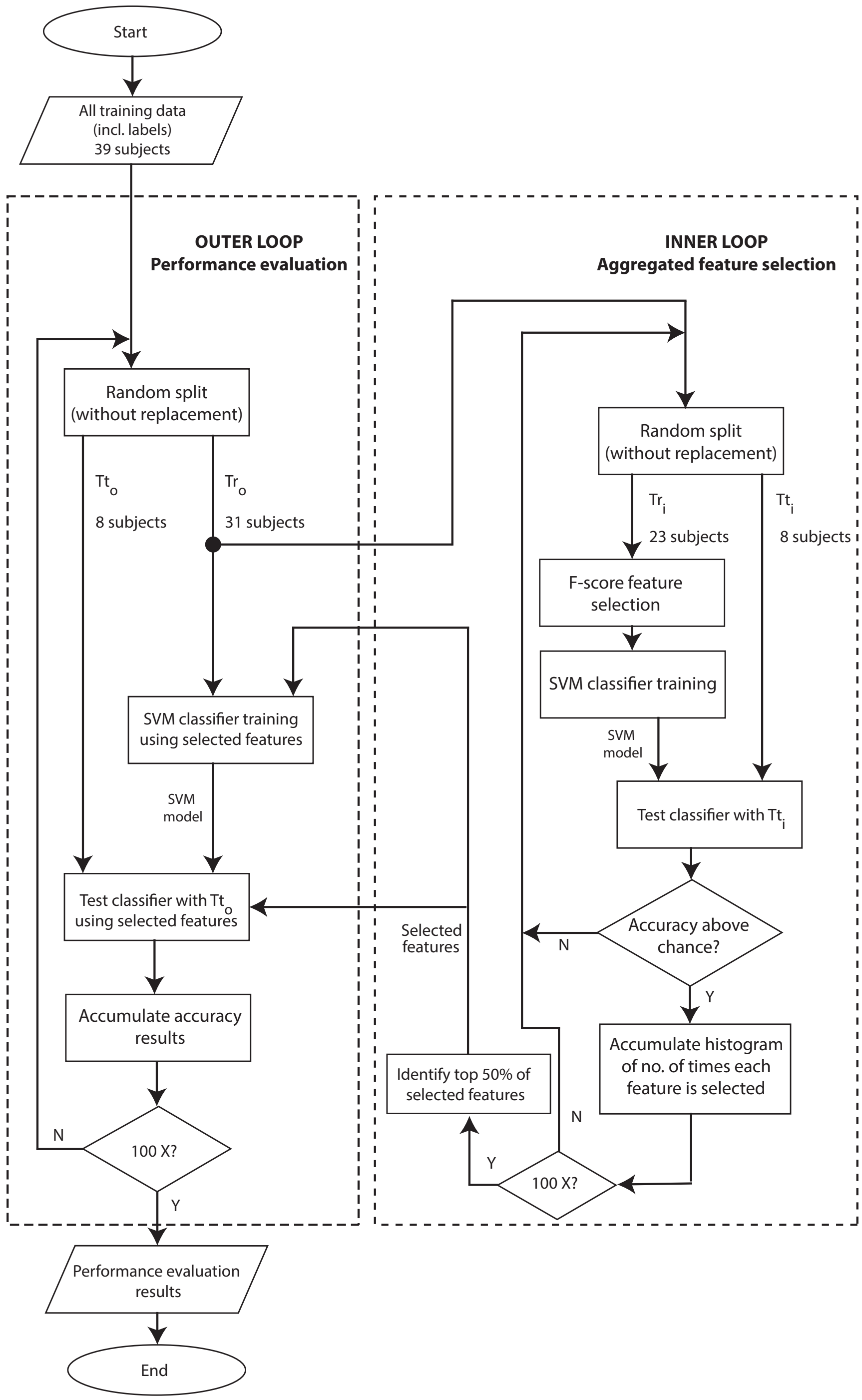

Figure 3.2: A flowchart outlining the overall machine learning procedure used in this study.

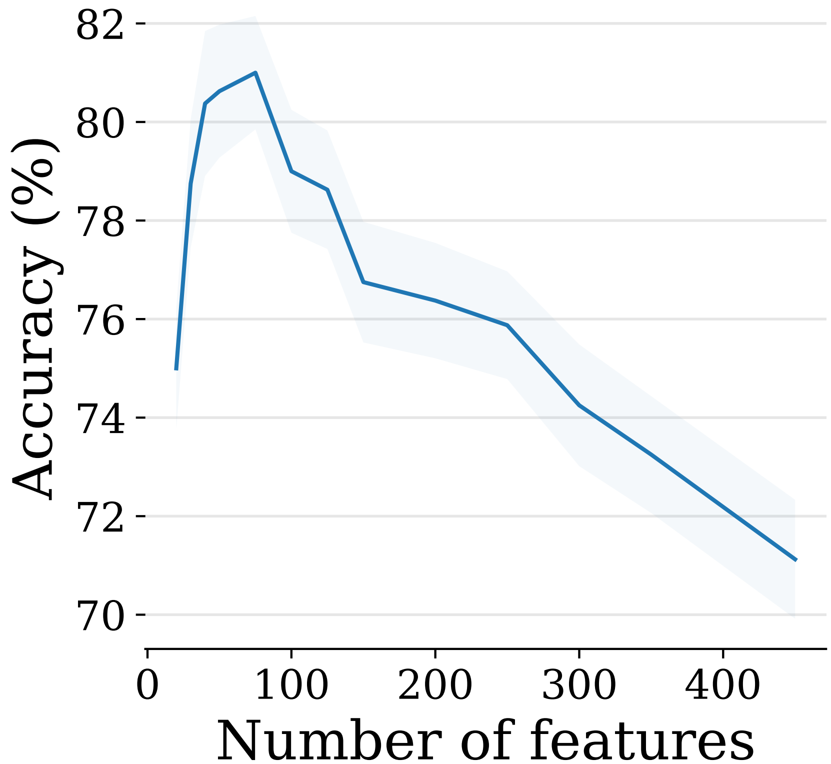

Figure 3.3: Average classification accuracy vs. the number of selected features. Shaded region indicates the standard error of the mean across the cross-validation steps.

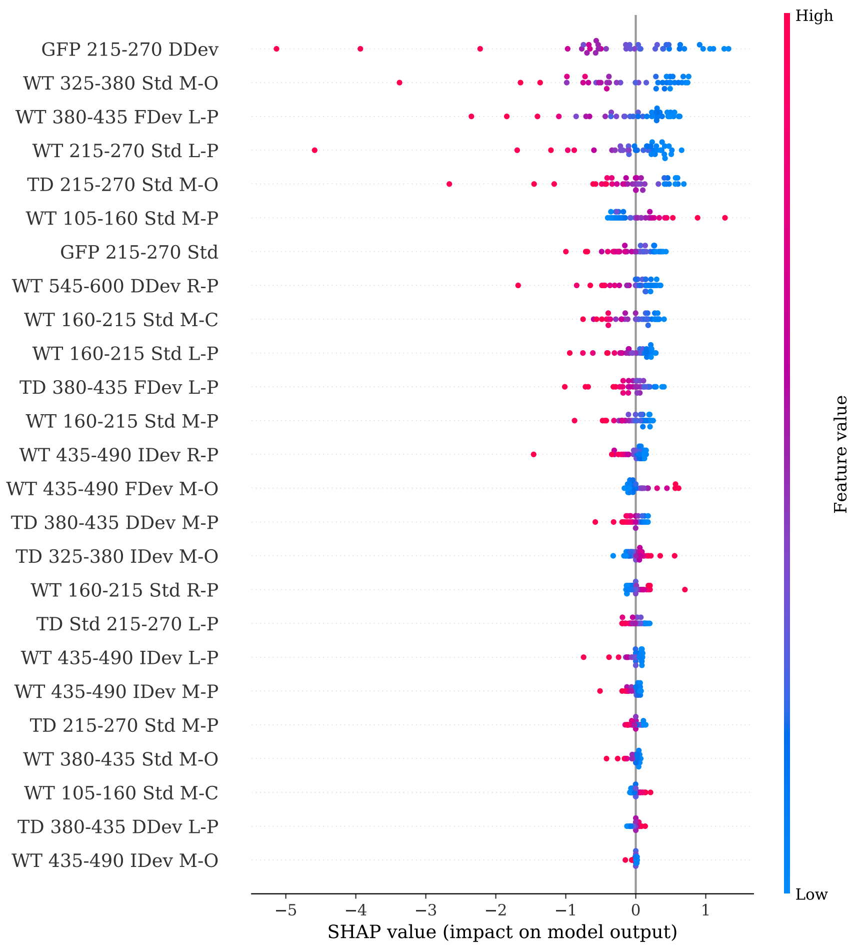

Figure 3.4: The SHAP values of all subjects for the 25 most-used features. Features are ranked top to bottom (top being the highest ranked). A single point represents a subject's SHAP value for a corresponding feature (ordinate). A positive (negative) SHAP value indicates the feature's impact towards classifying a subject as concussed (control). Color indicates the true value of each feature, as opposed to the derived SHAP value, from blue (low) to red (high; see color bar on the right). Combined with the distribution on the abscissa, the feature values (color) for all subjects indicates the directionality effect of a particular feature.

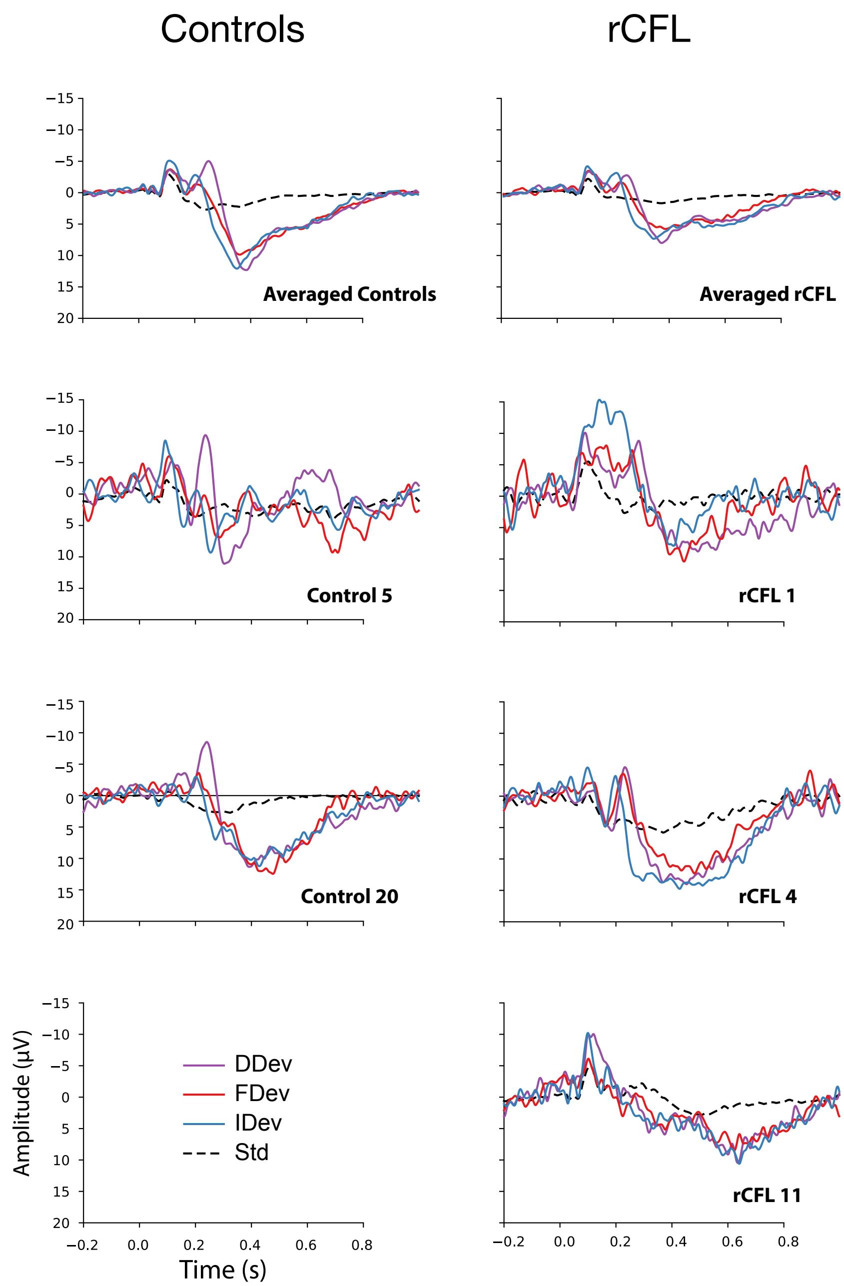

Figure 3.5: Subject averages of responses to all experimental conditions of the five commonly misclassified subjects. The averaged response for the two groups is presented in the first row. Waveforms represent data as recorded from the Pz electrode. Figure legend presented in the bottom left corner.

Bibliography

Armanfard, N., Komeili, M., Reilly, J. P., and Connoly, J. (2018). A machine learning framework for automatic and continuous mmn detection with preliminary results for coma outcome prediction. *IEEE journal of biomedical and health informatics*.

Blankertz, B., Acqualagna, L., Dähne, S., Haufe, S., Schultze-Kraft, M., Sturm, I., Uščumlic, M., Wenzel, M. A., Curio, G., and Müller, K. R. (2016). The Berlin brain-computer interface: Progress beyond communication and control. *Frontiers in Neuroscience*, 10(NOV).

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Macciocchi, S. N., and Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, 60(6):1050–7; discussion 1057–8.

Broglio, S. P., Moore, R. D., and Hillman, C. H. (2011). A history of sport-related concussion on event-related brain potential correlates of cognition. *International Journal of Psychophysiology*, 82(1):16–23.

Broglio, S. P., Pontifex, M. B., O'Connor, P., and Hillman, C. H. (2009). The Persistent Effects of Concussion on Neuroelectric Indices of Attention. *Journal of Neurotrauma*, 26(9):1463–1470.

Brush, C. J., Ehmann, P. J., Olson, R. L., Bixby, W. R., and Alderman, B. L. (2018). Do sport-related concussions result in long-term cognitive impairment? A review of event-related potential research. *International Journal of Psychophysiology*, 132(March 2017):124–134.

Cao, C. and Slobounov, S. (2010). Alteration of cortical functional connectivity as a result of traumatic brain injury revealed by graph theory, ICA, and sLORETA analyses of EEG signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(1):11–19.

Cao, C. and Slobounov, S. (2011). Application of a novel measure of EEG non-stationarity as 'Shannon- entropy of the peak frequency shifting' for detecting residual abnormalities in concussed individuals. *Clinical Neurophysiology*, 122(7):1314–1321.

Cao, C., Tutwiler, R. L., and Slobounov, S. (2008). Automatic classification of athletes with residual functional deficits following concussion by means of EEG signal using support vector machine. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(4):327–335.

Chan, A. M., Halgren, E., Marinkovic, K., and Cash, S. S. (2011). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *NeuroImage*, 54(4):3028–3039.

Combrisson, E. and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, 250:126–136.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*.

De Beaumont, L., Brisson, B., Lassonde, M., and Jolicoeur, P. (2007a). Long-term electrophysiological changes in athletes with a history of multiple concussions. *Brain Injury*, 21(6):631–644.

De Beaumont, L., Henry, L. C., and Gosselin, N. (2012). Long-term functional alterations in sports concussion. *Neurosurgical Focus*, 33(6):E8.

De Beaumont, L., Lassonde, M., Leclerc, S., and Théoret, H. (2007b). Long-term and cumulative effects of sports concussion on motor cortex inhibition. *Neurosurgery*, 61(2):329–336.

Demiralp, T., Ademoglu, A., Istefanopulos, Y., Başar-Eroglu, C., and Başar, E. (2001). Wavelet analysis of oddball p300. *International journal of psychophysiology*, 39(2-3):221–227.

Dhindsa, K. and Becker, S. (2017). Emotional reaction recognition from EEG. In *2017 International Workshop on Pattern Recognition in Neuroimaging, PRNI 2017*.

Dhindsa, K., Carcone, D., and Becker, S. (2017). Toward an Open-Ended BCI: A User-Centered Coadaptive Design. *Neural Computation*, 29(10):2742–2768.

Fickling, S. D., Smith, A. M., Pawlowski, G., Ghosh Hajra, S., Liu, C. C., Farrell, K., Jorgensen, J., Song, X., Stuart, M. J., and D'Arcy, R. C. (2019). Brain vital signs detect concussion-related neurophysiological impairments in ice hockey. *Brain*, 142(2):255–262.

Frishkoff, G., Frank, R., Sydes, J., Mueller, K., and Malony, A. (2011). Minimal information for neural electromagnetic ontologies (mi-nemo): A standards-compliant workflow for analysis and integration of human eeg. *Standards in Genomic Sciences (SIGS)*, 5(2).

Gosselin, N., Bottari, C., Chen, J.-K., Huntgeburth, S. C., De Beaumont, L., Petrides, M., Cheung, B., and Ptito, A. (2012). Evaluating the cognitive consequences of mild traumatic brain injury and concussion by using electrophysiology. *Neurosurgical Focus*, 33(6):E7.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*.

Harmon, K. G., Drezner, J. A., Gammons, M., Guskiewicz, K. M., Halstead, M., Herring, S. A., Kutcher, J. S., Pana, A., Putukian, M., and Roberts, W. O. (2013). American medical society for sports medicine position statement: concussion in sport. *Br J sports med*, 47(1):15–26.

Henry, L. C., Tremblay, S., and De Beaumont, L. (2017). Long-Term Effects of Sports Concussions: Bridging the Neurocognitive Repercussions of the Injury with the Newest Neuroimaging Data. *Neuroscientist*, 23(5):567–578.

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.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., and Chen, J. H. (2010). EEG-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 57(7):1798–1806.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F. (2018). A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces: A 10-year Update. *Journal of Neural Engineering*, pages 0–20.

Lundberg, S. and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems*, 16(3):426–430.

Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., Kim, J., and Lee, S.-I. (2018). Explainable machine learning predictions to help anesthesiologists prevent hypoxemia during surgery. *Nature Biomedical Engineering*.

McAllister, T. W., Sparling, M. B., Flashman, L. A., and Saykin, A. J. (2001). Neuroimaging Findings in Mild Traumatic Brain Injury *. *Journal of Clinical and Experimental Neuropsychology*, 23(6):775–791.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R. C., Cassidy, D., Echemendia, R. J., Castellani, R. J., Davis, G. A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C. C., Guskiewicz, K. M., Herring, S., Iverson, G. L., Johnston, K. M., Kissick, J., Kutcher, J., Leddy,

J. J., Maddocks, D., Makdissi, M., Manley, G. T., McCrea, M., Meehan, W. P., Nagahiro, S., Patricios, J., Putukian, M., Schneider, K. J., Sills, A., Tator, C. H., Turner, M., and Vos, P. E. (2017). Consensus statement on concussion in sport—the 5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, (October 2016):bjsports–2017–097699.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., and Cantu, R. (2009). Consensus Statement on Concussion in Sport – The Third International Conference on Concussion in Sport Held in Zurich, November 2008. *The Physician and Sportsmedicine*, 37(2):141–159.

McCrory, P., Meeuwisse, W. H., Aubry, M., Cantu, B., Dvořák, J., Echemendia, R. J., Engebretsen, L., Johnston, K., Kutcher, J. S., Raftery, M., Sills, A., Benson, B. W., Davis, G. A., Ellenbogen, R. G., Guskiewicz, K., Herring, S. A., Iverson, G. L., Jordan, B. D., Kissick, J., McCrea, M., McIntosh, A. S., Maddocks, D., Makdissi, M., Purcell, L., Putukian, M., Schneider, K., Tator, C. H., and Turner, M. (2013). Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, 47(5):250–258.

Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, (February):1–11.

Munia, T. T., Haider, A., Schneider, C., Romanick, M., and Fazel-Rezai, R. (2017). A Novel EEG Based Spectral Analysis of Persistent Brain Function Alteration in Athletes with Concussion History. *Scientific Reports*, 7(1):1–13.

Nuwer, M. R., Hovda, D. A., Schrader, L. M., and Vespa, P. M. (2005). Routine and quantitative EEG in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025.

Ojala, M. and Garriga, G. C. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research*, 11(Jun):1833–1863.

Parvar, H., Sculthorpe-Petley, L., Satel, J., Boshra, R., D'Arcy, R. C., and Trappenberg, T. P. (2015). Detection of event-related potentials in individual subjects using support vector machines. *Brain Informatics*, 2(1):1–12.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of machine learning research*, 12:2825–2830.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b.

Prichep, L. S., Jacquin, A., Filipenko, J., Dastidar, S. G., Zabele, S., Vodencarevic, A., and Rothman, N. S. (2012). Classification of Traumatic Brain Injury Severity Using Informed Data Reduction in a Series of Binary Classifier Algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(6):806–822.

Quiroga, R. Q., Rosso, O. A., Başar, E., and Schürmann, M. (2001). Wavelet entropy in event-related potentials: a new method shows ordering of eeg oscillations. *Biological cybernetics*, 84(4):291–299.

Ravan, M., Hasey, G., Reilly, J. P., MacCrimmon, D., and Khodayari-Rostamabad, A. (2015). A machine learning approach using auditory odd-ball responses to investigate the effect of clozapine therapy. *Clinical Neurophysiology*, 126(4):721–730.

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135—-1144.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Sculthorpe-Petley, L., Liu, C., Ghosh Hajra, S., Parvar, H., Satel, J., Trappenberg, T. P., Boshra, R., and D'Arcy, R. C. (2015). A rapid event-related potential (ERP) method for point-of-care evaluation of brain function: Development of the Halifax Consciousness Scanner. *Journal of Neuroscience Methods*, 245:64–72.

Todd, J., Michie, P. T., Schall, U., Karayanidis, F., Yabe, H., and Näätänen, R. (2008). Deviant Matters: Duration, Frequency, and Intensity Deviants Reveal Different Patterns of Mismatch Negativity Reduction in Early and Late Schizophrenia. *Biological Psychiatry*, 63(1):58–64.

Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., and De Lucia, M. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, 136(1):81–89.

CHAPTER 4

Neurophysiological Correlates of Concussion: Deep Learning for Clinical Assessment

Preface

The chapter includes a follow-up study to the previous chapter's, detailing a ML investigation of ERP effects in participants with more recent injuries. Due to the availability of a larger dataset, the previous ML tools were expanded to a more fine-grained single-trial analysis utilizing an adapted CNN architecture. The study presents a full automated extraction of features, as opposed to engineered ones in the previous chapter, and achieved high accuracies. Further, analyses of a longitudinal subset of concussed participants provided additional evidence for a dissociation between symptomatology and neurophysiological signs of concussion. The chapter is a reformatted from a manuscript prepared for submission in Scientific Reports with the following authors list:

Boshra, R., Ruiter, K. I., DeMatteo C., Reilly, J. P., & Connolly, J. F.

Abstract

Concussion has been shown to leave the afflicted with significant cognitive and neurobehavioural deficits. The persistence of these deficits and their link to neurophysiological indices of cognition, as measured by event-related potentials (ERP) using electroencephalography (EEG), remains restricted to population level analyses that limit their utility in the clinical setting. In the present paper, a convolutional neural network is extended to capitalize on characteristics specific to EEG/ERP data in order to assess for post-concussive effects. An aggregated measure of single-trial performance was able to classify accurately (85%) between 26 acutely to post-acutely concussed participants and 28 healthy controls in a stratified 10-fold cross-validation design. Additionally, the model was evaluated in a longitudinal subsample of the concussed group to indicate a dissociation between the progression of EEG/ERP and that of self-reported inventories. Our results form a first-step towards the clinical integration of neurophysiological results in concussion management and motivate a multi-site validation study for a concussion assessment tool in acute and post-acute cases.

Introduction

Traumatic brain injury (TBI) impacts upwards of 2.8 million individuals annually in the united states alone (Taylor et al., 2017). Concussions (henceforth used synonymously with mild TBI; mTBI) form a considerable subset of that figure and are defined as closed-head injuries that leave the affected with functional and cognitive deficits (McCrory et al., 2017; Langlois et al., 2006). The current understanding of underlying mechanisms in concussion remains lacking, with echoing concerns both in the identification and management of the condition (Broglio et al., 2017). An expansive body of work has targeted the multiple facets of concussion, offering different means of elucidating the cognitive deficits caused by concussion and its co-morbid sequelae (Brush et al., 2018). Electrophysiology is one tool with promising applications in concussions. Specifically, event-related potentials (ERPs) as recorded by electroencephalography (EEG) have shown persistent changes in concussed individuals in the post-acute stage and decades after insult (Gosselin et al., 2010, 2012; De Beaumont et al., 2007; Broglio et al., 2011; Ruiter et al., 2019).

ERPs are non-invasively-recorded indices of cognitive function (Duncan et al., 2009). The P300, a positive-deflecting response peaking approximately 300 ms after stimulus onset, is a commonly studied component in neurophysiology that is associated with attentional resource allocation, orientation, and memory (Polich, 2007). The P300 was found to be impacted by concussion immediately after occurrence and decades post injury (Fickling et al., 2019; Gosselin et al., 2010; De Beaumont et al., 2007; Ruiter et al., 2019; Broglio et al., 2011). P300 effects were observable when patients were symptomatic as well as after symptom resolution and were affected cumulatively following a series of concussive blows to the head in comparison to a single hit (Gosselin et al., 2006; Gaetz et al., 2000). The N2b is an ERP often linked to executive function manifesting as a fronto-central negative deflection 200 ms after stimulus onset (Folstein and Van Petten, 2008). Similar to the P300, the N2b was affected after sustaining hits to the head (Broglio et al., 2009; Gaetz et al., 2000; Gosselin et al., 2012; Moore et al., 2014; Ruiter et al., 2019). Research has demonstrated the versatility and sensitivity of both the P300 and N2b to concussion; however, a transition from controlled, group-level findings to individual assessment is required before clinical adoption is made feasible.

Machine learning (ML) has gained significant traction in the clinical field, offering a cost-efficient way of replicating expert judgements and decisions in a setting overloaded with data (Obermeyer and Emanuel, 2016; Rajkomar et al., 2019). ML introduces a dynamic process that is able to ingest high-dimensional clinical data and learn complex patterns that might also be difficult to detect or visualize for a human expert (Obermeyer and Emanuel, 2016; Rajkomar et al., 2019). Despite some scrutiny due to black-box solutions and susceptibility to bias in misapplication, machine learning remains a great tool for exploiting resources to improve clinical standards (Chen and Asch, 2017; Miotto et al., 2017; Obermeyer and Emanuel, 2016; Lundberg et al., 2018). EEG data are characterized by their rich high-dimensionality that requires certain degrees of aggregation to simplify for a human observer – quite possibly at the cost of losing critical information. That complexity has made ML a valuable method in EEG analysis (Boshra et al., 2019; Tzovara et al., 2013; Cao et al., 2008; Lawhern et al., 2018; Schirrmeister et al., 2017; Cecotti and Gräser, 2011; Opałka et al., 2018; Sturm et al., 2016; Connolly et al., 2019).

In the present study, we developed the TRauma ODdball Net (TRODNet), a deep learning network that uses convolutional layers in extracting information from single-trial EEG/ERP data to identify signs of concussion. The network learns a set of topographical maps that characterize different ERPs elicited in a multi-deviant oddball paradigm designed to elicit both the P300 and the N2b responses. The temporal activation of these maps form a set of automatically extracted features to predict a single-trial's label. TRODNet is trained and assessed using 10-fold class-stratified cross-validation on a dataset of 54 participants (28 controls). All concussed participants were clinically diagnosed and were symptomatic at the time of testing. Supplementary self-reports were collected to investigate concussive and depressive symptomatology as captured by the post-concussion symptom scale (PCSS) and the Children Depression Inventory 2 (CDI), respectively. Nineteen of the 26 concussed subjects returned for a follow-up test, nine of which reported full symptom recovery (PCSS of 0). Analyses on the longitudinal samples were run in parallel to assess whether symptom resolution was identifiable by the trained model. Model interpretation is a critical factor for integrating machine learning into the clinical setting (Miotto et al., 2017). Thus, trained models were interpreted using the SHapley Additive exPlanations (SHAP) method, a recent introduction to the field with demonstrated success in clinical applications (Lundberg and Lee, 2017; Lundberg et al., 2018; Boshra et al., 2019).

The study was designed to investigate two primary hypotheses. First, the study examined whether single-trial classification can be aggregated for each subject to provide a viable tool of detecting concussion-related neurophysiological effects using minimal feature engineering. Second, the model's judgements on longitudinal datapoints were examined. It was postulated that performance would deteriorate after symptom resolution due to a normalization of the recovered subjects' neurophysiological responses, as opposed to consistent performance in those who retained their symptoms. Model interpretability was prioritized to ensure a transparent representation of learned information and to serve as a confirmatory step for the model's results.

Results

Concussion Identification

As the model was trained (and tested) on single trials, aggregation of the TRODNet output was performed to create a prediction on the subject-level (see Methods for more details). As such, if more than 50% of a subject's trials were classified as concussed, the subject was predicted as belonging to the concussed group. The TRODNet model was able to achieve a single-subject cross-validation accuracy of 85%. Specifically, four control subjects were misidentified as concussed while four concussed subjects were misclassified as controls. This put the model's sensitivity to concussive effects at 84.6% and its specificity at 85.7%. Single-trial cross-validation accuracy was recorded at 74.4%; however, this figure should be assessed with care as discussed below. A detailed list of the model's single-trial accuracies; PCSS and CDI scores; demographics; and number of days since injury for each subject in the concussed group, including the longitudinal results, is reported in table 4.1.

Longitudinal Factors

Assessing the model's single-trial accuracy for the concussed subgroup that participated in the follow-up test yielded a significant drop in accuracy (F(1,17)=8.93, p ¡ 0.01) in the second test compared to the first. A significant main effect of Recovery (symptom resolution [SR] vs. no symptom resolution [NSR]) was also found (F(1,17)=4.84, p = 0.04), indicating a lower accuracy for the NSR group. Lastly, no significant Recovery × Testing Date interaction was observed (F(1,17)=0.17, p = 0.69). Overall, the model assessed 14 of the 19 subjects as concussed at the second testing date. The interaction plot is presented in figure 4.1, showing a clear main effect of Testing Date that is not influenced by Recovery. Additionally, it can be observed that subjects that didn't report symptom recovery had lower single-trial accuracies overall.

Injury Acuteness and Correlation Analyses

The effect of days since injury on perceived results was inconclusive for the first day of assessment (see figure 4.2 and table 4.1). For the second date, self-reported symptoms seemed to increase as days since injury increased for the no symptom resolution (NSR) group. This effect was equally observable in the PCSS and CDI scores. Although the two measures are inherently confounded, this result proposes a layer of subjectivity indicating a worsening of effects as an individual is subjected to symptom persistence. Conversely, no clear effect of days since injury was noted on the EEG/ERP results when accounting for symptom resolution.

Insights from Model Explanations

Upon interpreting the model with SHAP, TRODNet highlighted areas of interest overlapping with previously demonstrated effects in the literature (Baillargeon et al., 2012; Ruiter et al., 2019). The mean absolute SHAP values, indicative of feature importance, were reshaped for display on a 64-channel EEG plot for each condition (see figure 4.3). The two deviants had the most prominent features with important ones forming a bimodal distribution in the posterior regions, morphing into a unimodal shape in the frontal areas. The first and second peaks correspond in time and topography to the P300 and N2b, respectively (Folstein and Van Petten, 2008; Polich, 2007). Features tended to be uniformly important bilaterally, with slightly higher importance for the right side. Responses to the standard condition showed smaller and more dispersed distributions of feature importance, an unexpected finding considering an earlier study on chronic effects of concussion that showed early discernible effects to the standard tones (Boshra et al., 2019).

Discussion

Our results demonstrated the efficacy of an acute/post-acute automated system for concussion identification. In contrast to earlier work in concussion, the utilization of deep learning and convolutional networks enabled an end-to-end solution with minimal feature-engineering (Boshra et al., 2019; Cao et al., 2008; Munia et al., 2017; Prichep et al., 2012). Additionally, the hypothesis that single-trials offer a more granular and effective method of assessing EEG/ERP data was supported.

Results relating symptomatology and neurophysiological effects were negative.

Despite the misalignment between the present study's hypothesis and the data, symptomatology has been previously shown to have little correlation to EEG/ERP effects, especially as neuropsychological measures completely return to baseline in most cases (Gosselin et al., 2010; Ruiter et al., 2019; Baillargeon et al., 2012; Martini et al., 2017). This disagreement extends to other assessment modalities such as quantitative EEG (Nuwer et al., 2005; Munia et al., 2017). Lastly, it is noteworthy that the model's performance drop may be attributable to the time-elapsed since injury, a finding that agrees with a regression study conducted in parallel to the present one (in preparation). These results highlight the need to examine the multiple stages of concussion progression and their effects with care as some may potentially be observable strictly at a particular stage of injury and/or recovery. Moreover, in the longitudinal subset, the model predicted trials of subjects that exhibited symptom resolution as concussed more than the subjects with persisting symptoms. Interestingly, that difference was observed irrespective of Testing Date (1^{st} vs. 2^{nd}; figure 4.1). These results introduce the possibility that a subject's future recovery may be inferred from a participant's EEG/ERP results during their symptomatic stage; however, no strong evidence could be drawn given the constraints of the present dataset.

The interpretability layer on our neural network model confirmed our results' origins as pertaining to neurophysiological signals commonly affected by concussion. Primarily, in the deviant conditions, TRODNet's most important features, as extracted by SHAP, corresponded to the 100-500 ms window, encompassing both the N2b and the P300 (see figure 4.3). Topographical examination of feature importance showed the effects to be predominantly central, with an earlier effect that is marginally lateralized to the right. Examination of the standard condition showed a small parieto-occipital effect in the 100-300 ms range, likely related to the N1-P2 complex. While this finding agrees with previous work on chronic neurophysiological effects of concussion observable in responses to the standard tones in an oddball paradigm, the features show low and dispersed importance measures compared to what was observed in the earlier study (Boshra et al., 2019). This is compatible with a hypothesis that alterations in earlier responses (in the mismatch negativity or the N1/P2 complex) may correspond to irreversible effects of concussion and are strictly prominent in chronic cases (Ruiter et al., 2019; Boshra et al., 2019).

The study exhibits two primary limitations. First, the difference in age between the two groups can be argued to contribute to the model's ability to discern between the two experimental groups. Although there have been several reports of age-related differences in ERPs and resting-state EEG, the evidence supports little to no differences in the range of our two groups (15.04 and 19.3;Stevens et al., 2009; Johnstone et al., 1996; Amenedo and Dıaz, 1998). Thus, we argue that an effect pertaining to the presented age-range is minimal, if not unlikely. Secondly, as correlations between model output and symptomatology was conducted post-hoc, further work is required to confirm the relationships between time-elapsed since injury and ERP effects.

In sum, a strong case for the clinical utility of ERPs in individual assessment of acute/post-acute concussion patients has been presented. The current findings improve upon those from resting-state and quantitative EEG to establish a modality that is able to capture the effects of concussion immediately after insult and years post-injury (Munia et al., 2017; Prichep et al., 2012; Boshra et al., 2019). The intent of this research was not directed at the mechanisms of progression and symptom manifestation, which remain unclear. However, a major step in that direction has been achieved in the translation of a complex, multi-trial EEG signal that was successfully able to provide an accurate identification of concussion incidence. The proposed model, TRODNet, was able to capture distinguishing features without the need for feature engineering, enabling further application to prospective different population ages and pathologies.

Methods

Data Collection and EEG Recordings

Participants

Data were collected from 26 (7 male) adolescents (mean age = 15.04) with a recently sustained and clinically diagnosed concussion (mean days since insult = 20.15). A comparative group of 28 (5 male) participants (mean age = 19.3) acted as healthy controls, reporting no previous head injuries. All participants reported no neurological or auditory problems. The study was reviewed and approved by the Hamilton Integrated Research Ethics Board (HiREB), Hamilton, Ontario, Canada. Prior to study participation, all participants provided informed consent in accordance with the ethical standards of the Declaration of Helsinki.

EEG Stimuli and Experimental Conditions

ERPs were collected to a multi-deviant auditory oddball paradigm (Ruiter et al., 2019; Todd et al., 2008). A 600-tone sequence was presented across two blocks of 300 each. Three deviant tones were presented pseudo-randomly in a continuous stream of standard tones. The standard tone was presented 492 times (82%) at 1000 Hz, 80 dB sound pressure level (SPL), and a duration of 50 ms. Each deviant was presented 36 times (6%) and differed from the standard tone in only one sound characteristic. The frequency deviant was 1200 Hz, the duration deviant was 100 ms, and the intensity deviant was 90 db SPL. Participants were tasked to respond using one button to the standard and another button to all deviants. Due to technical issues, data from the intensity deviant were discarded during analysis.

Procedure

Participants were seated facing a computer screen in a dimly-lit, sound-attenuated room. Auditory stimuli were controlled and sequenced using Presentation software (Neurobehavioural Inc.). Stimuli were presented using noise-cancelling insert earphones (Etymotic ER-1). Participants were instructed to respond to the stimuli as accurately as possible. The protocol was 10 minutes long and was the first of a series of other protocols not pertinent to the present study.

EEG Recording and Preprocessing

Continuous EEG was recorded from 64 Ag/AgCl active electrodes (Biosemi ActiveTwo system) placed according to the extended 10/20 system using an elastic cap. Data were passed through an online bandpass filter of 0.01-100 Hz and referenced to the driven right leg. Data were digitized and saved at 512 Hz. Five external electrodes were recorded with the same settings. Three were placed on the mastoid processes and on the tip of the nose. The last two were placed above and over the outer canthus of the left eye to record eye movements. Stimuli markers were recorded and saved synchronously with the EEG data.

Data were processed offline using a 60 Hz notch and a 0.1-30 Hz (24 dB/oct) bandpass filters before re-referencing to the averaged mastoids. Artifacts were rejected manually using visual inspection followed by independent component analysis (ICA) decomposition. The two components found to correlate with horizontal eye movements and blinks were removed before recomputing sensor data. Trials with correct behavioural responses were segmented to 1200 ms intervals starting 200 ms before stimulus onset. Finally, segments were baseline corrected (-200 to 0 ms) and grouped into their respective experimental conditions before exporting the single trials. All EEG preprocessing was conducted using Brain Vision Analyzer (v2.01; Brain Products GmbH).

Statistical Analyses

Mixed effects analysis of variance (ANOVA) was used to examine the effects of Testing Date (2 levels: First and Second) and Recovery (2 levels: symptom resolution [SR] and no symptom resolution [NSR]) on the accuracies reported by TRODNet.

Machine Learning Procedure

Input Structure

The number of trials t_i^d, such that the superscript d indicates condition, extracted from each subject i was set to 36 to match the design's maximum for each deviant condition. In the standard condition, 36 trials were sampled without replacement for each subject. In cases when rejected data reduced the number of a deviant's trials below 36, bootstrapping was conducted to ensure $t_i^d = 36$. The deep learning classifier concurrently processed a single trial of data from each condition as input observation $O \in \mathbb{R}^{N \times S}$ where N was number of EEG channels (C) $\times$ the number of conditions (D), and S was the number of samples in each segmented trial. Passed samples were restricted to the 50-700 ms window such that $S = 332$. C was 64 channels and D was 3 conditions, yielding $N = 192$. Before dataset split, there were $T_{main} = t_i^d \times \#Subjects = 1944$ unique observations across the two classes, as well as $T_{longitudinal} = 684$ longitudinal observations collected from concussed subjects on their second day of testing. We denote the main dataset tensor as $X \in \mathbb{R}^{T_{main} \times N \times S}$. All EEG data manipulation was conducted using the Python MNE package (Gramfort et al., 2013).

Training and Validation

Stratified 10-fold cross-validation was applied to estimate the generalization accuracy of the trained models. X was split into X_{train} and X_{test} before standardizing both sets based on X_{train}, removing the mean and scaling to unit variance for each feature. Observations from one subject were contained exclusively in either X_{train} or X_{test} to ensure no performance inflation due to subject-specific idiosyncrasies. The learner was batch-trained on X_{train} for 500 epochs where each epoch passed a batch of $B = 160$ randomly-picked observations from X_{train}. The resultant model predicted the labels of each observation in X_{test} to produce the trial $accuracy_t$. A thresholded version of $accuracy_t$ evaluated the $accuracy_s$ of all trials from a single subject. If more than 50% was achieved, the $accuracy_s^i$ for subject i tallied as correct. In instances where X_{test} contained one or more subjects that have undergone a second day of testing, the subjects' second set of trials were evaluated in parallel to assess their follow-up test's accuracy similar to what's described above. This procedure was done to ensure an identical training-set for both testing dates as well as eliminate the possibility of within-subject bias.

Neural Network Architecture and Hyperparameters

Following the notion that a multi-channel EEG signal is the evolution of certain topographies across time, TRODNet utilized convolutional layers to learn commonly occurring topographical maps (Tzovara et al., 2013; Cecotti and Gräser, 2011; Schirrmeister et al., 2017; Lawhern et al., 2018). The present architecture, based on EEGNet and an EEG ConvNet, expanded to account for multiple conditions in the same input observation (Lawhern et al., 2018; Schirrmeister et al., 2017). The network had five layers in total (in addition to input).

- L_{input}: This describes the input layer. The input tensor is of size $B \times N \times S$ and is reshaped to $B \times N \times S \times 1$ before passing to the next layer.

- L_1: The input tensor was split across three separate convolutional filters such that each was tasked with learning $M = 5$ maps that are specific to the condition. Kernel size was set to (64, 1). The output from each of the three sub-layers was of size $B \times 1 \times S \times M$. The outputs were concatenated across the last dimension before passing to the next layer.

- L_2: A maxpooling layer was applied with both a pool size and stride of (1, 10) and (1, 5), respectively.

- L_3 *and* L_4: the next two layers were dense feed-forward layers of sizes 50 and 100, respectively.

- L_{output}: The output layer acted as the label predictor with softmax activation to separate classes *concussed* and *control*.

All layers but L_{output} had a rectified linear activation unit (ReLU). L_2 regularization was applied on all weights with $\lambda = 0.25$. The Adam optimizer was used during training with $\alpha = 5\text{e}{-4}$. Training for a cross-validation iteration was stopped after 500 complete epochs. These hyperparameters were set to optimize a separate dataset collected using the same EEG/ERP protocol and were not modified throughout training (Ruiter et al., 2019; Boshra et al., 2019).

Model Interpretation

The Deep Learning Important FeaTures (DeepLIFT) implementation using Shapley values was applied post-hoc on a model trained on all data to explain a model's decision on single-subject averages (Shrikumar et al., 2017; Lundberg and Lee, 2017). An overall estimate of all features' influence on classification was calculated as the mean of the absolute SHAP values for all single-subject averages. The values were overlaid across the head to represent a 64-channel plot as commonly used in EEG/ERP studies. For visual clarity, each experimental condition was plotted independently.

Acknowledgements

The authors gratefully acknowledge the participants for their time, Ms. Chia-Yu Lin for coordinating and organizing this study, as well as the entire Back-2-Play team for their support. This work was supported by the Canadian Institutes of Health Research (JFC and CD), Canada Foundation for Innovation (JFC), Senator William

McMaster Chair in Cognitive Neuroscience of Language (JFC), the MacDATA fellowship award (RB), the Vector Institute postgraduate affiliate award (RB), and the Ontario Ministry of Research and Innovation (RB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions statement

R.B. conceived the machine learning experiment and wrote the initial draft; J.F.C., R.B., and K.I.R. conducted the experiment(s); and R.B. analysed the results. All authors reviewed and edited the manuscript.

Data Availability

The input set was imported and formatted using Python MNE package version 0.16.1 running on Python 3.5.2 (Gramfort et al., 2013). Cross-validation and scaling were applied using scikit-learn 0.19.1 (Pedregosa et al., 2011). Deep learning used Tensorflow (v1.8.0;Abadi et al., 2016). All code is made available at https://github.com/boshra/TRODNet. Statistical analysis was conducted using R statistical software (v3.5.3) and the ez package (v4.4-0). Result storage, correlational plots, and feature importance visualizations were conducted using the pandas (v0.24.1), seaborn (v0.9.0), and Python MNE packages, respectively. The single-trial data used to train the models of this study are available upon request from one of the corresponding authors (J.F.C. or R.B.).

Competing Interests

The authors declare no competing interests.

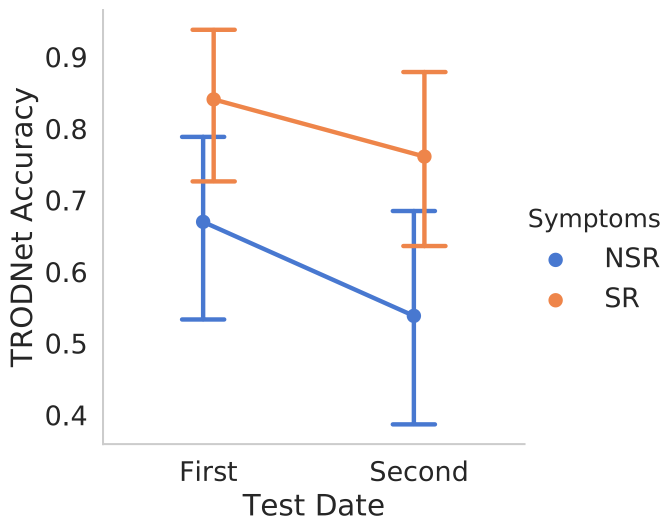

Figure 4.1: The interaction effect of Recovery and Testing Date on the TRODNet results as seen on the longitudinal subgroup. While there were main effects of both factors, no reliable interaction was found. Points represent mean prediction from TRODNet's result, where 0 (1) is a classification of control (concussed). Vertical extended lines indicate the 95% confidence intervals.

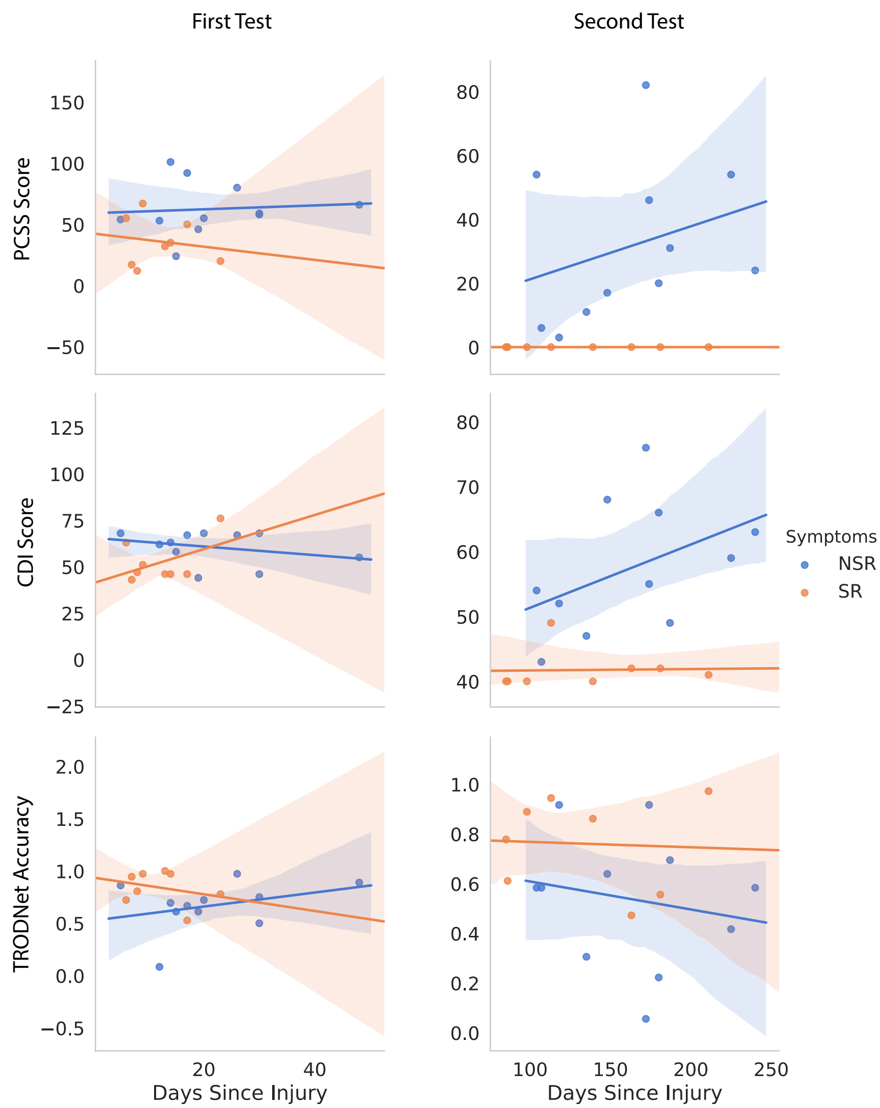

Figure 4.2: Interactions between days since injury and symptomatology (first row), depressive symptoms (second row), and TRODNet single-trial results in the longitudinal sample of our presented dataset (third row). The symptom resolution (SR) subgroup conveyed no identifiable patterns both in the first (left column) and second (right column) tests. The subgroup that did not have symptoms resolve (NSR) showed an increase in symptomatology and depressive signs as days since injury increased for the second test. Shaded regions signify the 95% confidence intervals.

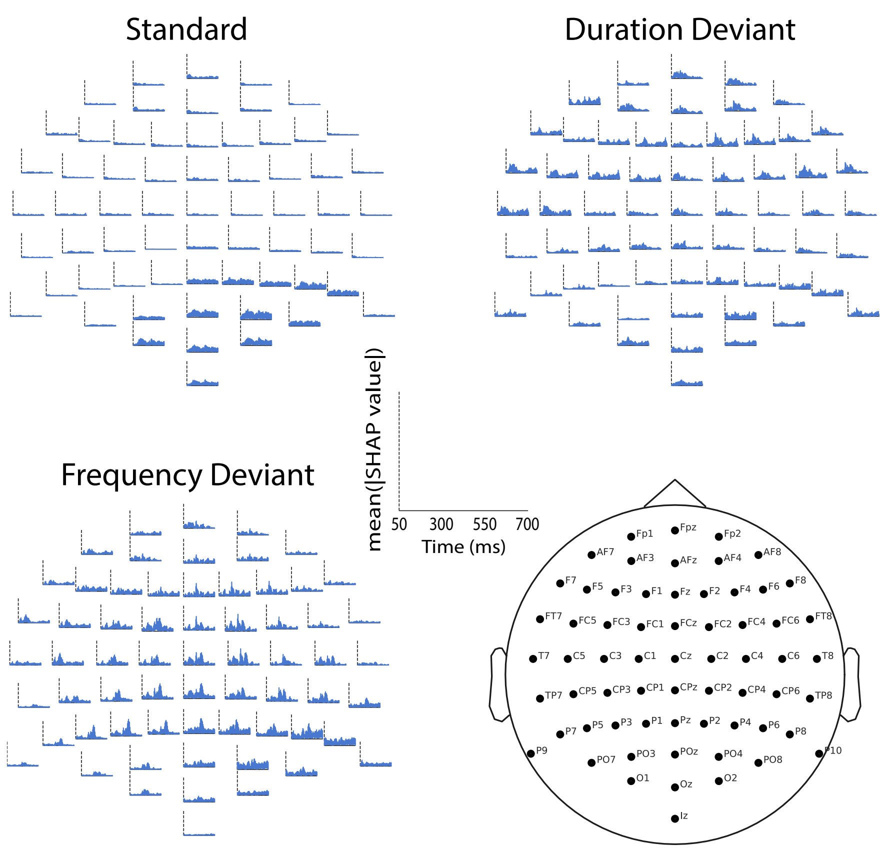

Figure 4.3: The mean of the absolute SHAP values for single-subject averages overlaid on the head for each condition and electrode. The abscissa denote time where 0 is the stimulus onset. The ordinate represents the mean absolute SHAP value at the indicated electrode, time, and condition. The figure shows a robust identification of ERPs of interest, particularly in the frequency (FDev) and Duration (DDev) deviants. An interesting effect can be observed to the standard condition where the parieto-occipital region has a widespread effect predominantly in the right hemisphere.

ID	Sex	Age	#Prev Concs	#Days since injury	PCSS	CDI	Accuracy
1	F	17	6	36	109	71	94.4
2	F	16	0	20 - 225	55-54	68-59	72.2 - 41.7
3	F	13	1	5 - 148	54-17	68-68	86.1 - 63.9
4*	M	14	2	23 - 86	20-0	76-40	77.8 - 61.1
5	M	13	2	30	64	49	86.1
6	M	16	2	7	33	57	5.6
7*	F	17	2	14 - 139	35-0	46-40	97.2 - 86.1
8	F	16	6	8	94	52	66.7
9	F	15	1	17 - 107	92-6	67-43	66.7 - 58.3
10*	F	15	1	9 - 211	67-0	51-41	97.2 - 97.2
11*	F	17	1	17 - 163	50-0	46-42	52.8 - 47.2
12	M	13	5	14 - 104	101-54	63-54	69.4 - 58.3
13	F	15	1	13	41	43	72.2
14	F	17	4	15 - 240	24-24	58-63	61.1 - 58.3
15	F	17	3	30 - 135	58-11	46-47	75.0 - 30.6
16*	F	13	0	7 - 98	17-0	43-40	94.4 - 88.9
17*	F	13	1	8 - 85	12-0	47-40	80.6 - 77.8
18	M	15	2	19 - 187	46-31	44-49	61.1 - 69.4
19	F	17	1	12 - 180	53-20	62-66	8.3 - 22.2
20	M	14	1	58	55	49	36.1
21	F	14	0	30 - 172	59-82	68-76	50.0 - 5.6
22	F	17	2	39	60	71	66.7
23	F	16	2	26 - 174	80-46	67-55	97.2 - 91.7
24*	F	14	1	6 - 181	55-0	63-42	72.2 - 55.6
25*	M	13	1	13 - 113	32-0	46-49	100 - 94.4
26	F	14	1	48 - 118	66-3	55-52	88.9 - 91.7

Table 4.1: Table detailing the symptomatology and depression scores for all concussed participants. Bolded subject IDs (19) represent the ones who returned for a follow-up EEG test. Where applicable, second testing (t2) values are presented after the first (t1) values as in: t1, t2. Asterisks (*) on subject IDs denote the concussed subgroup that reported full symptom recovery (PCSS of 0) by their second assessment.

Bibliography

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., and Others (2016). Tensorflow: A system for large-scale machine learning. In *12th \${\$USENIX\$}\$ Symposium on Operating Systems Design and Implementation (\${\$OSDI\$}\$ 16)*, pages 265–283.

Amenedo, E. and Dıaz, F. (1998). Automatic and effortful processes in auditory memory reflected by event-related potentials. age-related findings. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 108(4):361–369.

Baillargeon, A., Lassonde, M., Leclerc, S., and Ellemberg, D. (2012). Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Injury*, 26(3):211–220.

Boshra, R., Dhindsa, K., Boursalie, O., Ruiter, K. I., Sonnadara, R., Samavi, R., Doyle, T. E., Reilly, J. P., and Connolly, J. F. (2019). From group-level statistics to single-subject prediction: Machine learning detection of concussion in retired athletes. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Moore, R. D., and Hillman, C. H. (2011). A history of sport-related concussion on event-related brain potential correlates of cognition. *International Journal of Psychophysiology*, 82(1):16–23.

Broglio, S. P., Pontifex, M. B., O'Connor, P., and Hillman, C. H. (2009). The Persistent Effects of Concussion on Neuroelectric Indices of Attention. *Journal of Neurotrauma*, 26(9):1463–1470.

Brush, C. J., Ehmann, P. J., Olson, R. L., Bixby, W. R., and Alderman, B. L. (2018). Do sport-related concussions result in long-term cognitive impairment? A review of event-related potential research. *International Journal of Psychophysiology*, 132(March 2017):124–134.

Cao, C., Tutwiler, R. L., and Slobounov, S. (2008). Automatic classification of athletes with residual functional deficits following concussion by means of EEG signal using support vector machine. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(4):327–335.

Cecotti, H. and Gräser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):433–445.

Chen, J. H. and Asch, S. M. (2017). Machine learning and prediction in medicine—beyond the peak of inflated expectations. *The New England journal of medicine*, 376(26):2507.

Connolly, J. F., Reilly, J. P., Fox-Robichaud, A., Britz, P., Blain-Moraes, S., Sonnadara, R., Hamielec, C., Herrera-Díaz, A., and Boshra, R. (2019). Development of a point of care system for automated coma prognosis -a prospective cohort study protocol. *BMJ Open*.

De Beaumont, L., Brisson, B., Lassonde, M., and Jolicoeur, P. (2007). Long-term electrophysiological changes in athletes with a history of multiple concussions. *Brain Injury*, 21(6):631–644.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., and Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, 120(11):1883–1908.

Fickling, S. D., Smith, A. M., Pawlowski, G., Ghosh Hajra, S., Liu, C. C., Farrell, K., Jorgensen, J., Song, X., Stuart, M. J., and D'Arcy, R. C. (2019). Brain vital signs detect concussion-related neurophysiological impairments in ice hockey. *Brain*, 142(2):255–262.

Folstein, J. R. and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1):152–170.

Gaetz, M., Goodman, D., and Weinberg, H. (2000). Electrophysiological evidence for the cumulative effects of concussion. *Brain Injury*, 14(12):1077–1088.

Gosselin, N., Bottari, C., Chen, J.-K., Huntgeburth, S. C., De Beaumont, L., Petrides, M., Cheung, B., and Ptito, A. (2012). Evaluating the cognitive consequences of mild traumatic brain injury and concussion by using electrophysiology. *Neurosurgical Focus*, 33(6):E7.

Gosselin, N., Saluja, R. S., Chen, J. K., Bottari, C., Johnston, K., and Ptito, A. (2010). Brain functions after sports-related concussion: Insights from event-related potentials and functional MRI. *Physician and Sportsmedicine*, 38(3):27–37.

Gosselin, N., Thériault, M., Leclerc, S., Montplaisir, J., and Lassonde, M. (2006). Neurophysiological Anomalies in Symptomatic and Asymptomatic Concussed Athletes. *Neurosurgery*, 58(6):1151–1161.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*.

Johnstone, S. J., Barry, R. J., Anderson, J. W., and Coyle, S. F. (1996). Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *International Journal of Psychophysiology*, 24(3):223–238.

Langlois, J. A., Rutland-Brown, W., and Wald, M. M. (2006). The Epidemiology and Impact of Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*, 21(5):375–378.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 15(5):1–30.

Lundberg, S. and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems*, 16(3):426–430.

Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., Kim, J., and Lee, S.-I. (2018). Explainable machine learning predictions to help anesthesiologists prevent hypoxemia during surgery. *Nature Biomedical Engineering*.

Martini, D. N., Eckner, J. T., Meehan, S. K., and Broglio, S. P. (2017). Long-term Effects of Adolescent Sport Concussion Across the Age Spectrum. *American Journal of Sports Medicine*, 45(6):1420–1428.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R. C., Cassidy, D., Echemendia, R. J., Castellani, R. J., Davis, G. A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C. C., Guskiewicz, K. M., Herring, S., Iverson, G. L., Johnston, K. M., Kissick, J., Kutcher, J., Leddy, J. J., Maddocks, D., Makdissi, M., Manley, G. T., McCrea, M., Meehan, W. P., Nagahiro, S., Patricios, J., Putukian, M., Schneider, K. J., Sills, A., Tator, C. H., Turner, M., and Vos, P. E. (2017). Consensus statement on concussion in sport—the 5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, (October 2016):bjsports–2017–097699.

Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, (February):1–11.

Moore, R. D., Broglio, S. P., and Hillman, C. H. (2014). Sport-related concussion and sensory function in young adults. *Journal of Athletic Training*, 49(1):36–41.

Munia, T. T., Haider, A., Schneider, C., Romanick, M., and Fazel-Rezai, R. (2017). A Novel EEG Based Spectral Analysis of Persistent Brain Function Alteration in Athletes with Concussion History. *Scientific Reports*, 7(1):1–13.

Nuwer, M. R., Hovda, D. A., Schrader, L. M., and Vespa, P. M. (2005). Routine and quantitative EEG in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025.

Obermeyer, Z. and Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine*, 375(13):1216.

Opałka, S., Stasiak, B., Szajerman, D., and Wojciechowski, A. (2018). Multi-channel convolutional neural networks architecture feeding for effective EEG mental tasks classification. *Sensors (Switzerland)*, 18(10):1–21.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of machine learning research*, 12:2825–2830.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b.

Prichep, L. S., Jacquin, A., Filipenko, J., Dastidar, S. G., Zabele, S., Vodencarevic, A., and Rothman, N. S. (2012). Classification of Traumatic Brain Injury Severity Using Informed Data Reduction in a Series of Binary Classifier Algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(6):806–822.

Rajkomar, A., Dean, J., and Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14):1347–1358.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38(11):5391–5420.

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning important features through propagating activation differences. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3145–3153. JMLR. org.

Stevens, M. C., Pearlson, G. D., and Calhoun, V. D. (2009). Changes in the interaction of resting-state neural networks from adolescence to adulthood. *Human brain mapping*, 30(8):2356–2366.

Sturm, I., Lapuschkin, S., Samek, W., and Müller, K. R. (2016). Interpretable deep neural networks for single-trial EEG classification. *Journal of Neuroscience Methods*, 274:141–145.

Taylor, C. A., Bell, J. M., Breiding, M. J., and Xu, L. (2017). Traumatic Brain Injury–Related Emergency Department Visits, Hospitalizations, and Deaths — United States, 2007 and 2013. *MMWR. Surveillance Summaries*, 66(9):1–16.

Todd, J., Michie, P. T., Schall, U., Karayanidis, F., Yabe, H., and Näätänen, R. (2008). Deviant Matters: Duration, Frequency, and Intensity Deviants Reveal Different Patterns of Mismatch Negativity Reduction in Early and Late Schizophrenia. *Biological Psychiatry*, 63(1):58–64.

Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., and De Lucia, M. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, 136(1):81–89.

CHAPTER 5 On the Time-Course of Functional Brain Connectivity: Theory of the Dynamic Progression of Concussion Effects

Preface

Throughout the analyses conducted in both the studies presented in chapters 4 and 5, as well as parallel studies on the same datasets using traditional analysis techniques, it became incrementally clearer that concussion's influence on the human brain was not as lucid as initially hypothesized. Despite the consistency of several key effects – namely, those of the P300 – the variance of other co-occurring ERP effects was not immediately explainable. The study presented in this chapter was the conducted in an effort to clarify whether the inconsistencies of the previous results, analogous to the literature, were due to idiosyncrasies manifesting in our particular datasets and whether a more defined view of mTBI could be framed. To maximally reduce any statistical biases inherent of reanalyzing the data, different measures of event-related functional brain connectivity were utilized. In addition to addressing the study's primary question and finding patterns analogous to what is often seen in neuroimaging studies, the present chapter put forth a complete framework of mTBI and its progression through different injury stages. The proposed model is believed to be a valuable tool for hypothesis creation to maximize the impact of studies in bettering the current understanding of the injury's mechanisms.

Abstract

The current literature presents a discordant view of mild traumatic brain injury and its effects on the human brain. This dissonance has often been attributed to heterogeneities in study populations, etiology, acuteness, and/or testing modalities. To investigate the progression of mTBI in the human brain, the present study used data from 93 subjects (48 healthy controls) representing both acute and chronic stages of mTBI. The effects of concussion across different stages of injury were measured using two metrics of functional connectivity (FC) in segments of electroencephalography (EEG) time-locked to an active oddball task. Results show an increase in FC in the acute stage after mTBI, contrasted with significantly reduced FC in chronic stages of injury. This finding indicates a non-linear time-dependent effect of injury. In order to understand this pattern of changing FC in relation to prior evidence, we propose a new model of the time-course of the effects of mTBI on the brain that brings together research from multiple neuroimaging modalities and unifies the various lines of evidence that at first appear to be in conflict.

5.1 Introduction

Concussions and mild traumatic brain injury (mTBI) are difficult to detect via common brain imaging methods due to their predominantly functional effects on the brain. Work spanning a multitude of testing modalities supports a connection between concussions and multiple forms of cognitive dysfunction (McCrory et al., 2017). With increased awareness of the condition, concussion diagnosis is at an unprecedented high; however, clinical practices for identification, treatment, and management remain inadequate (Broglio et al., 2017, 2007). Refinement in assessment tools has targeted the disentanglement of concussion's constellation of symptoms with several methodologies showing promise such as: eye-movements (Johnson et al., 2012; Heitger et al., 2009), balance assessments (Broglio et al., 2017), functional brain imaging (McAllister et al., 2001), and electrophysiology (Broglio et al., 2009; De Beaumont et al., 2009; Ruiter et al., 2019). Of the last category, event-related potentials (ERP) as recorded with electroencephalography (EEG) were shown to be altered following concussion (De Beaumont et al., 2007; Broglio et al., 2011; Gosselin et al., 2006; Gaetz et al., 2000). The extent of these alterations was demonstrated to correlate with number of hits to the head, severity of injury, and time elapsed since injury (De Beaumont et al., 2007; Broglio et al., 2011). Most prominent in concussion research, the N2b and the P300 have been shown to be particularly affected after insult, either in response latency, amplitude, or both (Ruiter et al., 2019; De Beaumont et al., 2007; Gosselin et al., 2006; Broglio et al., 2009; De Beaumont et al., 2013). The association between these ERPs and specific cognitive processes has provided a valuable tool to pinpoint the functional and cognitive effects of concussion that have been reported to linger after symptom resolution (Gosselin et al., 2006; De Beaumont et al., 2007; Baillargeon et al., 2012). However, it remains unclear how the effects of concussion progress over the years following injury.

Most findings on the neurophysiological effects of concussion to date describe an attenuation or delaying of brain responses (Ruiter et al., 2019; De Beaumont et al., 2007; Gosselin et al., 2006; Broglio et al., 2009; De Beaumont et al., 2013). Contrary to these findings, several studies have reported hyperactivation post-injury during mentally-taxing tasks (McAllister et al., 2001), making inadequate a simple explanation that concussion leads to a static reduction in brain activity. Increase in activation has since been posited to result from allocation of extra resources to compensate for injury (McAllister et al., 2001; Iraji et al., 2016). This compensatory mechanism was argued to reflect the discrepancy often found between behavioural assessments and neurological measurements from electrophysiology (Ruiter et al., 2019; De Beaumont et al., 2012, 2009) and functional hemodynamics (McAllister et al., 2001; Hocke et al., 2018a). Subsequently, the notion of a cognitive reserve (see Stern, 2009; Kesler et al., 2003) has been hypothesized to explain the overall neurocognitive decline of previously concussed individuals through aging (De Beaumont et al., 2009, 2012). According to the cognitive reserve theory, a previously-injured brain loses the ability to sustain its compensatory mechanisms with age, resulting in an abnormal aging process with a resurgence of symptoms and other cognitive deficits (De Beaumont et al., 2009, 2012). For the purposes of the present study, we define three broad stages of concussion progression: a) *acute*, denoting the time directly after injury and extending to 3 weeks after (21 days) wherein the primary symptoms of mTBI are most apparent, b) *post-acute*, pertaining to the time after the acute stage to late-stage symptom resurgence that takes place with aging, during which the behaviourally-observable symptoms of mTBI appear to be most resolved, and c) *chronic*, referring to the state of injury decades after insult that is seen in late-adulthood.

While there is consensus that brain function is altered following concussion, there has been little work clarifying the progression of post-concussive effects throughout aging, the relationship between those effects to observable symptomatology, or to consolidate the results from different imaging methods. The notion that neurophysiological sequelae of concussion progress in a non-linear fashion is not new. A study by Zhu et al. (2015) reported increased functional connectivity (defined below) directly after a concussive episode followed by a return to normal levels 30 days after injury. Similar findings were reported in the literature where tests in the acute phase revealed a hyperconnected brain (Sours et al., 2015; Bernier et al., 2017; Shumskaya et al., 2012; Bharath et al., 2015; Nakamura et al., 2009; Messé et al., 2013; Iraji et al., 2016), whereas tests conducted after the acute stage of injury has elapsed showed reduced FC (Hocke et al., 2018a; Robinson et al., 2015), and one study reported no effect of mTBI on FC (Churchill et al., 2017). Of note, these studies utilized a wide array of brain imaging methodologies and experimental protocols, including recording brain activity in resting-state designs or under active cognitive load induced by different tasks.

Functional connectivity (FC), as opposed to structural connectivity, broadly describes statistically correlated activity and synchronization between brain regions (Bastos and Schoffelen, 2016). This correlated activity can be due to information transmission between communicating brain regions, or simply due to different brain regions contributing to a common task. In EEG analysis, FC can be measured using a variety of methods, each contributing different information about how brain activity is synchronized across regions (Vinck et al., 2011; Blain-Moraes et al., 2016; Bastos and Schoffelen, 2016). These can be computed across EEG sensors or estimated brain sources, and can describe temporally correlated changes in the power spectrum or phase-coupling. Two complementary methods are used in this study: magnitude-squared coherence, and the weighted phase-lag index (WPLI). Both of these methods have been widely used in the literature to describe non-directed connections between two signals (i.e., they do not specifically indicate which signal influences the other).

Magnitude-squared coherence, or simply coherence, is a measure of FC that describes the degree of linear similarity between two signals (in our case, two channels) in terms of their power spectrum (Nolte et al., 2004; Murias et al., 2007; Kumar et al., 2009). Specifically, it is the normalized magnitude-squared cross-spectral density of two signals in a given frequency band f, and thus primarily describes amplitude-related synchronization. Coherence is calculated as

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)} \tag{5.1}$$

where P_{xy} is the cross-spectral density of two signals x and y, and P_{xx} and P_{yy} are their respective power spectral densities. However, unlike WPLI, coherence is susceptible to influence from volume-conduction and noise (Vinck et al., 2011; Bastos and Schoffelen, 2016).

WPLI is a measure of FC that provides information primarily about phase-synchronization, and is designed to be robust in the presence of volume-conduction and noise (Vinck et al., 2011) – all features which are crucial to channel-space analysis of EEG. WPLI is calculated using the imaginary component of the cross-spectrum, which we denote as $ImC_{xy}(f)$, and which gives the phase-synchronization between two signals. It is then weighted by its sign, $\text{sign}(ImC_{xy}(f))$, which indicates which of the two signals leads or lags in phase-space. Putting these together, WPLI is calculated as follows:

$$WPLI = \frac{|E\{ImC_{xy}(f)\}|}{E\{|ImC_{xy}(f)|\}} \tag{5.2}$$

FC as measured using EEG has been investigated for application in mTBI with several reports of deficits observed in resting state EEG after concussion (Cao and Slobounov, 2010); although, the effects of post-acute mTBI on resting-state EEG have been questioned in the literature (Nuwer et al., 2005). Related work on more severe TBI showed a reduced FC effect that was unique to a working memory task, as opposed to resting state (Kumar et al., 2009). This work raised the issue that FC effects in mTBI might require an active cognitive load for the separation to be observed, a hypothesis that is compatible with fMRI research on mTBI (McAllister et al., 2001; Johnson et al., 2012, 2015). To date, no prior studies have been conducted to investigate the progression of neurphysiological FC effects of mTBI using EEG during active cognitive tasks. Moreover, there is limited work attempting to connect the effects found in EEG/ERP measures to the neuroimaging literature, which places a greater emphasis on FC. Here we show that functional brain connectivity observed in EEG at different time-periods following concussion provides valuable information on how concussion impacts brain activity over time.

In the present article, we investigate the effects of concussion on brain connectivity, as well as the development of these effects from youth into late adulthood. The study employed experiments designed to elicit canonical ERPs studied in concussion research (the N2b and P300) in order to examine a cognitively taxed brain using event-related FC. The study examined two hypotheses. First, in accordance with prior literature (Kumar et al., 2009), we expected a comparison between concussed subjects and age-matched controls to reveal a difference in FC. Second, we postulated that while some changes in connectivity would normalize after the acute stage, there would remain differences in FC between concussed subjects in the chronic stage and their age-matched controls. In addition to presenting our experimental findings, we propose a new theory of the effects of concussion on the brain that integrates cognitive reserve theory to explain how our data can be synthesized with the evidence found within the existing neuroscientific literature on mTBI in order to understand how concussion affects the brain over the lifespan.

5.2 Methods

5.2.1 Participants

To investigate the progression of FC changes from acute concussion to late after injury, data were collected from a total of 93 participants (48 controls) split across two age-groups in a cross-sectional manner. Data were collected as part of independent studies conducted using the same set of EEG/ERP paradigms as defined below.

Acute (AC) Dataset The first group was comprised of 26 participants (19 female) with an average age of 15.4 years that had sustained a head injury. All concussed participants were clinically diagnosed with a concussion and were subsequently tested an average of 20.15 days after injury. Data from 28 healthy participants (average 19.2

years old; 22 female) acted as a control group for the AC population.

Chronic (CH) Dataset A total of nineteen male retired football athletes were recruited with an average age of 57.6 years old. The retired athletes played an average of 7.84 years in the Canadian Football League (CFL), self-reported an average of 4.05 previous concussions, and indicated an average of 28.11 years between day of testing and the last reported head injury. A group of twenty healthy individuals (all male; mean age = 53.7) acted as controls for this age group.

All control participants reported no history of neurological disorders or head injury. All subjects reported normal hearing and provided written consent prior to participating in the study. The study was approved by the Hamilton Integrated Research Ethics Board (HiREB; Ontario, Canada) and was in accordance with the ethical standards of the Declaration of Helsinki.

5.2.2 EEG Stimuli

Each subject completed two separate EEG/ERP protocols designed to examine brain responses in both active and passive task conditions separated by a distractor task (Ruiter et al., 2019). Only data from the active protocol, presented first, was used in the present study. The protocol was a multi-deviant auditory-oddball task adapted from Todd et al. (2008). The standard tone (Std) was a 1000 Hz pure-tone presented at 80 dB sound pressure level (SPL). Three different tones, each similar to the Std in all but one sound characteristic, were presented as deviants to the standard tone sequence: frequency deviant (FDev; 1200 Hz), duration deviant (DDev; 100 ms), and intensity deviant (90 dB SPL). Standard tones were presented 492 times (82%) while each of the deviants was presented 36 (6%) times for a total of 600 stimulus presentations. Participants were tasked to actively attend to the tones and respond by pressing one button to standards and another button for all deviants. All segments with incorrect responses were discarded from further analysis. Technical issues rendered intensity deviants unusable in the AC recordings; thus, to facilitate comparisons between the two age-groups, responses to intensity deviants were discarded for all further analysis.

5.2.3 EEG Procedure

Participants were seated in a dimly-lit and sound-attenuated room. Noise-cancelling earphones (Etymotic ER-2) were used to deliver all binaural auditory stimuli. Participants were instructed to fixate on a cross placed in the centre of a computer monitor as they responded using two buttons to auditory stimulation as described above. Buttons were counterbalanced between participants. Presentation of all stimuli and respective EEG markers was done using Presentation software (NeuroBehavioral Systems; NBS). The protocol used for this study lasted a total of 10 mins in a addition to a break halfway through presentation. Participants were instructed to switch buttons for the second half of the protocol.

5.2.4 EEG Recording

Continuous EEG was recorded from 64 Ag/AgCl active electrodes using the Biosemi ActiveTwo system. Electrodes were placed according to the extended 10-20 system using an elastic cap and referenced online to the driven right leg circuit. Data from five external electrodes were recorded: two electrodes were placed separately to record eye movements placed above and over the outer canthus of the left eye, two electrodes on the mastoid processes, and one on the tip of the nose. Data from all electrodes were digitized at 512 Hz and passed through a 0.01-100 Hz bandpass filter.

5.2.5 EEG Data Preprocessing

Offline, data were passed through a 0.1-30 Hz (24 dB/oct) filter in addition to a 60 Hz notch filter. Using visual inspection, all segments containing non-ocular artifacts were marked for deletion. Independent component analysis was conducted on the remaining continuous data and components correlated with either external electrode placed around the left eye were removed. Following data cleaning, the EEG signals were re-referenced to the averaged mastoids and segmented for all experimental conditions from 200 ms before stimulus onset to 1000 ms after. All segments were baseline-corrected (-200 – 0 ms) before separating segments by condition and exporting to binary files for connectivity analysis. All the prior steps were conducted using Brain Vision Analyzer (v2.1; Brain Products GmbH).

5.2.6 Connectivity Analysis

We assessed brain connectivity as measured in sensor-space by computing pairwise FC for regions of interest that were defined similarly to Kumar et al. (2009). Fourteen electrodes were clustered based on their topographical location: laterally on the right (R) and left (L) hemispheres; and caudally at the frontal (F), temporal (T), and parietal (P) regions (see Fig. 5.1). The six clusters of 2-3 electrodes each were L-F: (F3, F5, and F1); R-F: (F4, F2, and F6); L-T: (T7 and CP5); R-T: (T8 and CP6); L-P: (P3 and P1); and R-P: (P4 and P2). Using the six clusters, four categories of connectivity were defined:

- Interhemispheric: described connections between the left and right hemispheres and was split to Frontal (between R-F and L-F), Temporal (between R-T and L-T), and Parietal (between R-P and L-P).
- Intrahemispheric, long-range: defined connections that spanned from the frontal region to the parietal region within the same hemisphere. In the right hemisphere they were defined as all combinations between R-F and R-P, whereas in the left hemisphere they were connections between L-F and L-P.
- Intrahemispheric, mid-range: described fronto-temporal and temporo-parietal connections in both hemispheres. In the frontal region: left fronto-temporal (between L-T and L-F) and right fronto-temporal (between R-T and R-F). In the parietal region: left temporo-parietal (between L-T and L-P) and right temporo-parietal (between R-T and R-P).
- Intrahemispheric, within-region: contained all pairwise comparisons within each electrode cluster described above.

For each category of connectivity, WPLI and coherence were calculated between each electrode-pair for each experimental condition using Python MNE (Gramfort et al., 2013). Connectivity was assessed across five canonical EEG bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-14 Hz), and beta (14-23 Hz). The spectral densities were estimated using the Morlet wavelet with central frequencies $f_i \in 1, 2, \ldots 23$ and corresponding cycles $c_i = min(0.75 \times f_i, 7)$. FC values were averaged for the duration spanning between stimulus onset and 800 ms after.

5.2.7 Statistical Analysis

Statistical analyses were adapted from Kumar et al. (2009) and modified to reduce the number of multiple comparisons. Mixed-effects analysis of variance (ANOVA; $\alpha = 0.01$) was used to examine whether concussion affected brain connectivity as measured by the WPLI and coherence independently. Separate univariate ANOVAs were run for each type of connectivity and spectral band such that the dependent variable was the connectivity metric (WPLI or coherence) to investigate the effect of Group (2 levels: control and concussed), Condition (3 levels: Std, FDev, and DDev), Site (as described above for each connectivity category), and Age (2 levels: young and old). In cases of Sphericity violations (assessed using Mauchly's Test), Greenhouse-Geisser estimates of epsilon were used to correct for the degrees of freedom. In instances of significant interactions, Bonferroni-corrected post-hoc analyses were conducted to investigate the effect.

5.3 Results

5.3.1 Coherence

Coherence was heavily influenced by Age differences. This manifested as a significant main effect of Age as well as a reliable effect of Age $\times$ Condition interaction in all connectivity types across all bands when observing coherence (see table 5.1). These effects are omitted for brevity from the detailed results below.

band ($p < 0.01$).

5.3.2 Intrahemispheric, long-range

In coherence, a significant Age $\times$ Site interaction was observed in all bands ($p < 0.01$; table 1). Additionally, a significant 3-way interaction of Group $\times$ Age $\times$ Site manifested in delta and theta ($p < 0.01$). WPLI analysis indicated a significant Group $\times$ Age interaction in both delta and theta bands ($p < 0.01$; figure 5.2). Additionally in theta, an Age $\times$ Site interaction was significant ($p < 0.01$). Post-hoc analysis in both bands showed that in the CH age-group, FC was lower in the left hemisphere than in the right hemisphere.

5.3.3 Intrahemispheric, mid-range

Results indicated a significant Group $\times$ Age interaction in coherence for the delta and theta bands ($p < 0.01$; see table 5.1). In WPLI, a main effect of Age was found in delta ($p < 0.01$) and theta ($p < 0.01$). A significant Group $\times$ Age interaction was found in delta and theta ($p < 0.01$; figure 5.2, B). An additional Age $\times$ Site interaction was observed in delta and theta ($p < 0.01$). Post-hoc analysis in both bands showed that FC in the AC group measured between at FT-L and FT-R was higher than TP-L and TP-L. That effect was reversed in the older CH group. The two 3-way interactions (Group $\times$ Age $\times$ Condition and Group $\times$ Age $\times$ Site) were found significant in the delta band ($p < 0.01$). Investigation of the first interaction showed that the Group $\times$ Age interaction was not observable in Std. Post-hoc analysis of the second interaction showed that FC observed from FT-L showed a significant effect of Group as well as a reliable Group $\times$ Age interaction in TP-L and TP-R.

5.3.4 Within-region

Analysis of coherence yielded significant Group $\times$ Age and Age $\times$ Site interactions in all bands ($p < 0.01$). An additional Group $\times$ Condition was observed in delta ($p < 0.05$). For WPLI, a main effect of Age was found in all but the beta band ($p < 0.01$), whereas a Group $\times$ Age interaction was observed in all bands but alpha ($p < 0.01$; figure 5.2). The Age $\times$ Site interaction was significant in the theta band ($p < 0.01$). Post-hoc analysis showed FC to be larger in the AC group compared to the CH group in all but the frontal sites (L-F and R-F).

5.4 Discussion

Our results introduce new evidence illustrating the evolution of the effects of mTBI on the human brain over time. We show that while an increase in connectivity from baseline may follow from concussion soon after injury, a reversal towards reduced connectivity may in part characterize the long-term effects (see Figure 5.2). In light of these findings, we argue for a critical component of our hypotheses: the effects of mTBI on the brain are dynamic and must be contextualized by how long the brain has been adapting to the effects of concussion. Concordantly, the interaction between history of concussion and age group had a significant effect on FC, while the fact that concussion had taken place in the past on its own did not.

Whether FC was measured with coherence, to capture power synchronization, or WPLI, to capture phase synchronization, the chronically-concussed group in the present study showed a widespread reduction in FC. While this result appears to conflict with what was found by Kumar et al. (2009), the findings may in fact be complimentary. Notably, Kumar et al. found reduced FC in concussed individuals an average of 2-3 months since last concussion. When comparing more recently concussed individuals (an average of just 21 days post-injury) with controls, we instead found increased levels of FC. One possible explanation is that there may be a sharp increase in FC shortly after injury, followed by a decrease in FC below pre-injury levels, a finding supported by longitudinal imaging studies (Zhu et al., 2015; Iraji et al., 2016). However, another explanation is that Kumar et al.'s study included subjects who suffered from more severe brain injuries than concussion that resulted in observable anatomical abnormalities (12 out of 30 subjects), which have been linked to severe decreases in FC (Davey et al., 2000).

5.4.1 A New Model of mTBI

We propose a nuanced view of the progressive effects of mTBI on the human brain. Specifically, we posit a model of the dynamic effects of mTBI on event-related FC that accounts for stage of injury (acute, post-acute, and chronic), age, and task-related cognitive load. A diagram of the theorized model, its three stages, and corresponding effects are illustrated in Fig. 5.3. We describe our model in terms of its three stages below.

In the acute stage of the model, the brain enters a state of hyperactivity and hyperconnectivity immediately after injury (though we acknowledge that the mechanism for this is not entirely understood) that is reflected in a number of studies in the acute stage of concussion (Johnson et al., 2015; Iraji et al., 2016; Shumskaya et al., 2012; Zhou et al., 2014; Zhu et al., 2015). The proposed model theorizes that an increased recruitment of neuronal resources is particularly observable after injury, requires minimal task complexity, and can even be observed at rest (see Fig. 5.3, red). Moreover, this stage is hypothesized to overlap with the period prior to clinical symptom recovery, as well as cognitive deficits observed in neuropsychological and other behavioural batteries (McCrory et al., 2017). Since the increase in activity and FC observed in the acute stage can be observed in fMRI while the subject is at rest, it might be unrelated to cognitive load, and may reflect the increased internal rumination and arousal often seen in recently concussed individuals (Zhu et al., 2015; Sours et al., 2015). Moreover, hyperactivity and hyperconnectivity has also been shown in animal studies of concussion, where the effect is described as part of a post-concussion metabolic cascade (Giza and Hovda, 2001). Our model instead suggests that this brain response plays a functional role in the brain adaptation to compensate for lost and impaired functionality. We posit this hyperactivity and hyperconnectivity to facilitate a search for suitable ways of rerouting information processing to enable optimal compensation for the reduced cognitive functioning that resulted from injury, akin to the hyperconnected state observed in infants prior to neural pruning and serving a similar purpose (Huttenlocher et al., 1987; Rakic et al., 1994).

In typical cases of concussion, the brain is able to adapt and progress past the acute stage, signified by a relief of symptomatology and other observable cognitive deficits (McCrory et al., 2017). Our model posits that this post-acute stage, and thus behavioural symptom resolution, co-occurs with the re-normalization of resting-state functional brain activation and connectivity. This is supported by a lack of significant differences in resting-state FC between controls and subjects with past concussion in the post-acute stage found in a number of fMRI and EEG studies alike (McAllister et al., 2001; Bharath et al., 2015; Churchill et al., 2017; Nuwer et al., 2005). However, the brain's solution to counteract mTBI-related cognitive loss is not perfect. A high processing load enforced by a complex task requires the injured brain to compensate, and thus activate more than the uninjured brain, in order to maintain task performance (see Fig. 5.3, green) (McAllister et al., 2001). Notably, we identify the mechanism of compensation as a likely explanation for the apparent normalization of behaviourally-measured cognitive functions after symptom resolution (Martini et al., 2017). As such, our model is in line with a previously posited hypothesis that suggests that even though behaviourally-observed symptoms of concussion appear to resolve, the brain does not truly recover from the injury (De Beaumont et al., 2009).

Lastly, as part of a typical aging process, the brain progressively loses its ability to sustain the compensatory strategy developed in the acute stage and maintained throughout the post-acute stage (see Fig. 5.3, blue). This reduction and resurgence of symptoms is posited to be strongly linked to the notion of a cognitive reserve (Kesler et al., 2003; Stern, 2009). This chronic stage marks a downward trend in brain activity and functional connectivity. As no fMRI FC research was found targeting the chronic stage of injury (Henry et al., 2017), we base this stage of the model on previous ERP studies conducted by our group and others (described below), the observed resurgence of behavioural and cognitive symptoms in old age, and the results presented in this paper. Additionally, we posit this stage may coincide with reported findings of tauopathy in retired athletes and potential link to chronic traumatic encephalopathy (CTE; Stern et al., 2019).

5.4.2 ERP-specific Implications

In addition to the neuroimaging research discussed above, there has been a considerable amount of research on the effect of mTBI on ERPs. One challenge in understanding mTBI is to explain why some ERPs appear to be affected in consistent ways across the different stages of injury, while the effects on other ERPs appear to change over time. Our model provides a means to interpret these findings.

There is broad consensus regarding the effect of concussion on the P300, which is manifest as an attenuated and often delayed peak (Brush et al., 2018). This effect is demonstrably consistent across acute, post-acute, and chronic stages of injury (Ruiter et al., 2019; Broglio et al., 2009; Baillargeon et al., 2012; inter alia). In contrast, the effects are not consistent on earlier ERP components, such as the mismatch negativity (Ruiter et al., 2019) and the N1/P2 complex (Boshra et al., 2019). These responses were unaffected in acute and post-acute concussed subjects (Ruiter et al., submitted; Boshra et al., submitted), but were attenuated in individuals that had sustained their concussions decades earlier (i.e., in the chronic stage of concussion). Our model predicts that this ERP component would be enlarged in the acute stage while at rest and in the post-acute stage while under cognitive load. Of note, it has been argued that the emergence of deficits in these earlier ERP components may be associated with an irreversible return of concussion-related cognitive decline (Ruiter et al., submitted).

The N2b response is also reported to be affected in a way that changes through the stages of concussion. This component was found to be unaffected in some cases in the post-acute stage (Moore et al., 2014), amplified in other cases in the post-acute stage (Ledwidge and Molfese, 2016; Moore et al., 2014, 2015), and attenuated in the chronic stage (De Beaumont et al., 2009; Ruiter et al., 2019; see Brush et al., 2018 for a review). The apparent modulation of the N2b proves similar to results from FC in fMRI. Our model predicts an enlarged N2b in the acute stage, as well as in the post-acute stage when task-related cognitive load is sufficient to elicit neural overcompensation (Ledwidge and Molfese, 2016; Moore et al., 2014, 2015). The model expects an unobserved N2b response when the task is sufficiently simple (see oddball task in Moore et al., 2014). Lastly, the model predicts an attenuated N2b in the chronic stage, concordant with De Beaumont et al. (2009); Ruiter et al. (2019).

In summary, there is an indication that ERPs may be affected by mTBI in one of three ways. For the P300, the effect appears to remain consistent throughout all stages of concussion. For the MMN and the N1/P2 complex, the effects seem to appear only in the chronic stage, possibly indicating more severe consequences of the injury (Ruiter et al., submitted;Boshra et al., 2019). Finally, for the N2b, the effect is more complex: the component appears to be affected in different ways depending on the stage of concussion and the degree of cognitive load under which the EEG is recorded. This makes the N2b a potential target for monitoring the progression of concussion. Furthermore, the time-course of the effects of concussion on the N2b most closely matches what has been observed in the fMRI literature and the FC findings in the present study showing hyperconnectivity immediately after injury followed by a tendency to normalize unless the participant is subjected to mentally-taxing task.

5.4.3 General Model Implications

To conserve simplicity, the model does not account for many of the factors argued to influence the state of concussion, such as the number of previous concussions, the severity and location of impact, age at the time of injury, and more. However, while the model does not provide precise predictive power at this lower resolution, it provides an overarching view of the brain's response to concussion, and may even extend to brain injury more broadly, depending on the severity and type of injury. Moreover, we argue that the model provides, for the first time, a falsifiable explanation of the time-course of the effects of concussion on the brain that directly relates three primary modalities used in concussion research: behavioural assessments, fMRI, and EEG/ERP. Our model also synthesizes some of the inconsistencies found in the broader literature, particularly the seemingly contradictory findings between studies exploring resting-state brain activity versus task-based studies that impose a high processing load on the participants. The model also relates to the literature on post-concussion syndrome (long-term persistence of symptoms of concussion) as well as of more severe TBI. Both of these phenomena can be described as types of injuries in which the post-acute phase is skipped; post-concussion syndrome because the brain fails, for one reason or another, to compensate for the injury, and TBI because the injury is too severe for the brain to compensate in principle. Instead, the brain progresses directly from the acute stage to the chronic stage, as evidenced by studies that show reduced FC levels and clinical symptoms that are consistent with the chronic stage of concussion as described in our model (Kumar et al., 2009; Hocke et al., 2018b; Messé et al., 2013; Robinson et al., 2015).

While the present study sheds some light on both the early and late stages of mTBI, there are important questions that require data from individuals across the full spectrum of age groups representing a continuity of stages of injury (from acute to chronic and in between). Further data will enable an investigation of when hyperconnectivity subsides and clarify the interaction between normalizing resting-state connectivity, symptom resolution, and processing loads. Additionally, it is pertinent to identify what initiates the transition to a hypoconnected state later in life (time elapsed since injury, biological disposition, age, etc.), and whether that deterioration can be counteracted. This investigation is critical, as the decline into hypoconnectivity late in adulthood has been linked to severe consequences in other pathologies and was associated with more advanced forms of neurodegeneration (Hillary and Grafman, 2017). Moreover, as argued in Boshra et al. (2019), experimental designs that facilitate single-subject analyses are necessary in order to understand the inter-subject variability of mTBI and its effects, as well as to build towards a clinical tool leveraging these findings.

5.4.4 Limitations

One potential confound in our study was the vast difference in age between the CH and the AC study groups, leaving open the possibility that differences in FC could be naturally occurring due to age as opposed to injury stage. Indeed, when looking only at spectral coherence, our less robust measure of FC, we found a significant effect of Age. This effect was observable for both the control group and the patient group, and thus appears to be dominantly age-related and not mTBI-specific. In contrast, measuring WPLI revealed a clear pattern of changes in FC that was unique to the concussed group as can be seen in Figure 5.2B. In almost all cases, there was no change in WPLI-derived FC across age groups for the controls, while a pronounced decrease was notable in almost all comparisons for the concussed group. The two cases showing a main effect of Age in FC in the control group showed increases in FC (mid-range intrahemispheric and within-region FC), where the same measures showed an age-related decrease in FC. Altogether, this indicates that phase synchronization, as measured by WPLI, may be significantly reduced through aging, specifically for those with a history of concussion.

The present study was made possible by aggregating data from two studies using the same methodology, the same equipment, and that were collected by the same group. We note that combining two datasets as done in this study comes with the risk that statistical differences found in the study are due to heterogeneities between the datasets themselves. Specifically, the study samples differ by age, gender composition, and profession. In our statistical analyses, we mitigated this risk by adapting analysis procedures from earlier work to include a reduced number of comparisons, and by adopting a more conservative significance threshold (Kumar et al., 2009). Moreover, inclusion of data from both age-groups in the same models was done in an effort to control for the potential confound of age, as discussed above.

Another limitation is the use of coherence as a measure of FC in channel-space, which is sensitive to volume conductance. We validated the results with found with coherence by using a complimentary measure of FC, WPLI, which is invariant to volume conduction. WPLI was particularly useful in highlighting the changing effect of concussion on FC across the age groups. In summary, we were not able to control for potential confounds as mentioned above; however, we argue that if there was an effect of these variables, it would be minimal given the strength of the reported results (figure 5.2B).

5.5 Conclusion

The present study introduces a novel model of the time course of mTBI that synthesizes the neuroimaging and EEG literature, while generating targeted hypotheses that direct future research towards a more coherent understanding of mTBI. We argue that several points of contention, and apparently conflicting data, with respect to altered brain activity following concussion can be explained by a complex, dynamic, and non-linear response by the human brain that involves pathophysiological reactions, healing mechanisms, and compensatory actions. The study reported event-related FC as measured by the EEG during an active oddball task from 93 subjects. Our results support previous work indicating that concussion alters brain functional connectivity as measured by the WPLI and spectral coherence. Strikingly, our results demonstrated a clear shift in connectivity as acute effects of concussion gave way to chronic neurophysiological alterations, causing a salient switch from an increased to decreased FC in concussed individuals relative to age-matched controls, respectively. The study presents the first report linking established trends in fMRI and neuroimaging literature to cognitive function manifesting in ERPs. We conclude that further work to elucidate the dynamics of that trajectory towards failure to compensate is critical to understanding the mechanisms behind late-adulthood symptom reemergence, which are thought to be reflective of severe neuropathological consequences of concussion.

5.6 Acknowledgements

The authors gratefully acknowledge the participants for their time; Ms. Chia-Yu Lin and the entire Back-2-Play team for coordinating and organizing the sub-acute injury study; and Mr. Steve Buist as well as the Hamilton Spectator for coordinating and supporting the chronic injury study. This work was supported by the Canadian Institutes of Health Research (JFC), the Natural Sciences and Engineering Research Council of Canada (JPR), The Hamilton Spectator, Canada Foundation for Innovation (JFC), Senator William McMaster Chair in Cognitive Neuroscience of Language (JFC), the MacDATA fellowship award (RB), the Vector Institute postgraduate affiliate award (RB), and the Ontario Ministry of Research and Innovation (RB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

			DELTA BAND				THETA BAND				ALPHA BAND				BETA BAND			
Interhemisphere Connectivity			Coherence		wPLI		Coherence		wPLI		Coherence		wPLI		Coherence		wPLI	
Predictor	dfNum	dfDen	F	p	F	p	F	p	F	p	F	p	F	p	F	p	F	p
Group	1	89	2.84	0.095	0.42	0.518	2.2	0.142	0.37	0.547	0.28	0.598	1.98	0.163	0.06	0.813	1.06	0.306
Age	1	89	2012	0	0.81	0.372	2507.94	0	2.32	0.131	1525.23	0	0.44	0.511	2084.18	0	1.24	0.269
Group x Age	1	89	4.92	0.029	2.26	0.137	6.7	0.011	7.36	0.008	2.99	0.087	1.85	0.178	0.58	0.448	0.51	0.476
Group x Condition	2	178	1.21	0.299	0.07	0.927	0.01	0.993	0.7	0.491	0.39	0.67	1.52	0.223	1.86	0.16	1.42	0.245
Age x Condition	2	178	79.27	0	1.61	0.204	104.6	0	0.43	0.643	84.59	0	0.77	0.459	187.81	0	1.52	0.222
Group x Site	2	178	0.76	0.467	0.66	0.516	1.01	0.362	0.21	0.772	0.86	0.418	0.49	0.585	1.81	0.173	0.74	0.461
Age x Site	2	178	55.43	0	1.3	0.274	208.11	0	1.31	0.27	335.65	0	2.26	0.116	471.64	0	0.33	0.69
Group x Age x Condition	2	178	0.48	0.621	2.03	0.136	0.36	0.693	0.05	0.943	0.9	0.404	1.44	0.24	1.6	0.206	0.54	0.581
Group x Age x Site	2	178	0.1	0.899	0.32	0.723	0.72	0.481	1.17	0.305	1.29	0.276	0.86	0.41	1.07	0.337	0.31	0.7
Intrahemisphere-long Connectivity			Coherence		wPLI		Coherence		wPLI		Coherence		wPLI		Coherence		wPLI	
Group	1	89	0.67	0.417	0.98	0.324	1.71	0.194	1.59	0.21	0.11	0.746	2.3	0.133	0.04	0.848	2.43	0.123
Age	1	89	594.03	0	0	0.944	1439.51	0	0.54	0.463	420.39	0	1.38	0.243	962.4	0	0.34	0.564
Group x Site	1	89	0.22	0.637	1.43	0.234	0.06	0.809	7.06	0.009	0.72	0.399	0.39	0.533	2.21	0.14	0.05	0.828
Age x Site	1	89	10.45	0.002	1.54	0.218	13.06	0	3.92	0.051	20.51	0	0.02	0.898	26.29	0	0.55	0.459
Group x Age x Site	1	89	8.14	0.005	5.21	0.025	7.03	0.009	1.11	0.295	1.73	0.192	0.01	0.932	0.33	0.565	3.27	0.074
Group x Condition	2	178	0.87	0.414	0.11	0.887	1.92	0.153	1.94	0.149	0.11	0.879	1.09	0.333	0.31	0.725	0.19	0.81
Age x Condition	2	178	110.46	0	0.1	0.896	77.89	0	3.21	0.046	6.37	0.003	0.79	0.441	84.93	0	0.69	0.492
Group x Age x Condition	2	178	0.47	0.616	2.03	0.136	0.35	0.686	4.08	0.02	0.44	0.621	0.6	0.531	0.34	0.705	0.16	0.842
Intrahemisphere-mid Connectivity			Coherence		wPLI		Coherence		wPLI		Coherence		wPLI		Coherence		wPLI	
Group	1	89	1.69	0.197	0.8	0.372	0.52	0.473	0.06	0.809	0.41	0.523	1.56	0.215	0.09	0.767	0.5	0.482
Age	1	89	1100.06	0	8.1	0.006	1806.52	0	26.98	0	918.3	0	5.6	0.02	1831.18	0	0.57	0.454
Group x Age	1	89	8.64	0.004	10.16	0.002	10.33	0.002	16.26	0	4.84	0.03	2.83	0.096	2.45	0.121	3.47	0.066
Group x Condition	2	178	1.3	0.273	0.22	0.797	0.64	0.526	0.3	0.738	1.26	0.284	2.72	0.076	1.38	0.255	0.1	0.895
Age x Condition	2	178	106.73	0	0.35	0.698	94.76	0	3.57	0.03	42.9	0	0.25	0.753	284.13	0	1.57	0.212
Group x Site	3	267	2.67	0.062	2.78	0.062	0.52	0.608	0.7	0.514	1.17	0.318	0.37	0.736	2.33	0.095	1.69	0.177
Age x Site	3	267	0.24	0.823	9.83	0	0.19	0.839	10.9	0	8.92	0	2.1	0.112	46.14	0	3.74	0.016
Group x Age x Condition	2	178	1.04	0.354	4.8	0.01	0.4	0.668	0.4	0.671	0.43	0.638	0.35	0.674	1.44	0.24	3.13	0.05
Group x Age x Site	3	267	3.65	0.021	4.87	0.008	0.68	0.519	0.66	0.536	0.26	0.818	0.7	0.525	1.74	0.175	0.64	0.569
Intrahemisphere-within Connectivity			Coherence		wPLI		Coherence		wPLI		Coherence		wPLI		Coherence		wPLI	
Group	1	89	0.83	0.363	1.15	0.287	0.21	0.645	0.14	0.713	0.8	0.374	2.36	0.128	0.18	0.674	0.79	0.376
Age	1	89	2795.02	0	37.52	0	4992.42	0	48.11	0	5638.83	0	11.11	0.001	5539.13	0	3.07	0.083
Group x Age	1	89	23.35	0	14.47	0	16.68	0	14.86	0	10.29	0.002	4.61	0.035	7.04	0.009	7.13	0.009
Group x Condition	2	178	3.8	0.025	0.45	0.635	0.48	0.62	0.07	0.936	1.15	0.317	2.63	0.077	1.01	0.363	0.1	0.899
Age x Condition	2	178	115.97	0	0.13	0.874	191.83	0	2.37	0.097	129.43	0	0.09	0.903	411.19	0	0.61	0.542
Group x Site	5	445	0.71	0.584	0.65	0.624	1.28	0.28	2.15	0.073	1.35	0.258	1.61	0.174	0.73	0.538	2.31	0.064
Age x Site	5	445	22.66	0	2.53	0.041	57.63	0	8.34	0	52.7	0	1.82	0.127	76.4	0	2.22	0.073
Group x Age x Condition	2	178	1.09	0.337	4.66	0.011	0.11	0.892	0.27	0.766	0.14	0.861	0.5	0.602	1.39	0.251	2.44	0.092
Group x Age x Site	5	445	0.77	0.542	1.3	0.27	1.59	0.18	2.1	0.079	0.85	0.472	1.21	0.305	0.23	0.886	2.78	0.031

Table 5.1: ANOVA tables for the connectivity types and bands for the WPLI and coherence values

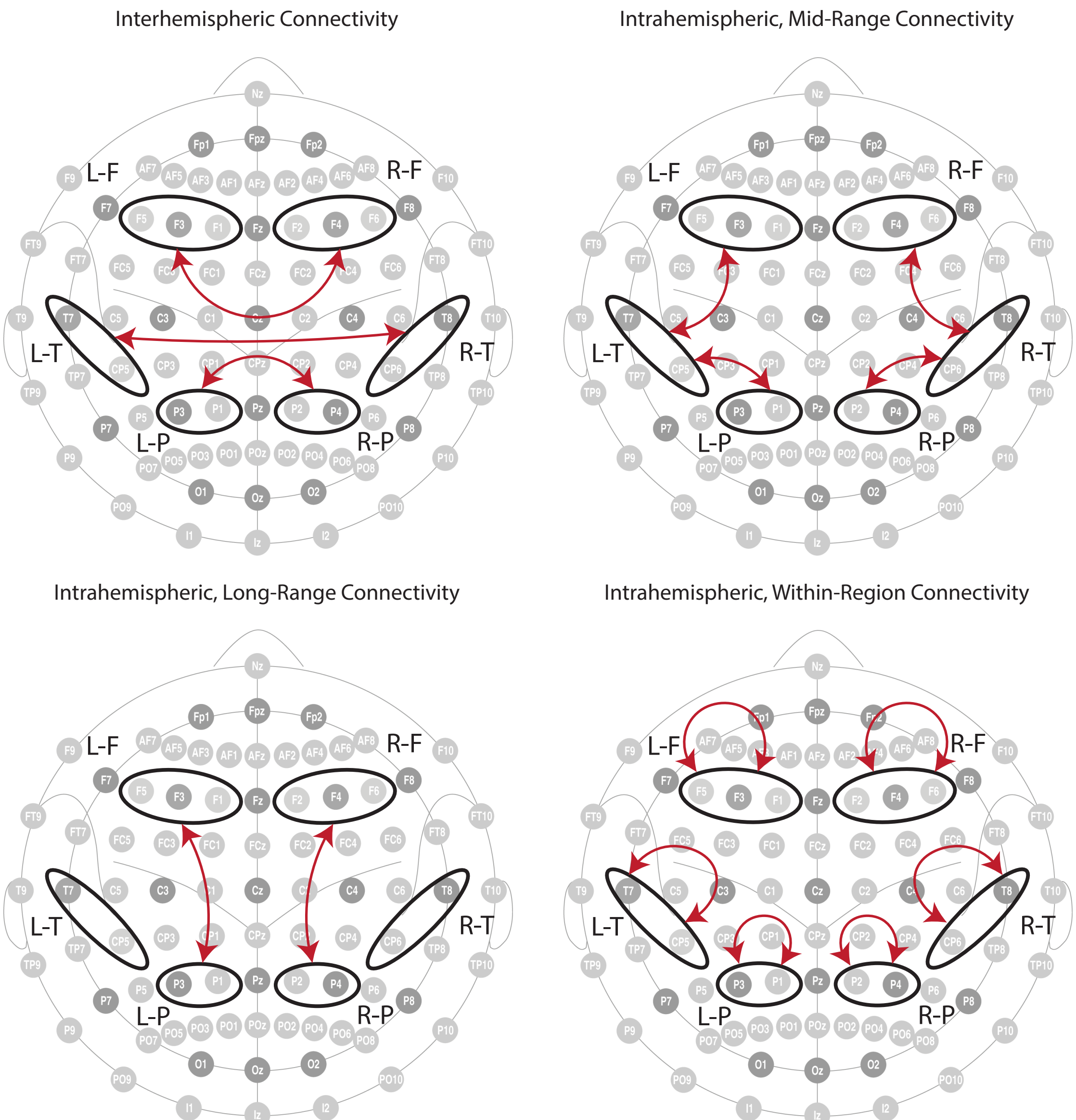

Figure 5.1: Different connectivity types and their respective electrode clusters as defined and adapted from Kumar et al. (2009).

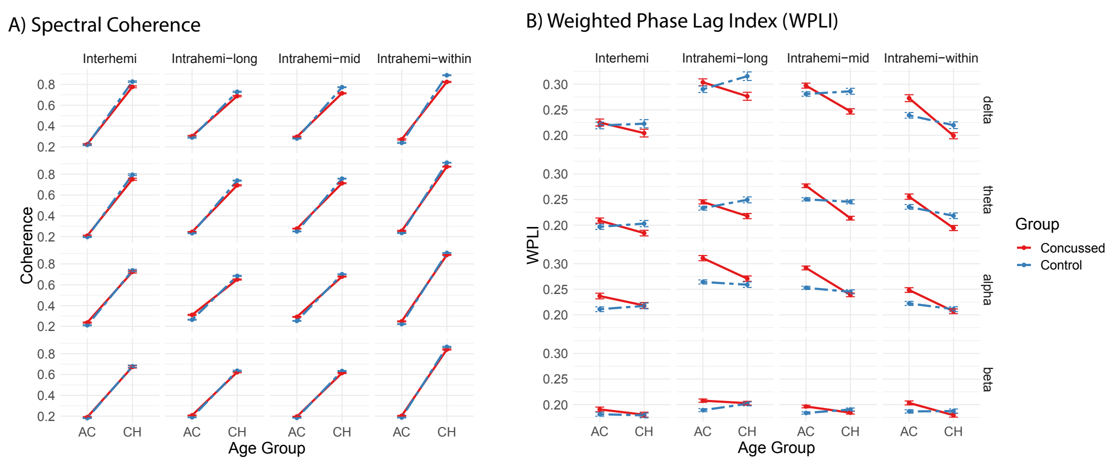

Figure 5.2: Group × Age interaction over different connectivity types as seen across the four bands in spectral coherence (A) and weighted phase-lag index (B). Error bars represent the 95% confidence intervals.

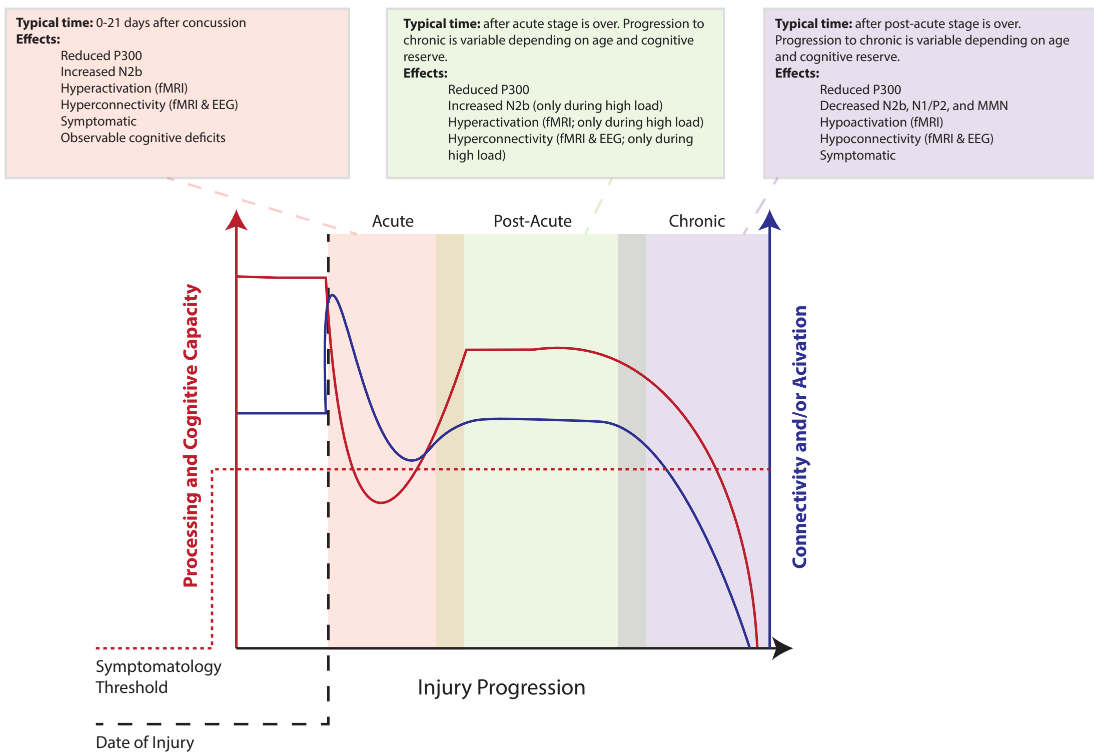

Figure 5.3: The theorized model with its three stages of injury progression: acute, post-acute, and chronic. Note the overlap between the stages in time, signifying an unclear transition point between them.

Bibliography

Baillargeon, A., Lassonde, M., Leclerc, S., and Ellemberg, D. (2012). Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Injury*, 26(3):211–220.

Bastos, A. M. and Schoffelen, J.-M. (2016). A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in Systems Neuroscience*, 9(January):1–23.

Bernier, R. A., Roy, A., Venkatesan, U. M., Grossner, E. C., Brenner, E. K., and Hillary, F. G. (2017). Dedifferentiation does not account for hyperconnectivity after traumatic brain injury. *Frontiers in Neurology*, 8(JUL):1–11.

Bharath, R. D., Munivenkatappa, A., Gohel, S., Panda, R., Saini, J., Rajeswaran, J., Shukla, D., Bhagavatula, I. D., and Biswal, B. B. (2015). Recovery of resting brain connectivity ensuing mild traumatic brain injury. *Frontiers in Human Neuroscience*, 9(September):1–13.

Blain-Moraes, S., Boshra, R., Ma, H. K., Mah, R., Ruiter, K., Avidan, M., Connolly,

J. F., and Mashour, G. A. (2016). Normal Brain Response to Propofol in Advance of Recovery from Unresponsive Wakefulness Syndrome. *Frontiers in Human Neuroscience*, 10(June):1–6.

Boshra, R., Dhindsa, K., Boursalie, O., Ruiter, K. I., Sonnadara, R., Samavi, R., Doyle, T. E., Reilly, J. P., and Connolly, J. F. (2019). From group-level statistics to single-subject prediction: Machine learning detection of concussion in retired athletes. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Macciocchi, S. N., and Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, 60(6):1050–7; discussion 1057–8.

Broglio, S. P., Moore, R. D., and Hillman, C. H. (2011). A history of sport-related concussion on event-related brain potential correlates of cognition. *International Journal of Psychophysiology*, 82(1):16–23.

Broglio, S. P., Pontifex, M. B., O'Connor, P., and Hillman, C. H. (2009). The Persistent Effects of Concussion on Neuroelectric Indices of Attention. *Journal of Neurotrauma*, 26(9):1463–1470.

Brush, C. J., Ehmann, P. J., Olson, R. L., Bixby, W. R., and Alderman, B. L. (2018). Do sport-related concussions result in long-term cognitive impairment? A review of event-related potential research. *International Journal of Psychophysiology*, 132(March 2017):124–134.

Cao, C. and Slobounov, S. (2010). Alteration of cortical functional connectivity as a result of traumatic brain injury revealed by graph theory, ICA, and sLORETA analyses of EEG signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(1):11–19.

Churchill, N., Hutchison, M. G., Leung, G., Graham, S., and Schweizer, T. A. (2017). Changes in functional connectivity of the brain associated with a history of sport concussion: A preliminary investigation. *Brain Injury*, 31(1):39–48.

Davey, M. P., Victor, J. D., and Schiff, N. D. (2000). Power spectra and coherence in the EEG of a vegetative patient with severe asymmetric brain damage. *Clinical Neurophysiology*, 111(11):1949–1954.

De Beaumont, L., Beauchemin, M., Beaulieu, C., and Jolicoeur, P. (2013). Long-term attenuated electrophysiological response to errors following multiple sports concussions. *Journal of Clinical and Experimental Neuropsychology*, 35(6):596–607.

De Beaumont, L., Brisson, B., Lassonde, M., and Jolicoeur, P. (2007). Long-term electrophysiological changes in athletes with a history of multiple concussions. *Brain Injury*, 21(6):631–644.

De Beaumont, L., Henry, L. C., and Gosselin, N. (2012). Long-term functional alterations in sports concussion. *Neurosurgical Focus*, 33(6):E8.

De Beaumont, L., Thoret, H., Mongeon, D., Messier, J., Leclerc, S., Tremblay, S., Ellemberg, D., and Lassonde, M. (2009). Brain function decline in healthy retired athletes who sustained their last sports concussion in early adulthood. *Brain*, 132(3):695–708.

Gaetz, M., Goodman, D., and Weinberg, H. (2000). Electrophysiological evidence for the cumulative effects of concussion. *Brain Injury*, 14(12):1077–1088.

Giza, C. C. and Hovda, D. A. (2001). The neurometabolic cascade of concussion. *Journal of athletic training*, 36(3):228.

Gosselin, N., Thériault, M., Leclerc, S., Montplaisir, J., and Lassonde, M. (2006). Neurophysiological Anomalies in Symptomatic and Asymptomatic Concussed Athletes. *Neurosurgery*, 58(6):1151–1161.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*.

Heitger, M. H., Jones, R. D., Macleod, A. D., Snell, D. L., Frampton, C. M., and Anderson, T. J. (2009). Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*, 132(10):2850–2870.

Henry, L. C., Tremblay, S., and De Beaumont, L. (2017). Long-Term Effects of Sports Concussions: Bridging the Neurocognitive Repercussions of the Injury with the Newest Neuroimaging Data. *Neuroscientist*, 23(5):567–578.

Hillary, F. G. and Grafman, J. H. (2017). Injured Brains and Adaptive Networks: The Benefits and Costs of Hyperconnectivity. *Trends in Cognitive Sciences*, 21(5):385–401.

Hocke, L. M., Duszynski, C. C., Debert, C. T., Dleikan, D., and Dunn, J. F. (2018a).

Reduced Functional Connectivity in Adults with Persistent Post-Concussion Symptoms: A Functional Near-Infrared Spectroscopy Study. *Journal of Neurotrauma*, 35(11):1224–1232.

Hocke, L. M., Duszynski, C. C., Debert, C. T., Dleikan, D., and Dunn, J. F. (2018b). Reduced Functional Connectivity in Adults with Persistent Post-Concussion Symptoms: A Functional Near-Infrared Spectroscopy Study. *Journal of Neurotrauma*, 35(11):1224–1232.

Huttenlocher, P. R. et al. (1987). The development of synapses in striate cortex of man. *Human neurobiology*, 6(1):1–9.

Iraji, A., Chen, H., Wiseman, N., Welch, R. D., O'Neil, B. J., Haacke, E. M., Liu, T., and Kou, Z. (2016). Compensation through functional hyperconnectivity: A longitudinal connectome assessment of mild traumatic brain injury. *Neural Plasticity*, 2016.

Johnson, B., Zhang, K., Gay, M., Horovitz, S., Hallett, M., Sebastianelli, W., and Slobounov, S. (2012). Alteration of brain default network in subacute phase of injury in concussed individuals: Resting-state fMRI study. *NeuroImage*, 59(1):511–518.

Johnson, B., Zhang, K., Hallett, M., and Slobounov, S. (2015). Functional neuroimaging of acute oculomotor deficits in concussed athletes. *Brain Imaging and Behavior*, 9(3):564–573.

Kesler, S. R., Adams, H. F., Blasey, C. M., and Bigler, E. D. (2003). Premorbid Intellectual Functioning, Education, and Brain Size in Traumatic Brain Injury:

An Investigation of the Cognitive Reserve Hypothesis. *Applied Neuropsychology*, 10(3):153–162.

Kumar, S., Rao, S. L., Chandramouli, B. A., and Pillai, S. V. (2009). Reduction of Functional Brain Connectivity in Mild Traumatic Brain Injury during Working Memory. *Journal of Neurotrauma*, 26(5):665–675.

Ledwidge, P. S. and Molfese, D. L. (2016). Long-Term Effects of Concussion on Electrophysiological Indices of Attention in Varsity College Athletes: An Event-Related Potential and Standardized Low-Resolution Brain Electromagnetic Tomography Approach. *Journal of Neurotrauma*, 33(23):2081–2090.

Martini, D. N., Eckner, J. T., Meehan, S. K., and Broglio, S. P. (2017). Long-term Effects of Adolescent Sport Concussion Across the Age Spectrum. *American Journal of Sports Medicine*, 45(6):1420–1428.

McAllister, T. W., Sparling, M. B., Flashman, L. A., and Saykin, A. J. (2001). Neuroimaging Findings in Mild Traumatic Brain Injury *. *Journal of Clinical and Experimental Neuropsychology*, 23(6):775–791.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R. C., Cassidy, D., Echemendia, R. J., Castellani, R. J., Davis, G. A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C. C., Guskiewicz, K. M., Herring, S., Iverson, G. L., Johnston, K. M., Kissick, J., Kutcher, J., Leddy, J. J., Maddocks, D., Makdissi, M., Manley, G. T., McCrea, M., Meehan, W. P., Nagahiro, S., Patricios, J., Putukian, M., Schneider, K. J., Sills, A., Tator, C. H., Turner, M., and Vos, P. E. (2017). Consensus statement on concussion in sport—the

5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, (October 2016):bjsports–2017–097699.

Messé, A., Caplain, S., Pélégrini-Issac, M., Blancho, S., Lévy, R., Aghakhani, N., Montreuil, M., Benali, H., and Lehéricy, S. (2013). Specific and Evolving Resting-State Network Alterations in Post-Concussion Syndrome Following Mild Traumatic Brain Injury. *PLoS ONE*, 8(6):1–10.

Moore, R. D., Broglio, S. P., and Hillman, C. H. (2014). Sport-related concussion and sensory function in young adults. *Journal of Athletic Training*, 49(1):36–41.

Moore, R. D., Pindus, D. M., Drolette, E. S., Scudder, M. R., Raine, L. B., and Hillman, C. H. (2015). The persistent influence of pediatric concussion on attention and cognitive control during flanker performance. *Biological Psychology*, 109:93–102.

Murias, M., Webb, S. J., Greenson, J., and Dawson, G. (2007). Resting state cortical connectivity reflected in eeg coherence in individuals with autism. *Biological psychiatry*, 62(3):270–273.

Nakamura, T., Hillary, F. G., and Biswal, B. B. (2009). Resting network plasticity following brain injury. *PLoS ONE*, 4(12).

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. (2004). Identifying true brain interaction from eeg data using the imaginary part of coherency. *Clinical neurophysiology*, 115(10):2292–2307.

Nuwer, M. R., Hovda, D. A., Schrader, L. M., and Vespa, P. M. (2005). Routine

and quantitative EEG in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025.

Rakic, P., Bourgeois, J.-P., and Goldman-Rakic, P. S. (1994). Synaptic development of the cerebral cortex: implications for learning, memory, and mental illness. In *Progress in brain research*, volume 102, pages 227–243. Elsevier.

Robinson, M. E., Lindemer, E. R., Fonda, J. R., Milberg, W. P., Mcglinchey, R. E., and Salat, D. H. (2015). Close-range blast exposure is associated with altered functional connectivity in Veterans independent of concussion symptoms at time of exposure. *Human Brain Mapping*, 36(3):911–922.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Shumskaya, E., Andriessen, T. M., Norris, D. G., and Vos, P. E. (2012). Abnormal whole-brain functional networks in homogeneous acute mild traumatic brain injury. *Neurology*, pages 175–182.

Sours, C., George, E. O., Zhuo, J., Roys, S., and Gullapalli, R. P. (2015). Hyper-connectivity of the thalamus during early stages following mild traumatic brain injury. *Brain Imaging and Behavior*, 9(3):550–563.

Stern, R. A., Adler, C. H., Chen, K., Navitsky, M., Luo, J., Dodick, D. W., Alosco, M. L., Tripodis, Y., Goradia, D. D., Martin, B., Mastroeni, D., Fritts, N. G., Jarnagin, J., Devous, M. D., Mintun, M. A., Pontecorvo, M. J., Shenton, M. E., and Reiman, E. M. (2019). Tau Positron-Emission Tomography in Former National Football League Players. *New England Journal of Medicine*, 380(18):1716–1725.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, 47(10):2015–2028.

Todd, J., Michie, P. T., Schall, U., Karayanidis, F., Yabe, H., and Näätänen, R. (2008). Deviant Matters: Duration, Frequency, and Intensity Deviants Reveal Different Patterns of Mismatch Negativity Reduction in Early and Late Schizophrenia. *Biological Psychiatry*, 63(1):58–64.

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.

Zhou, Y., Lui, Y. W., Zuo, X. N., Milham, M. P., Reaume, J., Grossman, R. I., and Ge, Y. (2014). Characterization of thalamo-cortical association using amplitude and connectivity of functional MRI in mild traumatic brain injury. *Journal of Magnetic Resonance Imaging*, 39(6):1558–1568.

Zhu, D. C., Covassin, T., Nogle, S., Doyle, S., Russell, D., Pearson, R. L., Monroe, J., Liszewski, C. M., DeMarco, J. K., and Kaufman, D. I. (2015). A Potential Biomarker in Sports-Related Concussion: Brain Functional Connectivity Alteration of the Default-Mode Network Measured with Longitudinal Resting-State fMRI over Thirty Days. *Journal of Neurotrauma*, 32(5):327–341.

CHAPTER 6 Summary and Conclusions

The current dissertation presented a series of independent studies that aimed to bring an expansive body of neurophysiology research another step towards a realized mTBI assessment tool. This work established the need to develop an expert model to capture EEG/ERP characteristics not readily attainable even by neurophysiology experts (Chapter 2). Motivated by these findings, the first ML study reported a successful application of an automated classification mechanism that is transparent in its decisions (Chapter 3). The learned model yielded promising performance and confirmed the hypothesis that the significant group effects from Ruiter et al. (2019) can be transferred to a more clinically-useful individualized assessment. The following study vastly expanded on the first, with an increased sample-size that enabled a more elaborate machine learning design. Results confirmed a definite divide between the concussed participants and controls on the single-subject level and established a fully-realized deep learning architecture, enabling an automated feature extraction procedure. The added layer of longitudinal comparison enabled direct comparison to potential correlations to symptomatology. Chapter 5 was a return to group-level statistics in an effort to answer a question of a wider scope – whether concussion induces dynamic or static effects on the human brain through time. The study found a critical change between immediate and long-lasting effects of mTBI. Further, a comprehensive view of the modern concussion literature was incorporated into a theoretical model of concussions. The model identifies a list of unknowns and offers a detailed framework for future question formulation with concrete falsifiable claims.

6.1 Summary of Findings

Taken as a whole, the work detailed in the present dissertation served to fill a number of gaps in the literature, replicate a number of findings, and create a preliminary theoretical model of mTBI based on an agglomeration of research studies. These findings are summarized here to address the dissertation's main goals (see Chapter 1).

1. While concussion's effect on ERPs is visually distinguishable on the group level, a human expert cannot reliably detect that with visual inspection on a single-subject basis. This is argued to result from a variety of factors that incrementally obscure an interpretation. First, while an expert can be argued to be the standard in detecting an all-or-nothing effect (per severe TBI cases), classification of a modulated signal is a vastly more complex problem; this is supported by the results in chapter 2. Second, given the non-static nature of mTBI effects on EEG/ERP, an expert might require information beyond the waveforms – e.g., age, time since injury, and medical history – to accurately identify response aberrations. Third, a clear understanding of the progression and its effects on EEG/ERPs is required, followed by a thorough transfer to clinical neurophysiologists to maximize their performance. Lastly, presented

results have shown that traditional practices in analyzing ERPs can be misleading when targeting the identification of a specific pathology. For instance, presented midline electrodes and single-deviant waveforms may simply not provide the necessary information to an expert observer (chapter 2). In essence, the work does not discount the necessity of human experts in formulating the parameters to maximize the performance of an EEG/ERP tool; however, an argument is made for an alleviation of the complexity of the problem using automated expert-systems especially designed for this application.

2. Explainable ML is of key importance in applications aimed at the clinical setting. Relating a black-box's output and input in a human-understandable format is essential for an expert's validation of the model. Moreover, model transparency enables the utilization of a trained model to identify features of interest. Our results emphasize this statement by a key finding from chapter 3. Investigating the decisions made by the trained model to achieve high performance indicated the implication of the N1/P2 complex, a low-level ERP, in long-lasting manifestations of concussion. Critically, we argue that the lack of reports on that effect is attributable to traditional ERP analysis' inherent bias towards a rigid set of ERPs investigated in concussion.

3. Identification of individual subjects afflicted with concussion is achievable using EEG/ERP in combination with ML methods. The presented work initiated and emphasized the utility of ML in ERP data in the detection of concussion across two critical populations: the recently afflicted and those with a history of repeated hits to the head. The promising performance across both groups urges for further work to incorporate and adopt the presented tools beyond research

studies, motivated by the potential critical benefit for the concussed population.

4. Responses to the same EEG/ERP designs alter dramatically at different states (acute vs. chronic) of injury. This was tested and partially confirmed using ML methods (unreported); however, a thorough examination of that interaction through event-related FC was detailed in chapter 5 and supported a complex view of mTBI's progression from the point of impact to ageing with the injury. Concretely, immediate effects of mTBI on the brain induce several layers of hyperconnectivity and hyperactivation whose mechanisms are not completely understood. These immediate effects are observable equally on RS EEG, EEG/ERP, neuroimaging, and behaviour. As the brain transitions into a more stable state, adapting with a non-transient yet unobserved damage to the brain, many of the previously mentioned effects fail to manifest unless the brain is put under a considerable cognitive load. That stage is referred to as the post-acute stage. Due to the consistent loss of cognitive capacity during aging, strategies to manage the unobserved damage become unsustainable, causing what is referred to as the chronic stage of mTBI. Thus, the author argues that a complex model of mTBI is necessary, both for a comprehensive understanding of the injury, as well as to guide future exploration of the problem-space using single-subject analyses.

5. Chapter 5 detailed the first investigation of WPLI in 94 subjects to examine the influence of concussion episodes on event-related FC. Results showed an increased overall connectivity observable directly after concussion. In contrast, subjects who had sustained their latest concussion an average of 28 years prior to testing exhibited a significantly decreased FC. In addition to complementing

the multiple facets of the proposed model in chapter 5, these results introduce event-related FC as an important tool in concussion research.

6.2 Scientific and Clinical Implications

In the move towards quantitative measures of concussion assessment, EEG/ERPs are one of the most promising, with studies observing concussion effects on ERPs across the entire timeline of injury progression and even after symptom resolution (Broglio et al., 2017; Ruiter et al., 2019; De Beaumont et al., 2009; Gosselin et al., 2012; De Beaumont et al., 2012). The present dissertation leveraged these ERPs in conjunction with ML to develop an accurate identification tool for concussion both in the acute stage (chapter 4) and later in the chronic stage (chapter 3). In contrast to previous work leveraging ML for concussion identification using RS, this methodology was based on a neurophysiological measure shown to sustain long after injury date (Nuwer et al., 2005; Rapp et al., 2015). Of interest, ERP results indicated higher performance even in acute cases compared to earlier ML work on mTBI (Prichep et al., 2012; Cao et al., 2008). This provides further evidence that the inherent complexity in the design and acquisition of EEG/ERP compared to RS EEG data is warranted and provides a significant benefit in mTBI identification.

Rapp and Curley (2012) have described the examination of ERPs in the context of mTBI as limited to amplitude and latencies of the defined components, suggesting that more intricate analysis of the waveforms may yield improvements in clinical utility. The present dissertation forms the closest instance in the literature of that exploration. Inspecting the waveforms beyond traditional analyses allowed for both a clearer view on the modulations of the ERPs to the pathology, as well as gave the motivation for an in-depth inspection of the progression of ERPs after injury and across the lifespan. Based on work presented in the earlier chapters, a theoretical model has been derived to represent the often incompatible reports from different mTBI assessment modalities: behavioural assessments, ERPs, RS EEG, and neuroimaging (primarily fMRI).

The model theorizes that immediately after concussion, the brain is in an unstable, hyperactivated state that is attempting to reconfigure and adapt to the injury (Iraji et al., 2016; Sours et al., 2015). In addition to the symptoms, cognitive deficits, and balance problems associated with concussive injury (McCrea et al., 2003; McCrory et al., 2017), most neuroimaging methods are able to observe this state of brain function as aberrant (Baillargeon et al., 2012; Ruiter et al., 2019a; chapter 4; chapter 5). Particularly, the brain is theorized to be in a general state of hyperactivity and hyperconnectivity (chapter 5). Beyond the acute stage, the model provides a concrete theory on the conflicting reports of behaviour vs. imaging, where numerous reports have shown behavioural signs of injury to normalize within a few weeks (McCrea et al., 2003; Martini et al., 2017; McCrory et al., 2017), while imaging has shown a sustained difference in observable effects despite symptom resolution (Cao et al., 2008; Guay et al., 2018; Broglio et al., 2011). The model attributes the normalization of symptoms to the brain's ability to reconfigure after the acute stage, enabling an allocation of resources to counteract the injury's damage. The ability of neuroimaging to still observe the effect is argued to be a factor of processing strain on the concussed brain (McAllister et al., 2001) – forcing the mechanism of allocation to execute and thus, its manifestation on ERPs such as the N2b. In contrast, the P300 was posited to be a more static ERP, with an attenuated effect that manifests and does not typically normalize after injury. Lastly, the model identifies the MMN and the N1/P2 complex, lower-level processing ERPs, as indicators of the brain's inability to sustain the compensatory allocation of cognitive resources, typically observed in older concussion cases (Ruiter et al., 2019; chapter 3). Despite the model's contextualization of several reports of disagreeing results, several points are acknowledged as unknown and require further research to clarify (discussed in depth below).

Subjective assessment of concussion has been numerously identified as fallible and inaccurate (Broglio et al., 2007). That is, a clinical judgment based on intangible evidence that is not quantifiable and rigorously optimized can be more akin to guessing the state of a patient (Broglio et al., 2017). Moreover, behavioural assessments, symptomatlogy scales, and pencil-and-paper tests are also susceptible to a variety of factors such as malingering, purposely scoring low on baselines, and a lingering degree of subjectivity. While adapting several tests for computer presentation was shown to improve sensitivity in some cases, there remains a degree of scrutiny regarding reliability and false-positive findings (Broglio et al., 2017; Ruiter et al., 2019). All the former issues notwithstanding, the detection of brain alterations despite symptom resolution and typical neurocognitive performance remains a concern for the dependence on behavioural measures (McAllister et al., 2001; Ruiter et al., 2019; Broglio et al., 2009). The present dissertation provided a detailed investigation of a methodology argued to be clinically underutilized in mTBI. With respect to EEG/ERP in general, the presented studies extend earlier reports of the cognitive responses' alterations after injury. Moreover, the extension of the results to single-subject analyses was able to concretely define sensitivity and specificity measures with over 80% in both acute and chronic cases of concussion. While EEG/ERP specifically wasn't the target of criticism in Nuwer et al. (2005), the authors presented a valuable, systematic analysis of how and why this methodology could be useful clinically. We argue that in conjunction with the expansive literature of ERP effects – despite some inconsistencies in the N2b, explainable by the theorized model – this dissertation provides strong evidence that qualifies EEG/ERP methods for larger-scale, multi-site validation studies to facilitate clinical adoption. Specifically, the combination of ERPs and state-of-the-art explainable machine learning models are argued to be a significant improvement over clinical standards of concussion management and identification.

6.3 Limitations

The four studies detailed in the present dissertation exhibited a number of limitations that warrant careful examination in light of the key implications highlighted above. While each of the studies contained more in-depth account of these limitations specific to the study, they are summarized here.

Due to logistical constraints commonly imposed on clinical research, the number of subjects that took part in the the studies presented here were limited. While the sample-size was sufficient for typical EEG/ERP studies, extending the methodology to leverage ML was not straightforward. This was especially pronounced in chapter 3 where a dataset of 39 observations (split across 2 classes) required rigorous enforcement of validation measures. Notably, this is a common problem faced by clinical application of ML (Combrisson and Jerbi, 2015; Miotto et al., 2017; Tzovara et al., 2013), rendering many of ML's state-of-the-art tools unusable and motivating the design of TRODNet in chapter 4. Despite the efforts to mitigate the sample-set problem, there are many research questions that can only be addressed with a more expansive dataset size.

Data heterogeneity was another factor limiting a lucid characterization of the findings as was discussed in chapter 5. Specifically, gender was not controlled in chapters 4 and 5; age varied between controls and the concussed group in chapter 4, and subsequently chapter 5; and etiology of injury was not controlled in all studies. While these are potential confounds with possible influence on the results, we have argued the effects to be minor, if not unlikely.

It is important to acknowledge a primary assumption for all the presented work: that group assignment (controls vs. concussed) was accurate. It can be argued that wrong assignment was less probable in the younger population study (chapters 4 and 5) due to them being clinical diagnosed prior to participation; however, the assumption that a professional football athlete had a history of concussion is equally plausible (chapters 2, 3, and 5). This is more difficult to argue for the two populations' respective control groups, where a participant's assignment was solely based on their report of not sustaining a head injury. Moreover, it is critical to acknowledge individual variation in a pathology that is largely not well-understood. Clearly, dissecting individuals into a binary problem (concussed vs. controls) is an oversimplification that is enforced to enable current work. With a lack of gold-standard for identifying and characterizing concussion, there remains considerable work before such a gradation can be realized (Broglio et al., 2017, 2007; McCrory et al., 2009, 2013, 2017).

6.4 Future Directions

The two datasets captured and analyzed in the present dissertation offer a wide view on the effects of concussion with a difference of more than 50 years in the participant ages between the two. However, there are many intermediate points that require further work to elucidate. In chapter 5, the theorized model leveraged previous work and current findings to predict the trajectory of ERP, event-related connectivity, and fMRI responses in the acute and post-acute stages broadly. However, it is critical to identify what defines the end of the acute stage and an empirically defined progression from that to the post-acute. Particularly, in addition to refining the proposed model of mTBI, potential correspondence between symptom alleviation and objective brain measurements can be invaluable in deciding concrete clinical guidelines of back-to-play and back-to-work.

Findings reported in chapter 5 offer a novel view of potential correspondence between EEG/ERP and fMRI results reported in the literature. The theorized model was developed based on an overview of previous work that did not directly record EEG/ERP and fMRI from the same set of participants. To confirm the posited linkage, a future study is needed such that a processing load is kept constant and brain responses are recorded, either simultaneously or sequentially from both modalities. Particularly, a single multi-modal investigation in acute patients has the potential to concretely establish a link between what is observed in individual ERPs, event-related FC, RS EEG, fMRI-derived FC, and fMRI brain activation, in addition to behavioural performance metrics and symptomatology scales.

The transition from post-acute to late-stage resurgence of adverse mTBI effects remains unclear. The theorized model predicts an unavoidable decline with aging; however, given that aging itself is very inconsistent across individuals, the model's prediction can be argued as specious. A prospective study on a concussed group late into adulthood is required to clarify the connection. We argue that studies such as that conducted by Martini et al. (2017), prospectively examining acute-stage mTBI to pre 40 year-old participants, can only observe behavioural effects once the timeline shifts later into adulthood. The proposed study is a substantial undertaking, with numerous variables that require capturing and proper statistical control; however, we argue that it is necessary in order to reach a mature understanding of concussion. Moreover, enabling long-term tracking beyond self-reports in chronic cases is critical in a world with an unprecedented increase in its geriatric population (Martel and Malenfant, 2010).

Irrespective of findings in intermediate stages between chronic and acute concussion, machine learning modeling based on ERP features has shown promise as a direct tool for a clinical application (see chapters 3 and 4). Rigorous validation and statistical confirmation metrics notwithstanding, a large-scale application of the developed models is necessary for a more accurate estimate of the models' generalization to unseen data. Particularly, a multi-site, multi-group study is proposed such that a model is applied to unseen data to measure accuracy, sensitivity, and specificity, deciding whether implementation of the models is clinically useful (Nuwer et al., 2005). Essentially, 3 factors should be assessed: 1) capacity for differential diagnosis/identification based on commonly occurring co-morbidities in mTBI (specificity), 2) performance reliability across different sites, and 3) improved sensitivity either compared to or in conjunction with other clinical tools. Moreover, a large scale study would confirm the validity of results given different etiologies, age-groups, and education levels, as well as provide insight on the impact of injury stage on model performance and most implicated ERP features.

The datasets analyzed in the present dissertation contained multiple paradigms; of which only one, the active multi-deviant oddball task, was used for all analyses discussed here. Of interest, the datasets contain responses to a passive variant of the oddball task with four times the number of trials. Previous work by our group has shown distinguishable long-lasting effects of mTBI in the retired athletes in the passive task (Ruiter et al., 2019). Moreover, the N1/P2 complex effects reported in the active task may be observed passively due to their elicitation irrespective of participant attentiveness (Näätänen and Picton, 1987). Future application of the present findings and modeling techniques on data from the passive oddball may provide useful insight, as well as improve performance when combined with data recorded actively. Further, responses to language comprehension have also been recorded. Some work has shown N400 effects following concussive hits to the head, motivating further investigation of these data (Fickling et al., 2019).

No investigation of single-subject effects was conducted using event-related FC differences found in chapter 5. An application of ML, as proposed in chapters 3 and 4, is warranted to investigate particular effects and potential clustering of FC effects on the individual level. Moreover, the effects found in chapter 5 suggest a potential improvement if FC features are to be added to candidate features for training, with further implications on the overall utility of this methodology in concussion.

All reported results argued for adverse effects of concussion on neurophysiology that are modulated by different injury stages. This hypothesis suggests a difficult generalization of a model trained on one group to apply to the other; however, as more data are collected and incorporated into model training, ensemble approaches can be invaluable in providing a comprehensive model that is applicable to a more heterogeneous population. A follow-up study incorporating data from both datasets is warranted, provided more data are collected to mitigate knowledge-leak bias, i.e., knowledge about the data as a whole additively and artificially increasing performance.

All ML methods applied in the presented investigations were supervised. Supervised ML imposes strict labeling on the dataset, which has been argued to be inherently inaccurate (see above). While supervised methods provide a simple framework to facilitate analyses, we posit that unsupervised learning is critical for an appropriate analysis of a pathology as complex as concussion. In brief, unsupervised methods provide the flexibility to capture patterns in the data not directly observable in the strict paradigm of supervised learning. For instance, cluster analyses can elucidate effects on the features beyond a binary classification of concussion history. Observed clustering patterns may relate to injury severity, injury location, biological disposition, age, and different co-morbidities – all important factors that add to concussion's complexity.

Information regarding prediction of outcome was limited provided the current datasets. Nonetheless, results from chapter 4 provided an interesting finding of participants that exhibited symptom resolution being distinguishable as such prior to that taking place. While this can be regarded as a critical finding, the exploratory nature of this result remains under scrutiny. Further, while identification of concussion at different stages is critical, a potential predictive capability of ERP or event-related FC may have serious beneficial consequences on the current standard of concussion management. This is most pronounced in two key cases. First, if ERP or FC signatures may indicate whether a concussion patient is on a good trajectory towards recovery, minimal intervention, and resources, would be allocated to their management. In contrast, early intervention for worse trajectories may prove beneficial to the patient. Second, an early detection of a marker for symptom resurgence later in life, despite the lack of symptoms at time of testing, may allow clinical intervention to either mitigate resurgence altogether or prolong the time of cognitive compensation. In essence, work is required to establish a direct neurological point of reference, invisible through behaviour, that enables empirical examination of intervention and rehabilitation efforts. Future work is required that prioritizes prospective cohort designs to provide the necessary data for such a seminal examination.

6.5 Concluding Note

The present dissertation provided the necessary groundwork to facilitate a viable migration of EEG/ERP from research context to the clinical setting. That was accomplished by training explainable machine learning models on a comprehensive set of datasets across two key injury stages: acute and chronic. Findings supported the ability of ML to accurately identify neurophysiological effects of concussion, while affirming the notion that concussion's adverse effects dynamically change after injury. The concluding study on the same data expanded the literature to encompass event-related functional connectivity, showing a hyperconnected brain directly after injury that diminishes to below-normal in the chronic stage. A theoretical, falsifiable model of mTBI and its progression has been proposed, providing a firm framework to synthesize future hypotheses spanning a multitude of modalities. While the work introduces a new set of unresolved questions and motivates a plethora of future studies, it provides a series of empirical, objective findings with the aim of revolutionizing current management strategies for the *silent epidemic*.

Bibliography

Baillargeon, A., Lassonde, M., Leclerc, S., and Ellemberg, D. (2012). Neuropsychological and neurophysiological assessment of sport concussion in children, adolescents and adults. *Brain Injury*, 26(3):211–220.

Broglio, S. P., Guskiewicz, K. M., and Norwig, J. (2017). If You're Not Measuring, You're Guessing: The Advent of Objective Concussion Assessments. *Journal of Athletic Training*, 52(3):160–166.

Broglio, S. P., Macciocchi, S. N., and Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, 60(6):1050–7; discussion 1057–8.

Broglio, S. P., Moore, R. D., and Hillman, C. H. (2011). A history of sport-related concussion on event-related brain potential correlates of cognition. *International Journal of Psychophysiology*, 82(1):16–23.

Broglio, S. P., Pontifex, M. B., O'Connor, P., and Hillman, C. H. (2009). The Persistent Effects of Concussion on Neuroelectric Indices of Attention. *Journal of Neurotrauma*, 26(9):1463–1470.

Cao, C., Tutwiler, R. L., and Slobounov, S. (2008). Automatic classification of athletes with residual functional deficits following concussion by means of EEG signal using support vector machine. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(4):327–335.

Combrisson, E. and Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, 250:126–136.

De Beaumont, L., Henry, L. C., and Gosselin, N. (2012). Long-term functional alterations in sports concussion. *Neurosurgical Focus*, 33(6):E8.

De Beaumont, L., Thoret, H., Mongeon, D., Messier, J., Leclerc, S., Tremblay, S., Ellemberg, D., and Lassonde, M. (2009). Brain function decline in healthy retired athletes who sustained their last sports concussion in early adulthood. *Brain*, 132(3):695–708.

Fickling, S. D., Smith, A. M., Pawlowski, G., Ghosh Hajra, S., Liu, C. C., Farrell, K., Jorgensen, J., Song, X., Stuart, M. J., and D'Arcy, R. C. (2019). Brain vital signs detect concussion-related neurophysiological impairments in ice hockey. *Brain*, 142(2):255–262.

Gosselin, N., Bottari, C., Chen, J.-K., Huntgeburth, S. C., De Beaumont, L., Petrides, M., Cheung, B., and Ptito, A. (2012). Evaluating the cognitive consequences of mild traumatic brain injury and concussion by using electrophysiology. *Neurosurgical Focus*, 33(6):E7.

Guay, S., De Beaumont, L., Drisdelle, B. L., Lina, J.-M., and Jolicoeur, P. (2018). Electrophysiological impact of multiple concussions in asymptomatic athletes: A re-analysis based on alpha activity during a visual-spatial attention task. *Neuropsychologia*, 108(October 2017):42–49.

Iraji, A., Chen, H., Wiseman, N., Welch, R. D., O'Neil, B. J., Haacke, E. M., Liu, T., and Kou, Z. (2016). Compensation through functional hyperconnectivity: A longitudinal connectome assessment of mild traumatic brain injury. *Neural Plasticity*, 2016.

Martel, L. and Malenfant, É. C. (2010). 2006 census: Portrait of the canadian population in 2006, by age and sex. *Statistics Canada. Web*, 10.

Martini, D. N., Eckner, J. T., Meehan, S. K., and Broglio, S. P. (2017). Long-term Effects of Adolescent Sport Concussion Across the Age Spectrum. *American Journal of Sports Medicine*, 45(6):1420–1428.

McAllister, T. W., Sparling, M. B., Flashman, L. A., and Saykin, A. J. (2001). Neuroimaging Findings in Mild Traumatic Brain Injury *. *Journal of Clinical and Experimental Neuropsychology*, 23(6):775–791.

McCrea, M., Guskiewicz, K. M., Marshall, S. W., Barr, W., Randolph, C., Cantu, R. C., Onate, J. A., Yang, J., and Kelly, J. P. (2003). Acute Effects and Recovery Time Following Concussion in Collegiate Football Players. *JAMA*, 290(19):2556.

McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R. C., Cassidy, D., Echemendia, R. J., Castellani, R. J., Davis, G. A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C. C., Guskiewicz, K. M., Herring, S., Iverson, G. L., Johnston, K. M., Kissick, J., Kutcher, J., Leddy, J. J., Maddocks, D., Makdissi, M., Manley, G. T., McCrea, M., Meehan, W. P., Nagahiro, S., Patricios, J., Putukian, M., Schneider, K. J., Sills, A., Tator, C. H., Turner, M., and Vos, P. E. (2017). Consensus statement on concussion in sport—the 5 th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, (October 2016):bjsports–2017–097699.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., and Cantu, R. (2009). Consensus Statement on Concussion in Sport – The Third International Conference on Concussion in Sport Held in Zurich, November 2008. *The Physician and Sportsmedicine*, 37(2):141–159.

McCrory, P., Meeuwisse, W. H., Aubry, M., Cantu, B., Dvořák, J., Echemendia, R. J., Engebretsen, L., Johnston, K., Kutcher, J. S., Raftery, M., Sills, A., Benson, B. W., Davis, G. A., Ellenbogen, R. G., Guskiewicz, K., Herring, S. A., Iverson, G. L., Jordan, B. D., Kissick, J., McCrea, M., McIntosh, A. S., Maddocks, D., Makdissi, M., Purcell, L., Putukian, M., Schneider, K., Tator, C. H., and Turner, M. (2013). Consensus statement on concussion in sport: the 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, 47(5):250–258.

Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, (February):1–11.

Näätänen, R. and Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*.

Nuwer, M. R., Hovda, D. A., Schrader, L. M., and Vespa, P. M. (2005). Routine

and quantitative EEG in mild traumatic brain injury. *Clinical Neurophysiology*, 116(9):2001–2025.

Prichep, L. S., Jacquin, A., Filipenko, J., Dastidar, S. G., Zabele, S., Vodencarevic, A., and Rothman, N. S. (2012). Classification of Traumatic Brain Injury Severity Using Informed Data Reduction in a Series of Binary Classifier Algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(6):806–822.

Rapp, P. E. and Curley, K. C. (2012). Is a diagnosis of mild traumatic brain injury a category mistake? *Journal of Trauma and Acute Care Surgery*, 73(2 SUPPL. 1):13–23.

Rapp, P. E., Keyser, D. O., Albano, A., Hernandez, R., Gibson, D. B., Zambon, R. A., Hairston, W. D., Hughes, J. D., Krystal, A., and Nichols, A. S. (2015). Traumatic Brain Injury Detection Using Electrophysiological Methods. *Frontiers in Human Neuroscience*, 9(February):1–32.

Ruiter, K. I., Boshra, R., Doughty, M., Noseworthy, M., and Connolly, J. F. (2019). Disruption of function: Neurophysiological markers of cognitive deficits in retired football players. *Clinical Neurophysiology*, 130(1):111–121.

Sours, C., George, E. O., Zhuo, J., Roys, S., and Gullapalli, R. P. (2015). Hyperconnectivity of the thalamus during early stages following mild traumatic brain injury. *Brain Imaging and Behavior*, 9(3):550–563.

Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., and De Lucia, M. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, 136(1):81–89.