UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Real-time Mental Workload Detection and Alert System with Brain Computer Interface for Augmenting Human Performance at Work

by Alka Rachel John

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of
Distinguished Professor Chin-Teng Lin

University of Technology Sydney
Faculty of Engineering and Information Technology

January 2023

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, *Alka Rachel John*, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the *School of Computer Science*, *Faculty of Engineering and IT* at the University of Technology Sydney. This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

SIGNATURE: Production Note: Signature removed prior to publication.

[Alka Rachel John]

DATE: 19th January, 2023

Real-time Mental Workload Detection and Alert System with Brain Computer Interface for Augmenting Human Performance at Work

A thesis submitted in partial fulfilment of the requirements
for the degree of

Doctor of Philosophy

by

Alka Rachel John

to

School of Computer Science
Faculty of Engineering and Information Technology
University of Technology Sydney
NSW - 2007, Australia

January 2023

Abstract

People often find it frustrating to perform high-demand tasks for prolonged periods of time. However, it is not always possible to avoid these types of tasks in certain workplaces, making it essential to design a work environment that effectively interacts with the capabilities of the human worker.

One of the key factors that impact the efficiency of human operators in complex, interactive work environments is the level of mental workload. When an operator is overwhelmed with mental demands, it can negatively impact their performance and lead to mistakes, which can harm the system's efficiency and safety. To ensure optimal performance, it is important to find a balance and avoid both overload and underload conditions, as the operator's performance can suffer in both situations.

This work explored the feasibility of employing electroencephalogram (EEG) data in measuring the workload of human operators in safety-critical work environments. Two safety-critical environments were considered: a stationary Air Traffic Control (ATC) and a more dynamic physical Human Robot Collaboration (pHRC) environment with uncontrolled, real-world physical interactions with an abrasive blasting robot. Further, we explored the error awareness of operators engaged in a physical interaction with the robot in under varying workload conditions.

We successfully uncovered EEG spectral power, eye, and heart rate variability correlates of mental workload variations for simple tasks of air traffic controllers, providing a comprehensive understanding of the workload demands in ATC tasks. Our preliminary findings in the ATC experiment pave the way to develop intelligent closed-loop mental workload-aware systems for ATC.

The systematic investigation into the impact of workload variations on operator's performance in a pHRC setting revealed that both task and physical performance degraded with increasing workload. Our study successfully isolated and retrieved the biomarkers of workload variations in a pHRC despite strenuous physical activity. Error awareness was found to deteriorate with increasing workload, exposing the significance of measuring and maintaining the human user's workload at an optimal level to ensure the safe and reliable use of BCI technology for intuitive physical human-robot collaboration.

Therefore, this research posits that the mental workload of a human operator can be predicted in real-time from physiological signals, and the estimated workload information can be used to provide safety alerts, enhancing the safety measures in place in many work environments. By utilizing newly discovered biomarkers, a tool can be developed to detect workload variations and estimate error processing capacity, facilitating a mental

workload-related safety alert system for real-world applications. Thus, this research aims to demonstrate the potential of creating an intelligent system that can detect mental workload in order to improve safety and efficiency in complex work environments.

Dedication

To my Yeshuappachen, ammachy, appas, ammas, achachan, Alin vava and Professor CT Lin for all that you have done for me ...

ACKNOWLEDGMENTS

In this section, I would like to express my sincerest gratitude to the individuals who have helped me bring this thesis to fruition. I am especially honored to have been mentored by one of the most respected and accomplished academics in my field, my principal supervisor, Distinguished Professor Chin-Teng Lin. His extensive research in BCI and AI provided me with a wealth of knowledge and resources, and he expertly guided me through the research process. I am grateful for his substantial contributions to BCI research, which motivated me to focus on developing a tool for mental workload assessment and detection. I am also grateful for the valuable insights and wisdom shared during our conversations, which continue to inspire new ideas and perspectives. His guidance has helped me think beyond my research and strive to solve real-world problems. His passion, patience, and persistence have been a constant source of inspiration throughout my PhD journey. I am truly grateful to him for being such an inspiring role model and for showing me that hard work truly pays off. I owe a great deal to my Professor!

I would like to express my gratitude to my co-supervisors, Dr. Avinash Singh and Dr. Yu-Kai Wang, for their guidance and support throughout my research process. Dr. Avinash provided valuable directions on my research and how to effectively present my findings. His perceptive inquiries during our research meetings refined the aspects of my work and presentation, and I am genuinely appreciative of how it enriched my research. I am deeply appreciative of Dr. Wang's encouragement and enthusiasm, which helped me to persevere even during challenging times. I also thank him for his valuable feedback on my research and his uplifting words.

I would like to extend my appreciation to Professor Klaus Gramann from TU Berlin for his guidance in my EEG analysis, his review of my manuscripts, and his detailed feedback which improved the quality of my work. I am also grateful to Professor Dikai Liu for his guidance in my study of physical human-robot collaborations, and for providing support and access to the robot for my experiments. Additionally, I would like to thank Professor Ami Eidels and Professor Eugene Nalivaiko for their valuable feedback on my research on Air Traffic Controllers. I am grateful for the feedback and fresh perspectives provided by my doctoral committee members, including Dr. Helen Lu, Dr. Yulei Sui, Dr. Jaime Garcia Marin, and future assessors of this thesis.

I consider myself very fortunate to have had the opportunity to work with such accomplished academics. I am grateful for their willingness to review each manuscript I prepared and for their valuable suggestions to improve my findings. Without their

dedicated time and efforts, this thesis would not have been possible. I extend my heartfelt thanks to all of them.

The research presented in this thesis builds upon the contributions of many BCI researchers, and I am deeply grateful to them for their work in advancing the field. I would like to especially thank the authors of the publications cited in this thesis for their significant efforts in furthering BCI research. I also extend my thanks to the journal editors and reviewers for their feedback on the manuscripts submitted from this thesis work.

I am deeply grateful to all the participants in the experiments reported in this thesis, as their participation was essential for the data that formed the basis of this work. I would like to express my appreciation to Dr. Nguyen Tien-Thong Do for his guidance with EEG analysis. I also would like to extend my thanks to Dr. Zehong Cao from the University of South Australia for introducing me to this field of research.

I am deeply grateful for the support and assistance provided by my friends in the Computational Intelligence and Brain Computer Interface (CIBCI) lab during this journey. I would like to express my appreciation to Dr. Jia Liu and Yanqiu Tian for their help and guidance in a wide range of areas, including research, and for the valuable insights and memories they shared with me. I would also like to thank all the other scholars of CIBCI and the School of Computer Science for their help throughout this journey. I would also like to acknowledge Professor Paul Kennedy, the Head of the School of Computer Science, for the facilities and support provided to me while I was a student in the school.

I am deeply grateful to the Australian Artificial Intelligence Institute for recognizing my research and for providing motivation to continue my efforts by awarding me the best paper in the Applications in Artificial Intelligence category in 2022. I also extend my appreciation to the Women in Engineering and Information Technology at UTS for recognizing my PhD research.

I am grateful to the University of Technology for providing me with the necessary resources and facilities to conduct this thesis. I also appreciate the financial support from UTS which made this research possible. I would like to express my appreciation to the Australian Research Council for funding my research throughout this journey.

I am deeply grateful to my family for their unwavering support and encouragement throughout this journey. I would like to extend my special thanks to my husband, who has been my rock and support in every sense. I am grateful for his ability to provide a sense of calm and stability in times of stress and for being my constant cheerleader. He knew how to comfort me during times of failure and motivated me to keep pushing forward. I am grateful for his unconditional love and support and for accepting all of my strengths and weaknesses. Without his support, this journey would not have been as fulfilling.

I extend my deepest appreciation to my grandmother for her love and support throughout this journey. She is a woman, a mother, and a grandmother that I look up to and aspire to be like. I thank my father for recognizing my abilities and for pushing me to achieve my dreams. He always encouraged me to do my best, and I am grateful for his efforts in preventing me from settling for mediocrity. I am deeply grateful to my

mother for her unwavering support and selflessness. She has been an outstanding role model, always putting my needs and desires before her own. Her tireless sacrifices have brought endless joy and happiness to my life. I am also grateful to my younger brother, Alin, for being an amazing brother. Despite his young age, he has always protected and supported me in countless ways, always respected my decisions, and went out of his way to let me know that I am always appreciated. I also want to express my gratitude to my sister-in-law for loving my brother and my family and for bringing so much joy to our family just by being a part of it.

I am truly grateful to my father-in-law for his love and support. His positive outlook on life and belief in me has been a constant source of inspiration and motivation. My mother-in-law's unwavering support and constant acts of kindness have been invaluable to me throughout this journey. I am grateful for the values she instilled in her children, which have greatly benefited me. I also want to thank the rest of my wonderful family - chechi, Jaicks achachan, Hanna chechi, Maccu, Sam, and Joseph vava for their patience, love, and wholehearted support throughout this journey.

I am incredibly proud of my family for their unwavering support throughout my research journey, despite the challenges posed by COVID-19 travel restrictions, which kept us apart for over three years. Their patience, love, and understanding have been invaluable, and I could not have achieved this without them.

Last but most importantly, I thank my Yeshuappachen for His amazing grace and steadfast love. Even when I am unworthy, He continues to love me and redeem me when I stray. From forming me in my mother's womb to this day, He has always carried me on His shoulders, in the light of His face. His signature is evident in everything that I have, and I will thank Him with all of my being.

Alka Rachel John

List of Publications

Related to the Thesis :

1. John, A. R., Singh, A. K., Do, T. T. N., Eidels, A., Nalivaiko, E., Gavgani, A. M., ... & Lin, C. T. (2022). Unraveling the Physiological Correlates of Mental Workload Variations in Tracking and Collision Prediction Tasks. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 770-781.

2. John, A.R., Singh, A.K., Do, T.T.N. & Lin, C.T. (2021), Exploiting Neurometrics to Design Intelligent Adaptive System for Efficient Air Traffic Control. Australasian Mathematical Psychology Conference.

3. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T., "Effective Connectivity of the Brain under Multi-Workload Conditions during Physical Interaction with Robot" (in preparation).

4. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T., "Effective Connectivity of the Brain under Cognitive Conflict in Different Workload Conditions during Physical Human Robot Collaboration" (in preparation).

5. Lin, C.T. and John, A.R, Towards the Design of BCI-based Accelerated Training System for Air Traffic Controllers, IEEE Brain Newsletter, 2022.

Under Review and Drafted :

1. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T. Human Awareness in Physical Human Robot Collaboration: A Survey, Submitted to the Special Issue on Artificial Intelligence for Human-Robot Interaction (AI-HRI), ACM Transactions on Human-Robot Interaction (under review).

2. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T. Mental Workload Correlates of Physical Performance in a Physical Human-Robot Collaboration, Submitted to the Human Brain Mapping journal.

3. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T. Effects of Mental Workload Variations on Cognitive Conflict during Unexpected Robot Behavior in a Physical Human-Robot Collaboration (under review in Journal of Neural Engineering).

4. John, A. R., Singh, A. K., Liu, D., Gramann, K., & Lin, C. T., "EEG-based Safety Alert Tool in Physical Human Robot Collaboration" (drafted).

5. John, A. R., Singh & Lin, C. T., "EEG-based task Performance drop Alert System for Simple Tracking and Collision Prediction Tasks" (in preparation).

OTHERS :

11. King, J.T., John, A.R., Wang, Y.K., Shih, C.K., Zhang, D., Huang, K.C. and Lin, C.T., 2022. Brain Connectivity Changes During Bimanual and Rotated Motor Imagery. IEEE Journal of Translational Engineering in Health and Medicine, 10, pp.1-8.

12. Cao, Z., John, A.R., Chen, H.T., Martens, K.E., Georgiades, M., Gilat, M., Nguyen, H.T., Lewis, S.J. and Lin, C.T., 2021. Identification of EEG Dynamics During Freezing of Gait and Voluntary Stopping in Patients With Parkinson‚Äôs Disease. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29, pp.1774-1783.

13. Huang, K.C., John, A.R., Jung, T.P., Tsai, W.F., Yu, Y.H. and Lin, C.T., 2021. Investigating brain activities and neural comodulations in motion sickness between passengers and drivers. IEEE Transactions on Neural Systems and Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society.

14. Lin, C.T., King, J.T., John, A.R., Huang, K.C., Cao, Z. and Wang, Y.K., 2021. The impact of vigorous cycling exercise on visual attention: a study with the BR8 wireless dry EEG system. Frontiers in neuroscience, p.70

15. Cao, Z., John, A.R., Chen, H.T., Martens, K.E., Georgiades, M., Gilat, M., Nguyen, H.T., Lewis, S.J. and Lin, C.T., Prediction of Onset of Freezing of Gait Parkinson‚Äôs Disease Patients. Applied Sciences.

TABLE OF CONTENTS

List of Figures

Figure **Page**

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Background

Individuals often shy away from tasks that push their abilities to the brink as they find it frustrating and stressful [3]. However, not all work environments afford this luxury, highlighting the importance of fostering good interaction between the human operator's abilities and the work environment [359]. Though human operators have the capacity to adapt to various work environments, perform multiple tasks, and utilize different equipment simultaneously, poorly designed work environments can lead to an overload of sensory information and an increase in workload.

Mental workload is a crucial factor that impacts the efficiency of human operators in complex interactive work environments. In recent years, the role of human operators has shifted to a supervisory level [299], requiring them to integrate multiple streams of information which leads to an increase in cognitive resources [269] and subsequently in mental workload for operators [196, 248]. Wickens and Tsang [355] define mental workload as the dynamic relationship between the cognitive resources required by a task and the operator's ability to provide those resources. Human operators have limited information processing abilities as their resources have a finite capacity [169, 192].

The theory of limited cognitive resources posits that exposure to demanding task conditions can impair performance by depleting resources [261] or limiting access to them [43]. High levels of mental workload can lead to human errors [294], compromising system efficiency and safety [368]. To maintain optimal performance, the mental work-

load should be kept at a balanced level, avoiding both underload and overload [41, 137], as performance can suffer in both of these scenarios [48, 138, 341, 371]. The dynamic adaptive theory suggests that the brain seeks a balance of resources and cognitive comfort, and extremely high or low task demands can hinder adaptability and performance [139]. By predicting mental workload and adjusting task allocation, it is possible to avoid a loss of situational awareness and maintain high performance. Accurate and reliable measurement of mental workload is crucial, especially in safety-critical work environments, in order to improve work environments and human-machine interactions [12, 46, 309].

Researchers have employed various strategies, including self-assessment, performance measures, and physiological metrics, to evaluate mental workload; however, each of these methods has its own set of advantages, and disadvantages [254]. The results from these different mental workload measurement methods are often dissociated [370] as the sensitivity of these measures depends heavily on the operator's workload [82]. A number of subjective measures, such as the Instantaneous Self-Assessment (ISA) questionnaire [44, 167, 180], NASA Task Load Index [142], and the Subject Workload Assessment Technique [296], are used to assess the operator's workload.

Mental workload is a complex concept that reflects the available cognitive resources and cannot be accurately assessed by using subjective measurements alone [83]. Additionally, the mental workload assessment method should not interfere with the task at hand or influence the mental state of the operator, which may be the case with subjective assessment strategies using questionnaires. Another commonly used workload assessment method is performance-based workload measurement, which, like the subjective assessment method, provides a retrospective workload assessment. However, performance-based measures can only provide a partial understanding as operators can achieve the same performance while experiencing a higher workload [16].

Over the years, physiological metrics have been used to evaluate workload [54, 57, 221] as it provides high sensitivity, diagnostic capability and is mostly non-intrusive [266, 380], giving a precise and real-time assessment of the operator's workload. The use of physiological data such as neurophysiological signals can assess mental workload in real-time without influencing the task as there is no explicit output [125, 263, 264]. Neurophysiological measures can also assess changes in the mental state that are not just evident in overt task performance [38, 264, 355**?**].

Neurophysiological measures such as the electroencephalogram (EEG) signal have been extensively used to estimate mental workload as the effects of task demand are

clearly visible in EEG rhythm variations [41, 45, 89, 125, 204, 221, 288, 290]. Researchers have also used EEG to reliably predict performance degradation from workload variations [127, 219] and noted that it is correlated with an increase in frontal theta power and a change in parietal alpha power, which relates to cognitive and memory performance [119, 124, 126, 274, 330]. Many EEG-based workload indices, such as the ratio of frontal theta to parietal alpha power [118, 154], the ratio of beta to theta and alpha [116], theta-beta ratio [235] reliably reflect workload.

However, EEG features of mental workload are found to be task-dependent; therefore, adding other modalities like eye activity data and heart rate data can lead to better results [173, 277]. Pupil size and blink rate have recently been recognized as reliable indicators of workload [216]. Heart rate variability (HRV) is yet another highly sensitive physiological index to mental workload variations [149, 171, 239, 241, 245]. The root mean square of successive differences (RMSSD) is considered the most robust time-domain HRV measure of workload [226].

Assessing an operator's mental workload can be used to create an adaptive system [279, 316] that adjusts its behavior based on the operator's current level of mental workload [168, 310]. This type of system should be able to respond to changes in the operator's workload without having to explicitly request assistance. When human operators work alongside automation, the operator expects the automation to act like a human coworker [12]. For this reason, adaptive automation should be responsive, stepping in at the right time, and should be able to understand and respond to the operator's needs, taking on tasks that may be overwhelming the operator. In this study, we aim to identify multiple physiological indicators of mental workload in order to create a real-time mental workload detection and alert system.

1.2 Research Aim and Research Objectives

This research thesis aims to discover advanced knowledge of mental workload and develop models and algorithms that enable intuitive mental workload adaptive systems. The first component of this project involves the identification of the neurophysiological behavior traits associated with mental workload variations. This component requires designing real-world tasks that induce controlled levels of mental workload. Workload metrics will be explored in two real-world scenarios: Air Traffic Control and physical Human Robot Collaboration (pHRC).

In this research, we focus on the role of Air Traffic Controllers (ATCs) in maintaining

safe and efficient air traffic flow. ATCs operate in a complex environment and are responsible for organizing the movement of aircraft to ensure they reach their destinations in an organized and timely manner. With the increase in air traffic, understanding the mental factors that contribute to ATC efficiency is becoming more crucial. We examined the mental workload features in the stationary air traffic control environment and used the experience to identify the biomarkers of mental workload in a more complex, dynamic, and physically active physical human robot collaboration. We considered the influence of workload variations, which were manipulated using an arithmetic task in a physical robot collaboration, which is a complex unstructured environment. EEG and physical performance of the participants in these tasks was also evaluated. The physiological effects of workload were systematically investigated in different settings of ATC and pHRC.

Along with the ATC environment, pHRC environment was considered for this project as there is still a major technical challenge to fill this gap in migrating from an interface with a computer (in BCI) to an interface with a robot (in BRI) when the robot is physically moving in unstructured, complex and dynamic environments. It is difficult to record EEG signals in these environments due to movement-induced artifacts when the user moves with the robot. The continuous and indeterminate movement of a robot in BRI also introduces uncertainty in timely detecting and precisely correlating the user's state while the robot is in motion, making it difficult for the robot to learn to calibrate itself to mitigate the user's cognitive workload in real time.

The second component systematically investigated the effects of workload variations on cognitive conflict induced by unexpected robot behavior in different settings of physical human-robot collaboration (pHRC). This component includes having the collaborative robot behave unexpectedly under various workload conditions, which would elicit cognitive conflict of differing intensity in the human operators. Effective communication between a user and a robot is crucial for improving the performance of physical human-robot collaboration (pHRC) in complex and unstructured environments. The ability to quickly and clearly convey a user's intentions to the robot can enhance efficiency and compliance in tasks involving both human and robot movement. In pHRC, the robot should be able to follow the intended motion of the human as they move together.

In physical human-robot collaboration (pHRC), cognitive conflict can enhance the intuitive interaction between humans and robots. This type of conflict arises when the actions of the robot deviate from the user's intended actions. By processing these conflicts, the brain can learn to adapt to automation control and improve the dynamics of human-

robot interaction [201, 306]. However, to the best of our knowledge, no research has been conducted to control the continuous movement of a robot in pHRC based on cognitive conflict. This is due to the difficulty in identifying discrete events from a continuous movement that leads to a brain response that can be used for closed-loop control of the system. As cognitive conflict responses in the human operators can be exploited to correct the robot behavior in the environment, the workload experienced by the operators needs to be maintained at optimal levels. To the best of our knowledge, no studies have explored how workload variations impact cognitive conflict. The physiological responses accompanying workload variations and conflict can be used to improve the collaboration between humans and robots. This brings us to the third and final component of this thesis.

The third component is to develop a safety alert tool based on real-time mental workload assessment. Once the first component was complete, the results of the component were used to explore different strategies to adapt the task conditions to the measured workload of an operator. The data collected as part of the first component can be used to develop a new deep neural network to extract a mental workload biomarker recorded in real-time pHRC by addressing factors within and between user variations, and spatio-temporal nonstationary and nonlinear brain dynamics.

The developed machine learning algorithms can be integrated into a brain robot interface, which can detect the user's workload or error processing capacity in real-time and calibrate the system/robot to meet the user's workload and/or intentions by optimizing the user's workload. Moreover, as a future work, an advanced reinforcement learning algorithm can be developed to predict the occurrence of a conflict and tune the control parameters of the robot in pHRC by treating cognitive conflict as an external reward signal.

This component of the research aims to develop new techniques for measuring and understanding changes in workload during physical human-robot collaboration (pHRC) tasks. By identifying cognitive workload and conflict, it aims to create models and algorithms that can adapt to human-robot interfaces and improve collaboration in complex, unstructured environments. Advanced machine learning and deep learning algorithms for classification are used to create prediction models for workload and conflict. The resulting models and algorithms will be integrated into a brain-robot interface for testing in real-world industrial tasks. The component will also be evaluated through online simulations to test the performance of the safety alert tool

In summary, the aim of this research project are as follows:

RA1: Identify the biomarkers of mental workload variations in a simple ATC task and a pHRC task with strenuous physical activity.
RA2: Discover how the workload variations impact cognitive conflict during a pHRC task
RA3: Design and evaluate the performance of a real-time workload assessment tool for safety in pHRC environments.

RO1: Design tasks that can induce workload in a controlled manner and develop an experiment that captures the physiological changes during the workload variations in ATC and pHRC tasks (RA1)

To accurately measure and induce workload, a controlled task that mimics real-world scenarios is necessary. For this purpose, the basic tasks of Air Traffic Controllers (ATC) were chosen. ATC operators are responsible for anticipating and preventing conflicts between aircraft by ensuring they adhere to International Civil Aviation mandated separation standards and managing the complexity that arises [300].

Although various factors such as environmental, display, traffic, and organizational factors can affect the complexity of ATC tasks [77, 234], the main functions of ATC operators are tracking and collision prediction. Therefore, tracking and collision prediction tasks were designed to investigate the physiological effects of workload variations in these basic ATC tasks. The experiment was designed as a classical cognitive paradigm with manipulation of workload levels (low, medium, high) and repeated stimuli, and physiological data such as EEG, eye activity and HRV were used to assess the mental workload of the operator while performing these basic tasks.

Another task design that was explored will be the pHRC scenario. This task will be in the form of interactive 'Clock Game'. This task emulates the operation of an abrasive blasting robot and builds a repetitive pHRC task. The participant will be asked to move the robot from the starting point to the target that needs to be blasted. The robot will be programmed to behave in an unexpected manner, which will provoke cognitive conflict in the user. Workload variations in the task were manipulated by varying the difficulty in identifying the target in each trial.

This objective will include the preparation for the experiment and the collection of data using the experiment. The preparation steps were designing an experimental protocol, seeking ethics approval, synchronizing all data measuring devices, recruiting participants, and conducting the data collection.

RO2 - Analyse and extract the data from RO1 to identify the biomarkers of

workload variations in ATC and pHRC tasks (RA1)

Once the data had been successfully collected, it was analyzed. As the task design successfully induced workload, prominent workload features were observable. The brain dynamics, eye activity, and HRV features associated with the workload variations in the ATC can be extracted. For pHRC task, which is a more realistic work environment, involving physical interactions with the robot, more strict data preprocessing strategies were employed to extract the biomarkers of workload variations.

RO3 - Analyse and extract the data from RO1 to discover how workload influences cognitive conflict responses in pHRC task (RA2)

As the design of the pHRC task was successful in inducing cognitive conflict conditions under different workload conditions, the influence of workload variations on the cognitive conflict response was easily retrieved. The impact of workload variations on cognitive conflict was also examined.

RO4 - Create a model for mental workload assessment (RA3)

Workload features were extracted from the experiment data to train a deep-learning model. This model was used to predict the workload experienced by an operator performing the tasks. These predictions were able to assess the workload for the pHRC tasks. A workload detection tool was designed based on the findings of this objective. The safety alert tools aim to be a deployable real-time system.

RO5 - Develop a mental workload-based safety alert tool (RA3)

The mental workload-based safety alert tool for ATC will involve intelligent mental workload assessment strategies to mitigate workload effects. The human user will be wearing an EEG device while moving the robot arm for ANBOT. This experiment also incorporates the previous work of RO4 and the designed safety-alert tool, where it evaluated the performance and effectiveness of the safety-alert tool in accurately detecting non-optimal workload conditions. The mental workload-based safety alert tool for pHRC will involve the blasting robot stopping its movement in the environment at safety breaches introduced by a non-optimal workload of the collaborating operator. The performance of this safety-alert tool will also be assessed to study its effectiveness.

1.3 Chapter Organization

The organization of this thesis is illustrated in Figure 1.1 and provides an overview of the key points and findings. It is divided into eight chapters that include an introduction, methodology, results, and potential future research applications.

1. Chapter 1, entitled Introduction, provides an overview of the focus on the mental workload of operators in safety-critical work environments and sets the context for the research goal of detecting mental workload variations in real time.

2. Chapter 2, entitled Literature Review, offers an overview of existing research on mental workload, detection algorithms, cognitive conflict, and the specific research in the field of air traffic control and physical human-robot collaboration.

3. Chapter 3, entitled Methodology, provides an outline of the research design and execution of the experiments in ATC and pHRC, including the background of similar studies, the experimental approach, and the methods used for data analysis.

4. Chapter 4, entitled Biomarkers of Mental Workload Variations in Basic Air Traffic Control Tasks, delves deeper into the research and presents the EEG, eye activity, and HRV biomarkers of workload variations in tracking and collision prediction tasks within ATC. The results of this chapter give an understanding of how effective the experimental design was in inducing variations in workload.

5. Chapter 5, entitled Neural Correlates of Workload in Physical Human Robot Collaboration, presents the findings of our pHRC experiments, including behavioral, performance, and EEG results.

6. Chapter 6, entitled Effects of Mental Workload Variations on Cognitive Conflict during Unexpected Robot Behavior in a Physical Human-Robot Collaboration describes and presents our results that demonstrate how workload variations influence cognitive conflict during a pHRC experiment. In this section, we present the results of the event-related potential (ERP) analysis and explore the implications of these findings for brain-computer interface (BCI) applications in real-world scenarios.

7. Chapter 7, entitled Intelligent Online Workload Assessment and Safety Alert Tool in Physical Human Robot Collaboration, outlines the exploratory prototype of our safety tool for pHRC that considers workload and error awareness of the operator to provide proper safety alerts.

8. Chapter 8, entitled Conclusion and Future work, brings together the key findings of the thesis, evaluates whether the research questions have been answered, and suggests potential areas for future research based on the results.

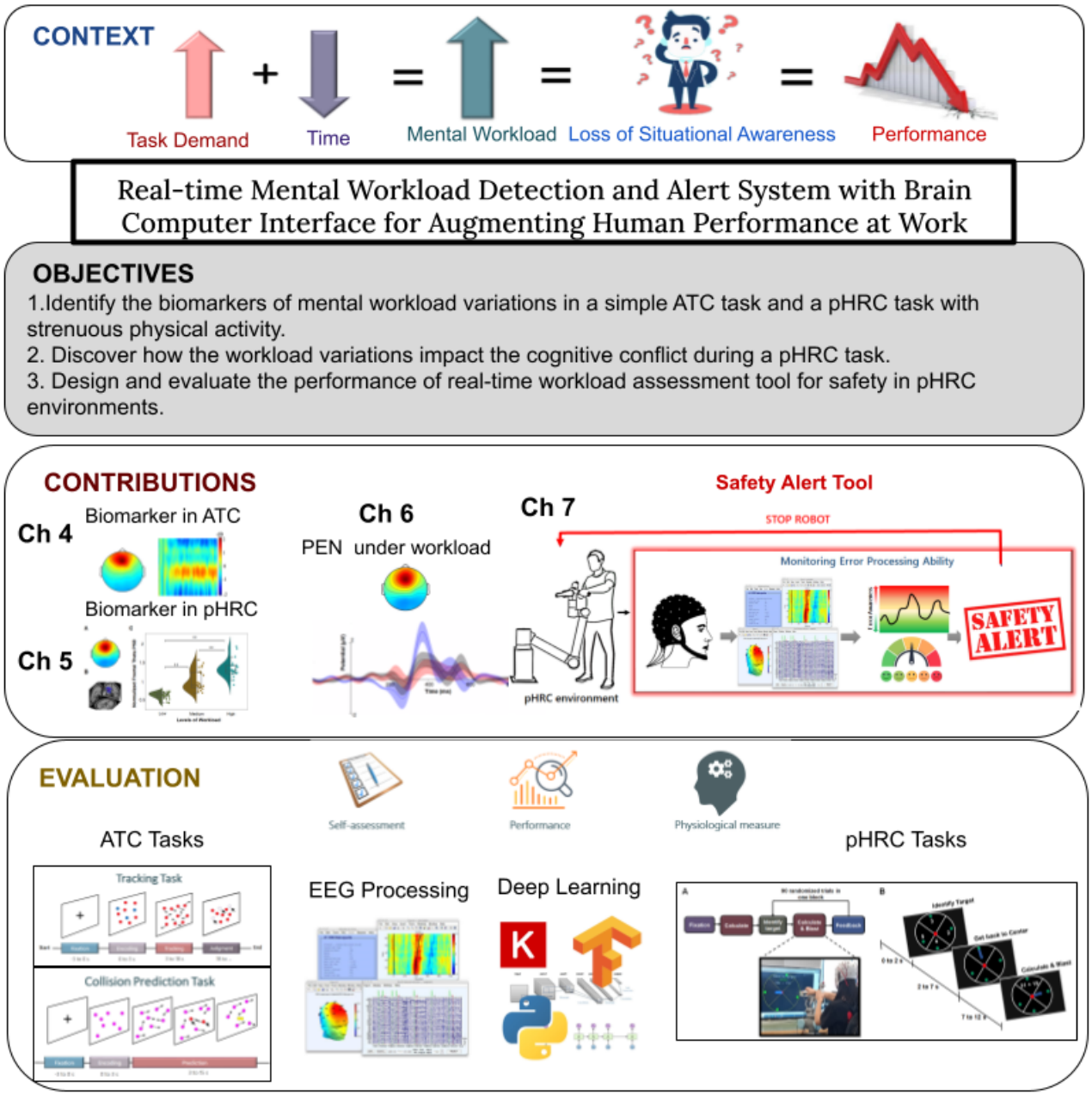

Figure 1.1: Thesis Organization.

CHAPTER 2

LITERATURE REVIEW

2.1 Mental Workload

People find "doing too many things at once" as frustrating and stressful, and they tend to avoid such tasks that push their capabilities beyond their limits [3]. However, all work environments do not offer that leisure. It is vital to have good interaction between work environments and human capabilities[355]. Even though, humans easily adapt to a variety of work environments, and perform a variety of tasks simultaneously, using different equipment, the greater the number of tasks and devices to use, the greater the workload on them. Moreover, digitization, advances in information technology and communications, and interactive work environments require human operators to integrate multiple streams of information in the complex world [269]. This causes an overload of sensory information, resulting in a higher cognitive load [196, 248].

The ability to gauge an individual's cognitive load is of great importance, especially in industries where safety is paramount and human performance can be variable. As the brain's demands increase, it can become more difficult for a person to execute their tasks with precision, resulting in a higher rate of errors. Thus, having dependable methods for assessing cognitive activity and related states, such as mental exertion and situational awareness, is crucial to preserving performance within acceptable standards.

Studies in cognitive psychology have revealed that mental workload has an inverse U-shape correlation with performance. In other words, some amount of mental strain can

aid in achieving a high-performance level, as it keeps the individual alert and focused [49]. On the other hand, a lack of mental stimulation can cause a dull and unengaging state, leading to a decrease in attention and cognitive resources. Furthermore, taking on too many tasks simultaneously can lead to an overload condition, resulting in an increased chance of mistakes [179].

Mental workload is a crucial aspect to take into account when dealing with complex team dynamics, technology and cognitive complexity, and multitasking. It is a transactional concept that describes the interaction between an individual's capabilities and the demands of the task. In simpler terms, mental workload is the amount of an individual's information processing capacity that is needed to meet system demands and represents the proportion of an individual's overall mental capacity used at a given time [10, 342].

Cognitive load, also known as mental workload, is the amount of effort placed on the working memory. It is defined as the cognitive demands of a task in relation to an individual's actual cognitive capacity [96, 169, 254, 355, 368]. There are three types of cognitive load [331]:

1. *Intrinsic load* - intrinsic load is related to the complexity of the task and its association with the operator.
2. *Extraneous load* - extraneous load is induced by how the task is presented.
3. *Germane load* - germane load is related to an individual's ability to absorb information.

As the cognitive load increases, it negatively impacts human trust perception [63], which is particularly vital for human robot interactions. Additionally, mental workload has a strong correlation with variations in individual performance, influencing the likelihood of human errors, the safety of the system, productivity, and the satisfaction of the operator [368]. Moreover, performing multiple tasks interferes with the attention of the human operator [231], which results in performance deterioration [247, 249, 333]. Accurate performance prediction is very important for human operators in the context of safety-critical behavior.

Mental workload is a complex and multi-faceted concept. As shown in Table 2.1, different models, such as the multiple resource model, attempt to explain mental workload as a cognitive process within a resource framework [354, 356, 357, 360]. According to the dynamic adaptive theory, the brain strives for equilibrium and cognitive ease. As a result, very low and very high task demands can harm adaptability and subsequently, performance [137].

Table 2.1: Mental Workload Models

Model	Reasoning	References
Multiple Resource Model	Mental workload is a cognitive process within a resource framework	[354, 356, 357, 360]
Dynamic Adaptive Theory	Brain strives for equilibrium and cognitive ease, which negatively affects the adaptability and performance	[137]
Limited Cognitive Resources	Exposure to challenging and prolonged task demands negatively impacts performance	[43, 261]
Malleable Resources Theory	Attention drifts during non-challenging tasks	[373].

According to the theory of limited cognitive resources, if an individual is exposed to challenging and prolonged task demands, it can negatively impact their performance. This can be due to the following:

- due to resource depletion according to local-sleep theory [261] or
- compromised access to resources [43].

However, it does not explain why even non-demanding tasks such as passive monitoring can lead to moments of mind wandering where attention shifts from the task to unrelated thoughts [92, 220, 222, 322, 323]. The malleable resources theory tries to explain this drift in attention during non-challenging tasks, but it is highly theoretical and hard to put into practice [373].

As the mental workload increases, it leads to a decline in situational awareness which results in poor performance and errors [99, 100, 208, 224, 242, 294]. To maintain attention, it is essential to maintain a high level of situational awareness. Situational awareness is the ability to perceive elements in the environment, understand their significance, and anticipate their future status [99]. There are three levels of situational awareness:

- *Level 1* - level 1 of situational awareness deals with the ability to perceive elements in the environment,
- *Level 2* - level 2 of situational awareness deals with the ability to understand the current situation and

- *Level 3* - level 3 of situational awareness deals with the ability to predict future status.

Situational awareness can be evaluated post-experiment using questionnaires such as SAGAT (Situation Awareness Global Assessment Technique) and SART (Situation Awareness Rating Technique) [101]. Another way to measure situational awareness is by using the SPAM (Situation Present Assessment Method) methodology [74]. However, these methods can only be used to measure situational awareness after the fact and not in real-time.

In recent times, the role of human operators has evolved from a traditional role to a supervisory role using various interconnected tools [299]. This change in the role requires operators to create and maintain a precise situational awareness model from a large amount of information. Additionally, mental workload and situational awareness are closely linked, with one influencing the other. Both mental workload and situational awareness are influenced by similar human factors such as limited working memory, individual variations, and system elements like the difficulty level of the task. In order to prevent a loss of situational awareness, it is crucial to anticipate mental workload and adjust system behavior by adjusting task allocation to maintain high operator performance.

Mental workload plays a crucial role in an operator's performance as it is closely tied to the limitations of the human information processing system [138, 221, 341]. The human brain's ability to process information is finite and has a limited capacity [192], leading to information overload and less efficient performance [10, 357].

Performance is the ability to successfully complete a task [334]. Research has shown that performance decreases at both low and high levels of workload [48, 138, 341, 371]. High task loads can cause performance degradation due to high mental workload and inadequate attention management [236]. Performance reduction in non-optimal workload situations is often preceded by negative neurocognitive states, such as mind-wandering, disengaging effort, repetition, and failure to pay attention.

Mental workload is a complex concept that is influenced by the user, task, and environment [138, 221, 342, 372]. It is crucial to reliably measure instantaneous workload over time, which can facilitate the creation of better work environments and better human-machine interactions [12, 46, 309]. Having a reliable estimate of mental workload is important to maintain workload within the permissible level, avoiding both underload and overload conditions [42, 137].

The mental workload can be measured by:

- Self-assessment,
- Performance or
- using the neurophysiology or psychophysiology.

Each method has advantages and disadvantages [254, 362], as shown in Table 2.2. The sensitivity of each of these methods varies based on the operator's workload [82], and these measurement methods are often dissociated [370].

Subjective measures of workload are generally questionnaires, and the assessment cannot be continuous or real-time. An instantaneous self-assessment (ISA) questionnaire can be used to determine subjective ratings of workload. This questionnaire was developed to assess the mental workload of air traffic controllers, and they can select whether they feel they were under-utilized, relaxed, comfortable, high, or excessive [167, 180]. The subjective workload can also be measured using the NASA Task Load Index [142]. Another subjective workload assessment method is the Subject Workload Assessment Technique [296]. Another self-reports is the Instantaneous Self-Assessment (ISA) technique [181]. All of these measures interrupt the current task and may be affected by the researcher's implicit objectives, and post-justification [34]. For example, they require the operator to focus on filling out the questionnaire regularly, which is invasive and not suitable for real-time evaluation. Furthermore, these measures have a subjective nature and are dependent on the operator which can lead to bias.

Mental workload reflects the cognitive resources that are available, including attentional resources and working memory capacity during the task. This cannot be captured using simple subjective measurement strategies alone [83]. Additionally, the method used to measure mental workload should not interfere with the task or change the operator's mental state. Subjective assessments using questionnaires may be problematic in this regard. Performance-based measures that evaluate primary and secondary performance can be used to assess workload [334], but they can only be used to measure workload retrospectively after an error has occurred. Furthermore, performance measures only provide part of the picture, as an operator can achieve the same performance while experiencing different levels of workload [16].

The mental workload can be evaluated using physiological measurements as well an increase in mental demands leads to increased physical response [152]. The physiological measure is found to be effective in measuring physiological signals [54, 57, 221], and brain, cardiac activity, respiratory activity, and eye activity are some of the physiological measurements of workload [82]. Mental workload measurement using physiological

Table 2.2: Mental Workload Assessment Methods

Assessment Methods	Advantages	Disadvantages
Self-assessment [142, 167, 180, 181, 296]	Ease of Use	• Not continuous or real-time, • Interrupt the task, • Maybe influenced by researcher's implicit objectives [34] • Could be biased
Performance [334]	Easy to implement	• Only retrospective measure, • Provides only part of the picture [16].
Physiological Measures [54, 57, 82, 152, 221]	• Improved Sensitivity and Diagnostic Ability [13, 38], • Non-intrusive [266, 380] • No impact on the task • can be measured on-line [124, 263]	

signals is better than performance or subjective measurement because of the improved sensitivity, diagnostic ability, and non-intrusive characteristics [266, 380].

Also, researchers have found that neurophysiological measures give a better measure of the mental state than performance-based measures [13, 38]. Furthermore, neurophysiological measures have no impact on the task performed by the user as there is no explicit input. It can also be measured online while the task is being performed [124, 263]. Neural measures reflect the decline in performance, which is a transition from a normal operational state to an impaired one, caused by the deactivation of the prefrontal cortex. The prefrontal cortex is a source of limited resources in the brain [233, 278, 292]. It is responsible for controlling function during routine cognitive operations such as decision-making and retrieving information from working memory [291, 298]. The prefrontal cortex becomes active during high levels of cognitive demand [21, 103, 122, 145, 286], and this dysfunction degrades performance [90, 307].

Recently, there has been an increased focus in neuroscience on understanding the mental states of users in operational settings through the use of neurophysiological signals. This research employs a variety of neuroimaging technologies and neurophysiological measures, including functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG), as well as other types of physiological signals such as heart rate, eye movement, etc [41, 291, 367]. Techniques like fMRI are particularly useful in examining how the brain adapts and responds to repeated exposure and practice of specific tasks.

However, the high costs, space requirements, and invasiveness of these methods make them less suitable for real-world working environments where a more practical and less invasive approach would be preferred. Techniques like fMRI require expensive equipment and high maintenance costs, in addition to their large size, which makes them difficult to use in real-world settings [47].

On the other hand, measurements of eye activity, heart rate, and skin response have been found to be correlated with certain mental states such as stress, mental fatigue, and drowsiness, but their effectiveness has been shown to be limited when used alone. They are more effective when used in combination with other neuroimaging techniques that directly measure activity in the brain [39, 41, 304]. Furthermore, research has shown that EEG recordings provide the most informative data for evaluating mental states compared to other peripheral physiological signals [152]. In light of these findings, EEG and fNIRs are considered to be the most viable and straightforward options for investigating brain activity in operational environments.

EEG and fNIRS are popular neuroimaging techniques used in evaluating mental states in operational environments. However, they have distinct physical characteristics, such as interface, size, weight, and power consumption, which can impact their suitability for use in real-world operational environments. Additionally, the presence of hair can negatively impact the accuracy of fNIRs on hairy brain regions [240]. EEG, on the other hand, is less affected by hair and has a higher temporal resolution, making it a more suitable option for use in operational environments."

There is a growing interest in academia and the applied sector in developing neural measures to index workload levels, as neural measures are well-suited to infer cognitive states and for driving brain-computer interfaces. Furthermore, neural measures can reveal changes in cognitive states that aren't obvious in task performance, and they can gather data in real-time without disrupting the operational environment [264, 355].

The mental effort required for a task is determined by a combination of factors, including both a person's inherent characteristics, like their skill level and intelligence, and their current state, such as their emotions, motivation, and fatigue level, as well as the demands of the task itself. Therefore, workload responses to the same task can vary among individuals [150]. Ultimately, everyone responds to workload differently, so we do not analyze workload itself, but we measure the effort put into the task or the subjective reaction to workload condition.

2.2 Mental Workload Classification

Researchers have explored developing mental workload classifiers from physiological data [23, 66, 221, 363, 380]. Artificial neural networks (ANNs) are very popular classification algorithms used for classifying EEG data. ANNs trained on within-difficulty manipulation types provide 85.8% classification accuracy [364]. ANNs have shown they can classify an operator's workload on multi-task combination using heart rate, respiration, eye movement, and EEG data with 98.5% average accuracy [365].

Artificial neural networks are most suitable if multiple physiological signals are used in the workload classification [335, 365]. Jimenez-Molina and colleagues [165] combined multiple physiological signals (electrodermal activity, EEG, and photoplethysmography (PPG)) in artificial neural networks and reported 93.7% accuracy in predicting workload. Wilson and Russell, [364] employed artificial neural network and stepwise linear discriminant analysis to monitor the mental state of air traffic controllers performing ATC tasks using EEG, heart rate, blink rate, and respiratory rate.

Classifiers based on the hierarchical Bayes model that handles multiple subjects have comparable classification accuracy as the subject-specific ANNs [349]. Support vector machine techniques have also been employed to classify the workload of a large number of aircraft pilots using the model that was trained on the change in theta and alpha frequency ratios w.r.t baseline condition. The predictions of the classifier were found to correlate with the pilots' report on the difficulty of the task [42].

Both Random Forest and k-Nearest Neighbor classifiers have been found to be effective and efficient options for classifying mental workload, both in offline and online settings [202]. A study also reported that a support vector machine was effective in identifying different levels of workload and did not require additional training for new subjects or tasks [287, 289]. Additionally, the research utilized personalized head maps that incorporated both spectral features and their spatial locations.

A method called Stepwise Linear Discriminant Analysis (SWLDA) was created that combines the use of forward and backward stepwise analyses. It uses least-squares regression to assign weight to each input feature in order to predict the target class labels [8, 73]. An advanced version of this method, named Automatic Stop Stepwise Linear Discriminant Analysis (assSWLDA), was developed to identify the most important features for different levels of mental workload. The classifier then assigns weights and biases to each feature and the algorithm automatically stops when the ideal number of features have been selected based on the p-value of the model [11, 13]. Studies have shown that this method can effectively classify mental states in drivers and pilots with a high accuracy rate of up to 90% using EEG, EOG, and heart rate data [35, 78].

However, classifiers are specific to each individual and session, meaning that they need to be trained for each person and session separately. It is unlikely that a workload detector with fixed parameters that can be applied to everyone will be developed in the near future [185]. The performance of a classifier is typically measured by its classification accuracy, precision, and recall [110]. Abbass and colleagues [1] stated, based on their work on mental workload classification of air traffic controllers, that it was better to focus on accurately distinguishing two levels of mental workload: low and high.

Moreover, researchers have reported that EEG-based workload estimation is not always very reliable [11]. Penaranda and Baldwin [272] report that the time interval between training and the classifier's test is one important determining factor of classification accuracy. Christensen and Estepp [65] reported that the classification accuracy decreases over days, probably because of overfitting and the high specificity of training

data. They proposed that a classifier that utilizes only a few spectral properties would likely be more selective in its classification.

Supervised machine learning models need labeled datasets of EEG and eye-tracking data for different workload conditions. Features such as alpha power, theta power, pupil diameter, heart rate, and the workload class label can be used to classify the low, medium, and high workload levels [23, 173, 185, 272, 365].

Moreover, selecting appropriate features for classifier training is very important, and it is crucial to limit the number of features to avoid overfitting and improve the stability of the classifier [11]. Further, the cognitive strategies used for performing the task can influence the classification results, intra- and inter-individually [223, 282]. A possible strategy to work around this problem is to cluster the subjects based on their performance or age, or even their experience.

Researchers have been utilizing various techniques to study EEG data. These include analyzing the data in terms of the time it was recorded, the frequency of the signal, and a combination of both time and frequency [162]. One example of this is looking at the amplitude and timing of event-related potentials in the time domain; however, this is not an effective method when the recording does not have a synchronized stimulus [37]. Other methods, like the Largest Lyapunov Exponent [102], and Hjorth parameters [149], allow researchers to investigate the statistical properties of the EEG signal. Additionally, Autoregressive coefficients [9] are also utilized, but interpreting the results in terms of the individual's physiology can be difficult because the results tend to vary from person to person.

Analyzing EEG data in the frequency domain is a common technique, as oscillations play a key role in the synchronization of neural activity in the brain [162]. Researchers often use techniques such as band power analysis, which involves using fast Fourier transforms and short-time Fourier transforms, to study these oscillations. Another well-established method is the spectral power density, which is calculated using Welch's method. This method provides insight into how the power of the signal is distributed across different frequency bands [24, 162]. Gasser and colleagues [121] suggest applying logarithmic transformations to spectral power to improve classification accuracies. Time-frequency methods, such as Hilbert-Huang Spectrum and Discrete Wavelet Transform, are suitable for analyzing short-duration data [68].

Accurate detection of an operator's workload level can drive the implementation of an adaptive automation system. This system adjusts its behavior according to the operator's physical and cognitive state. Some potential applications include:

1. Simplifying the interface to reduce visual distractions

2. Changing the interaction modalities to enable hands-free operation or to decrease the reliance on visual input

3. Modifying the data processing logic or decision-making process, such as using a less conservative detection logic to reduce the number of irrelevant alerts displayed.

Adaptive automation systems have the capability to take over control of the system in critical scenarios such as an operator's incapacitation and bring the system to a stable state. The idea of adaptive automation has been around for some time, but past attempts mainly focused on linking the information and functions provided on a human-machine interface to the situation's characteristics. With the recent advancements in neurophysiology, the systems that adapt based on the operator's cognitive state have become a significant area of research. In recent years, various BCI systems have been studied that use neurophysiological measures to trigger automation changes and have been analyzed for their impact on operator performance [81]. Studies have revealed that people not only view these systems as helpful assistants but even expect them to behave like humans.

While adaptive automation can offer many advantages, it also poses some challenges. The dynamic changes in the system's behavior can make it more complicated and harder to anticipate for the user. Therefore, it is important to try to reduce instances where the user is perplexed or caught off guard by the system's actions. The concept of adaptive automation may require adjustments to increase acceptance among users. An alternative solution, such as adaptable automation, where the user can decide when to enable advanced automation, could be considered. This would give users a sense of control, prevent confusion and allow them to regain control if the automation is unable to handle the situation.

Different strategies have been proposed for triggering automation [311]. Generally, the literature describes three main approaches:

- Critical-event strategy, it assumes that human workload may increase during critical events, and it triggers the automation accordingly [147];

- Performance-measurement strategy is used to gauge how well an operator is performing during a task. This helps to determine their current and predicted state and to see if they are feeling overwhelmed;

- Neurophysiological measurement strategy is employed to record the operator's physiological signals in order to understand their mental workload [13, 309].

In this field of study, using neural measures can help to address two crucial concerns as detailed in [12]:

1. The use of neural measures can help establish specific boundaries for when automation levels should shift, either up or down. These boundaries can be either clear-cut or more fine-tuned over a longer period, to avoid frequent adjustments and consider the cumulative effects. For instance, a moderate to high workload that continues for several minutes could have the same impact as a very high workload peak. Additionally, after a prolonged period of working with high automation, it might be beneficial to keep the same level of automation for a while, even if the workload is low, in order to aid in recovery.

2. Using neurometrics as a tool can offer scientific proof of the advantages of adaptive automation, for instance by presenting evidence of a decline in workload after the introduction of adaptive automation.

A Brain-Computer Interface (BCI) is a system that captures brain activity and converts it into a form that enhances the natural functioning of the brain. This improves the interaction between the brain and its external or internal environment [366]. A passive BCI, in particular, recognizes the mental state of the user and utilizes it to adapt the interaction between the operator and the system. They are well suited for adaptive automation [265, 375], but there have been only a few studies that have explored the use of passive BCI for triggering adaptive automation solutions, all of which have been conducted in a laboratory setting.

EEG is particularly beneficial in creating passive brain-computer interfaces [13]. Scientists have developed a closed-loop system that uses the operator's own EEG signals to decrease their workload [279, 316]. Comstock and Arnegard [71] were one of the first to create an EEG-based adaptive automation system, which alternates between automatic and manual modes for tracking tasks in a Multiple-attribute Task Battery.

Prinzel and colleagues [279] show a significant decrease in mental workload after activating automation, which was confirmed by EEG and subjective ratings. Berka and colleagues [32] used the Ageies simulator, a military simulation environment, to demonstrate the usefulness of EEG-based workload indexes in reallocating tasks and system aids in real time. Abbass and colleagues [1] used the controller,Äôs working

position adaptive to the EEG signals and task complexity. It was observed that operators achieved better performance with adaptive automation, which used the theta-to-beta ratio over the entire scalp as the feature. However, the study only had four subjects, which is not statistically significant. Arico and colleagues [13] used a passive brain-computer interface system integrated with an ATM simulator that triggers adaptive automation in real time based on workload, avoiding both underload and overload.

Mental workload adaptive systems have challenges that limit their successful implementation [173]. It could be because individual differences in cognitive abilities affect performance [93, 94, 130, 224]. Mental workload estimated, especially from EEG, is task-specific and the classifier trained on a task, may fail for a similar task [173, 277].

2.3 Cognitive Conflict and Prediction Error Negativity

Cognition is the process of acquiring information through senses, experience, and thought processes. It involves several cognitive functions, such as learning, thinking, memorizing, and reasoning [284]. Researchers have extensively explored most of the components of these intellectual functions, including the ability to predict a future state on the basis of experiences in the past.

Human beings continuously predict future outcomes and consequences while performing any task [303] based on an internal cognitive model. This internal cognitive model works by forming a simulation of the response to a task that will predict its outcome [115]. Whenever a mismatch occurs between reality and this internal cognitive model, which is inevitable, a phenomenon called cognitive conflict happens [140].

Cognitive conflict occurs from the mismatch between reality and the internal cognitive model or perception. These internal models are continuously used by the human neuromotor control system to estimate the outcomes or actions; this phenomenon of cognitive conflicts occurs when the expectation is unmet or when these internal models are found to be wrong. Therefore, cognitive conflict response can be an indirect way to measure the intuitiveness of any system.

Cognitive conflict is a very fundamental cognitive response in real-world interactions. Detecting and measuring this phenomenon can improve the interactions human beings have in any work environment, especially in an environment that involves interactions and collaborations with robots. This cognitive conflict phenomenon can be exploited to guide a robot in a task environment.

EEG, or electroencephalogram, is a method of measuring the electrical activity of the brain through sensors placed on the scalp of the user. It is considered more advantageous than other neurophysiological measurement techniques as it is relatively inexpensive, offers high temporal resolution, is completely non-invasive, and can be used in real-world settings where the user has freedom of movement. As a result, EEG measurements have been widely utilized by researchers in cognitive neuroscience [117] and particularly in creating brain-robot interface systems [214].

The idea behind Error-related potential (ErrP) first came to light in the 1990s during a study where individuals were observed making errors in a task that required quick decision-making [106, 123, 252]. Essentially, ErrP is a way to compare the brain's electrical activity during correct actions and errors, it's essentially the contrast between an error-related negativity (ERN) signal and a correct-related negativity (CRN) signal [107]. The ERN is a form of brain activity that can be recorded through EEG and it is activated when a person makes or detects a mistake while participating in an experimental task [107].

When an error is made, the brain's electrical activity in the frontocentral area, known as the ERN signal, dips negatively at around 50-200 milliseconds after the error occurs. This is followed by a positive shift in the central area of the brain at around 200-500 milliseconds later [107, 136, 343, 348]. On the other hand, when a correct response is made, a similar waveform to the ERN is observed, it's called the correct-related negativity (CRN), but it is of a lower magnitude [136, 348].

Studies on the Error-related Potentials (ErrP) indicate that this signal is only activated in certain task scenarios and the different types of error-related potential are shown in Table 2.3. For instance, a Response ErrP is likely to happen when the person is asked to respond as fast as possible [257, 343, 348]. A Feedback ErrP happens when a person becomes aware of an error upon receiving feedback on a task [205, 369]. An Interaction ErrP occurs when a person is interacting with a machine and the machine misunderstands an instruction [36, 112]. An Observation ErrP happens when a person notices an error made by a machine or external system [301, 306]. Recently, three new types of errors: Target, Outcome, and Execution ErrPs have been reported in the literature [230, 327].

Over the years, researchers have identified the negative response in ERP in different task settings, and it has been referred to different terminologies depending on the task conditions. For instance, Stagg and colleagues [328] identified the cognitive conflict experienced during a visual effect monitoring task and termed it as visual mismatch

Table 2.3: Types of Error Related Potential

Types	Reference
Response ErrP	Happens when the person is asked to respond as fast as possible [257, 343, 348]
Feedback ErrP	Happens when a person becomes aware of an error upon receiving feedback on a task [205, 369]
Interaction ErrP	Happens when a person is interacting with a machine and the machine misunderstands an instruction [36, 112]
Observation ErrP	Happens when a person notices an error made by a machine or external system [301, 306]
Target ErrPs	[230, 327]
Outcome ErrPs	[230, 327]
Execution ErrPs	[230, 327]

negativity (vMMN). Kopp and colleagues [188] found another feature of ERP called the N2 or N200, which is elicited by visual stimuli. In an experiment on price and reward prediction, feedback-related negativity was observed.

Recently, in an experiment that involved active motor movement, Singh and colleagues [320] observed prediction error negativity (PEN). The Prediction Error Negativity is caused by an error in the task being performed, it is different from Event-related Negativity as it occurs during the task rather than after it has been completed or while it is being observed. Prediction error negativity is the negative deflection in the brain wave between 100-300ms that is associated with a mismatch between a predicted and an actual system state. Additionally, EEG measurements can be used to detect the prediction error negativity, which is similar to cognitive conflict, by measuring event-related potentials [207, 252].

Identifying and understanding cognitive conflict can lead to better interactions in various settings, such as when a robot works alongside a human to complete a task. In such an environment, cognitive conflict will be invoked whenever the robot functions unexpectedly. The cognitive conflict during human robot interactions has been explored by researchers. In a study, the human operators were required to observe the behavior and actions of the robots and judge them as correct or wrong [159]. They observed negativity in the ERP around 400 ms after the robot performed a wrong activity. They

used this data to minimize cognitive conflict by training the robot based on reinforcement learning.

A BCI is a method of communication between a person and a device by interpreting the user's intention through the brain signals recorded by EEG [199, 206]. When the system misinterprets the user's intent it leads to an error, which triggers an Error-related potential (ErrP) [170, 378]. This ErrP signal can be combined with the standard BCI system to create a hybrid BCI system that initiates steps to correct the error, and avoid it from occurring again, hence enhancing the overall effectiveness of the BCI system [62, 75, 275].

Using ErrP as a feedback mechanism in an adaptive BCI can help the system to improve by learning from its mistakes, which will decrease the chances of interpreting the user's intent incorrectly [17, 61, 97, 158]. Studies have shown that ErrP signals are an innate human feedback mechanism that can be generated in the brain without any additional training after an error is made [177]. The non-invasive nature of ErrP makes it easy to combine with other brain-computer interfaces (BCIs) where users intentionally control their brain activity to control the system dynamics [376]. Studies have shown that ErrP signals can be accurately captured, making it feasible for use in real-time BCI systems [62, 111, 267, 315].

Penaloza and colleagues [271] exploited the error-related negativity information from the human operator who was observing the robot navigating in the task environment to activate the safety measures in the robot. Ehrlich and Cheng, [97] observed ErrP (Error-related potentials) in the event-related potential of the human users who observed robots operate in the task environment while randomly performing some erroneous actions. Salazar-Gomez and colleagues [306] used the ErrP resulting as a repose to the robot's erroneous actions to correct the robot's activity in the task environment. Even cognitive conflict phenomenon has been explored in human robot interactions to control and correct the robot behavior, no researchers have explored this cognitive conflict behavior in real-world settings but only with human operators observing the robot activity passively.

The cognitive conflict phenomenon of real-world physical interactions with robots has not been explored because of the noise and artifacts in EEG from physical movement and other confounding factors [281]. Singh and colleagues [320] studied the cognitive conflict phenomenon in a task involving active motor movement. The study identified a unique indicator of cognitive conflict, labeled PEN, in the ERP data by creating a task in which participants utilized hand movements to select 3D objects within a virtual reality

setting.

Our lab has conducted some preliminary studies in physical human-robot collaboration (pHRC), a collaboration involving human operators and semi-autonomous robots that are actively physically guided by the human operator [5]. This study found evidence of cognitive conflict in physical human robot interactions.

Cognitive conflicts are accompanied by a physiological response and activate different brain regions in humans. Many studies have noted a significant amount of activity in the dorsal anterior cingulate and prefrontal areas, which are responsible for monitoring and resolving conflicts respectively, as reported in [109]. Physiological measurements such as electroencephalography (EEG) reflect certain cognitive conflict states and can be used for BCI in controlled environments [97].

Traditional BCI approaches to robot control, such as P300 potentials [253], steady state visually evoked potential [201], motor imagery [164] and need constant human attention in operation and usually require training for users to learn to modulate thoughts appropriately. In contrast, the neural features accompanying cognitive conflict are generated automatically and do not require prior training. Thus, brain dynamics associated with conflict processing can foster an intuitive interaction between humans and robots in pHRC [201, 306].

However, there is still a major technical challenge to fill this gap in migrating from an interface with a computer (in BCI) to an interface with a robot (in BRI) when the robot is physically moving in unstructured, complex and dynamic environments. It is difficult to record EEG signals in these environments due to movement-induced artifacts when the user moves with the robot. The continuous and indeterminate movement of a robot in BRI also introduces uncertainty in timely detecting and precisely correlating the user's state while the robot is in motion, making it difficult for the robot to learn to calibrate itself to mitigate the user's cognitive conflict in real-time.

2.4 Air Traffic Controllers

Making sure that air travel is as safe as possible is a top priority in the industry. This goal can be achieved not only through updating and improving technical systems but also by studying the human factors that play a role in any incident where safety is compromised. A comprehensive understanding of all the factors involved is necessary to maintain the highest level of safety.

Human factors, which refer to the various characteristics and abilities of people,

can greatly impact performance, efficiency, and reliability. These factors, which are determined by psychological and physiological conditions, are believed to play a role in as many as 80% of aviation accidents [256, 318]. This is not just limited to pilots, but also air traffic controllers, who may be affected by a number of factors that can impact their focus, performance, and the potential for errors that can lead to safety hazards [56].

Air traffic control is a service that ensures the safety and efficiency of aircraft by preventing conflicts between them and managing traffic flow through the organization and provision of information and support to pilots. Air traffic controllers are responsible for one particular sector, where they accept aircraft, give instructions and clearances, and hand over the aircraft to the next sector or the tower sector. They regularly handle a large number of aircraft coming from different directions, heading to different destinations, and at various speeds and altitudes [131]. The primary goal of air traffic controllers is to regulate air traffic and provide information and support to pilots from control towers, to maintain a safe and orderly flow of air traffic, and avoid conflicts among aircraft [203].

The main objective of the air traffic control system is to guarantee the safety and smooth operation of air traffic. To accomplish this, international civil aviation organizations establish standards for separation between aircraft that must be followed. The air traffic controllers are in charge of making sure that aircraft comply with these standards, including maintaining a vertical distance of at least 1000 feet and a horizontal distance of 5 nautical miles from other aircraft. They also need to keep an eye on air traffic and detect any possible conflicts between aircraft, take action to resolve them, and minimize any disruption to the flow of air traffic, in order to allow aircraft to reach their destinations in an orderly and timely manner.

Air traffic control is a complex task that can be affected by a variety of factors, including environmental, organizational, traffic, and display factors [77, 234], stress, fatigue, and overall health and mental condition [283]. To address these challenges, researchers have conducted studies to understand human behavior and responses to stress and to develop strategies to eliminate potential safety risks [50].

As the number of airplanes in the sky continues to rise, air traffic control is facing more pressure. According to the European Agency for Safety and Health at Work, heavy workloads and fatigue are major factors in transportation accidents [72]. Air traffic controllers have three main responsibilities: keeping track of all the planes in the air, predicting and preventing problems, and resolving conflicts [203, 336]. These duties can be demanding, especially during times of high traffic or when dealing with flights that need special attention [148].

Air traffic controllers are constantly facing high pressure, even during routine operations. Their job of managing aircraft, such as keeping them separated, coordinating their movements, and resolving conflicts, requires a lot of mental focus and attention. They play a crucial role in dealing with the complexity of the air traffic system, as their main goal is to anticipate and handle any unexpected situations that may arise. The air traffic control system is not only complex in its environment but also in its communication systems such as radio, phone, radar displays, and computers. All these factors combined, make the controller's job even more challenging.

Air traffic control work often involves performing a variety of tasks in addition to routine operations, such as communicating with aircraft upon arrival and passing them off to other control sectors. These tasks may be familiar, but their order and timing can be unpredictable due to factors like weather, traffic demand, and operational conditions. This unpredictability, in addition to the high level of responsibility and complexity of their job, makes it necessary to systematically evaluate workload levels to ensure safety. To get an accurate understanding of the workload involved in air traffic control, it's important to individually examine each of the subtasks that controllers perform.

The measurement of workload in air traffic controllers has gained significant attention in the air traffic management field due to its importance in human resource planning, operational efficiency, and, most importantly, the safety of the air traffic system. Recently, traditional air traffic management has faced several challenging issues, such as a substantial increase in the number of flights, more complex airspace, sensitivity to weather changes, and the interaction of these factors.

EEG-based studies have been conducted to assess mental workload in ATC work environments and the research is summarized in Table 2.4. Research has shown that it is possible to understand the level of mental effort being exerted by air traffic controllers (ATCOs) by analyzing their brain activity using electroencephalography (EEG) signals. In recent years, EEG has been recognized as the most accurate neurophysiological method for measuring mental workload. For example, Brookings et al. [45] conducted a study that looked at the changes in mental workload among eight military ATCOs as the difficulty level of Terminal Radar Approach Control (TRACON) tasks in a simulated air traffic management system varied. The research conducted a thorough evaluation of controller workload by utilizing multiple metrics, including performance, subjective, and physiological measurements. The study revealed that the level of difficulty had a significant impact on the theta band power in certain regions of the brain, as the power increased with the increase in task difficulty. It also found that the alpha band

Table 2.4: EEG-based Mental Workload Assessment in ATC

Research Team	Contribution
Brookings et al. [45]	Increased task difficulty resulted in increased theta power and decreased alpha activity.
Shou et al. [319]	Frontal theta EEG was identified as accurate indicator of workload
Arico and colleagues [13, 14]	Developed a brain Workload Index to effectively distinguish different levels of difficulty in ATC tasks.
Arico and colleagues - passive BCI in ATC [14]	Revealed a strong correlation between the EEG-based mental workload index and the subjective mental demands, as measured by ISA.
Arico and colleagues - adaptive automation in ATC [13]	EEG-based Adaptive Automation in an ATM research simulator, which effectively turns on the adaptive automation during high-demand conditions, ultimately reducing the mental workload for the ATCOs.

activity decreased as a result of the interplay between the level of difficulty and traffic manipulation. These conclusions align with those of other studies that have employed similar techniques to gauge ATCO mental workload.

A team of researchers led by Shou et al. [319] assess the level of mental fatigue and workload during an air traffic control task. The study found that measuring the activity of the frontal theta EEG was an accurate and dependable way to evaluate workload and the effect of time spent on the task, on a minute-by-minute basis. Furthermore, the study also pointed out that this method is particularly useful in identifying neural activity from ongoing EEG recordings in real-world tasks.

Recently, researchers have found that by using machine learning techniques and analyzing brain activity data from EEG, they can calculate a brain Workload Index [13, 14]. This index has been proven to effectively distinguish different levels of difficulty in air traffic management tasks. Furthermore, this technique has also been utilized to assess the effect of different avionic systems on the cognitive load of helicopter pilots [40]. Not only that but it has also been shown to be a reliable measure of mental workload in real-world scenarios where the task at hand is constantly changing [15, 87]. One example of this is a study where professional air traffic control operators were tasked

with performing a realistic air traffic management task.

A similar research project enlisted the participation of 12 professional air traffic controllers. These individuals were presented with a variety of simulated ATM scenarios that ranged in difficulty level [14]. The scenarios were designed to mimic real-life situations and featured a dynamic level of difficulty, starting off easy and gradually increasing before returning to an easier level. The results of the study revealed a strong correlation between the mental workload index, as measured by EEG, and the mental demands experienced by the operators throughout the task, as measured by ISA.

Recently, a team of scientists has created an EEG-based system that turns on Adaptive Automation in an ATM research simulator, using the actual mental workload of the operator as a reference. The system was put to the test by 12 Air Traffic Control Officers as they performed high-realistic ATM scenarios of varying difficulty levels. The study found that the proposed system was effective in activating AA during high-demand conditions, ultimately reducing the mental workload for the ATCOs. Furthermore, the system was designed to not activate AA when the workload level was below a certain threshold, thus preventing the operator from experiencing low-demand conditions that could be dangerous [13]. The brain features related to mental workload, identified in this and previous studies, was found within the frontal theta, parietal alpha, and occipital theta rhythms [14, 15, 87, 88].

2.5 Physical Human Robot Collaboration

The integration of robots into our daily lives is becoming increasingly prevalent as automation gains momentum. The utilization of robotics goes beyond just industrial settings and is now making its way into human-centric environments. For robots to truly become a seamless part of our lives, safety, and dependability must be given priority. Personal robots have the potential to be just as common in households as personal computers are today, but as of now, robots are still perceived as unwieldy and unsafe machines. One of the major hindrances to widespread acceptance is the lack of well-defined safety standards in the scientific community for physical human-robot collaboration.

Ensuring safety for both humans and robots is a crucial aspect when it comes to the coexistence of both in the same physical space, where they share workspace and cooperate physically. Industrial robots are typically equipped with safety measures such as safety guards or locked doors to keep the workspace of the robot separate from

humans [329]. In case these safety measures are tampered with, the robot is shut down immediately to prevent any harm.

As more and more people are looking to use service robots in their homes and workplaces, it's becoming increasingly important to make sure that these robots are safe to interact with humans. This is especially true when physical contact between humans and robots is not only desired but sometimes necessary. To ensure the safety of human users, it's crucial to objectively measure the potential for injury that these robots may cause, and make sure that the robot's actions will not exceed a safe limit if physical contact occurs. This includes both regular operation and in case of any faults that may occur.

As technology advances, more and more systems are being developed to share the same physical space with humans, such as self-driving cars and collaborative robots. The most important aspect of these systems is ensuring that they are safe for people to interact with. This has been a key concern for quite some time, as seen in science fiction literature, where the principle that robots should not cause harm or allow harm to occur to humans through inaction has been emphasized [18]. To ensure safety when interacting with robots, it's important to make sure that the worst-case scenario would only result in mild injuries. Safety in physical human-robot collaboration has been widely studied, and the results of these studies have been summarized in various research papers.

In recent years, there has been an increasing number of scenarios where autonomous systems share the same physical space as humans, such as collaborative robots and self-driving cars. The most crucial element in these systems is making sure they are safe to interact with people. This has been a significant concern for quite some time, as seen in science fiction literature where it is stated that robots should not cause injury to humans or through inaction allow humans to come to harm [18] long before the development of robot actuators [237]. A way to ensure safety when interacting with robots is to make sure that in worst-case scenarios, only mild injuries may occur [135]. The topic of safety in human-robot collaboration has been widely studied, and the results of these studies have been summarized in various research papers [69, 133, 273, 302].

Physical Human-Robot Collaboration (pHRC) is the teaming up of humans and robots with the aim of accomplishing a task. While human operators are highly skilled and possess an intuitive ability to discern elements in the work environment, robots lack this intuitive and empathetic quality. Robots have greater mechanical strength and are completely by mathematical models, and when robots participate in a collaborative task, they are also called cobots [70]. When these humans and robots team up to accomplish a

task, this collaboration should benefit from the strength of both human operators and robots. However, the lack of intuitive communication between the human operator and the robot poses many safety challenges and complexities.

When humans and robots collaborate in a manner that includes physical contact, such an interaction is called physical Human Robot Collaboration (pHRC). The scope of this collaboration can be defined by the collaborative task on which the human operator and robot are working and also by the nature of this collaboration.

In physical Human Robot Collaboration, the human operator and robot are in physical contact with one another either directly or through a rigid body in order to accomplish the task at hand. This might even be an exchange of forces between the human operator and robot, where the operator might be completely or at least partly controlling the robot's activities. This control of robot activities can be done using continuous force exchange, objects that can be manipulated, and haptic devices. Such physical human robot collaboration has been explored with industrial robots for material handling [120], lifting and holding tasks [4], and homokinetic joint assembly [64].

This physical human robot collaboration has also benefited rehabilitation robotics. Researchers have devised hand rehabilitation robots [218], an exoskeleton for shoulder and elbow rehabilitation [53], and a lower-limb exoskeleton [374].

Generally, human operators control robots in these collaborations by exchanging forces with robots which can be transformed into robot motions. Therefore, these controllers are usually force-based, and they can be either impedance-based control or admittance-based control, or a variant of these [297]. Impedance and admittance provide information on how force can be rendered to velocity in a physical human robot collaboration.

Mostly, the robot motion is also governed by many other components like assistive strategies, singularity, and collision avoidance [53, 80, 86, 243, 255]. As assistive strategies and singularity and collision avoidance components play a crucial role in robotic motion, these robot motions can never be explained on the basis of force applied by the operator. Therefore, this results in the human operator experiencing dynamic mechanical resistance in the task environment. The resultant mechanical resistance can be adjusted empirically based on intuitiveness and smoothness.

A direct consequence of the absence of an accurate representation of the relationship between intuitive and mechanical resistance is the difficulty in formulating a rigorous mathematical model. Even though researchers have explored different components of physical human robot collaborations, not many researchers have explored intuitiveness

assessments. And the few studies that explored intuitiveness in a physical human robot collaboration only considered subjective data from questionnaires and surveys [113].

Intuitiveness is the ability to know or understand things without conscious reasoning. It's a cognitive process that relies on previously acquired knowledge and experience. It's often described as a way of understanding things without the need for rational thought, as stated by Bastick [25].

As such, intuition is rooted in cognitive internal models and a human's ability to make predictions. Consequently, the intuitiveness of a human operator collaborating with a robot is directly proportional to the ease with which the human operator can predict or foresee robot behavior in the environment. Intuitiveness is found to increase with human operators' training and experience with collaborating with the robot, making the experience more natural. However, the robot's intuitiveness is always independent of the human operator's level of training or experience.

Villani [347] identified that only an intuitive human robot interaction can ensure a safe physical human robot collaboration, which will also fully optimize the system's performance. Several strategies have been explored to enhance the quality of this human robot interaction with the aim of establishing a natural and intuitive form of communication between the human operator and robot.

A robot can be aware of the condition of the human operator through its sensors. The advancements in biomedical devices, that measure the physiological state of a human operator, have made it possible to have a reliable estimate of the human state comfortably and more cheaply. Many research groups are investigating ways to integrate these physiological measurement devices into the robot to have an accurate estimation of the human state. Galvanic skin responses have been used to measure engagement [238] and comfort [227] in human-robot collaboration. Additionally, heart rate assessment equipment has been widely used to estimate the human state [161, 293, 347].

However, heart rate data can be affected by various psychological and physical factors. To have a more reliable estimation of the human state, it is better to use multiple physiological signals. Kulic and Croft used skin conductance, contractions of the facial muscle and heart rate [194], while Koenig used skin conductance/temperature and heart rate [184]. Researchers have so far employed these physiological signals to estimate the stress experienced by the human operator in human robot collaborative set-up.

However, in a physical human robot collaboration, the intuitiveness of communication is very crucial to ensure a safe and natural interaction between the human operator and robot. Researchers have explored this intuitiveness qualitatively and subjectively. An

objective and quantitative approach to assess the intuition of this communication is by measuring prediction error negativity or PEN based on the brain dynamics of the human operator.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter will provide an overview of the experiment details for this research project, including the materials used and the methodologies employed to gather and analyze data. The methodology for the ATC experiment, the pHRC experiment, and the online safety alert tool for pHRC will be outlined, supporting the findings presented in Chapters 4 to 7. The chapter organization is shown in Figure 3.1.

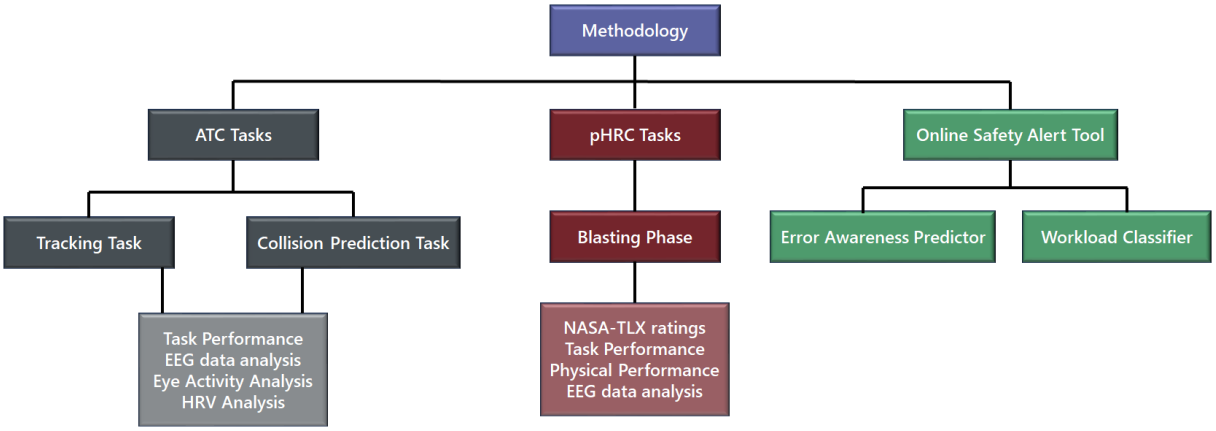

Figure 3.1: Organization of Chapter 3.

3.2 ATC Experiment Methodology

3.2.1 Participants

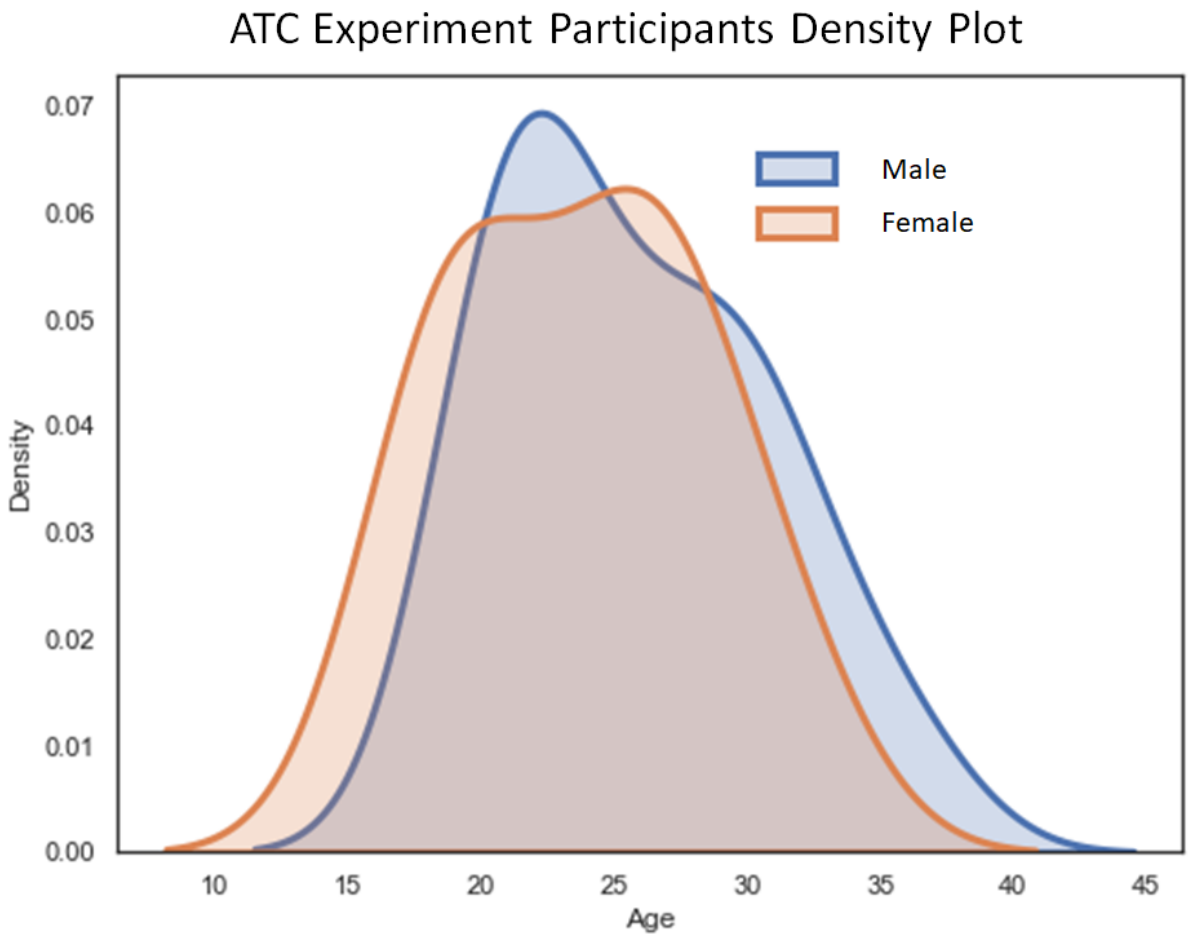

Figure 3.2: ATC Experiment Participant Distribution.

This experiment involved the participation of 24 individuals (average age 25 ± 5, with 17 males and seven females, all of whom were right-handed) who gave written informed consent. The experiment density plot is shown in Figure 3.2. All participants had normal or corrected vision and no prior psychological disorders that may have impacted the results. The experimental protocol was granted approval by the University of Technology Sydney Human Research Ethics Expedited Review Committee (ETH19-4197).

This data collection was conducted during the first COVID lockdown in Sydney, Australia. We took additional care to follow the additional COVID safe protocol based on New South Wales guidelines, which was further approved by the UTS Ethics Committee

and the University Risk Assessment team. There was difficulty in recruiting participants for this study during this time as most of the country was staying locked inside.

We collected EEG data using the SynAmps2 Express system from Compumedics Ltd. in Australia, which had 64 Ag/AgCl sensors. The electrodes were placed according to the extended 10% system [58] and we made sure that the impedance for each electrode was below 10 kΩ before each session. The data was recorded at a rate of 1000 Hz. Additionally, we used Pupil Labs Pupil Core, a wearable eye-tracking headset from Berlin, Germany, that has three cameras to record the participant's eye activity and field of view at 200 Hz and 30 Hz, respectively [172]. We also recorded Blood Volume Pulse (BVP) data using Empatica E4, a device that uses infrared plethysmography from Empatica Srl in Milano, Italy. We were able to synchronize events from the task scenario with the EEG, eye activity, and BVP data in real time using the Lab Streaming Layer [191].

3.2.2 Experimental Procedures

Our study featured two tasks: the multiple objects tracking task [156] and the collision prediction task. As shown in Figure 3.3(A), in the tracking task, participants were first presented with a fixation cross on the screen for three seconds, followed by a freeze phase where the dots, some of which were blue, and the rest were red, remained stationary. The blue dots were the targets that participants were instructed to track. After three seconds of freeze, the blue targets turned red, so that they were no longer distinguishable from the other dots and all the dots began to move randomly. The dots had a diameter of 14 pixels, and they moved at a rate of 15 frames/second.

In the task, participants were asked to keep track of the target dots (initially blue) for 15 seconds. Once the time was up, all the dots stopped moving, and participants were prompted to indicate the target dots by clicking on the dots that they had been keeping track of. The workload level of the tracking task was manipulated by varying the number of blue dots and the total number of dots. As shown in Table 3.1, in the low workload condition, there were 10 dots with one blue dot; in the medium workload condition, there were 12 dots with three blue dots; and in the high workload condition, there were 15 dots with five blue dots.

In the collision prediction task, as shown in Figure 3.3(B), participants were first presented with a fixation cross on the screen for three seconds. This was followed by a three-second-long freeze phase where the dots remained stationary. After that, all the dots started moving in a predefined uniform direction. Unlike the tracking task, all dots were of the same color (pink), and the participant's task was to predict the trajectory

of the dots and identify which pair of dots would collide. The trajectory of the dots was manipulated such that there would be only one collision in each trial. Participants were asked to identify the pair of dots that would collide and click on both dots before the collision happened, which usually occurred in the last three seconds of the trial.

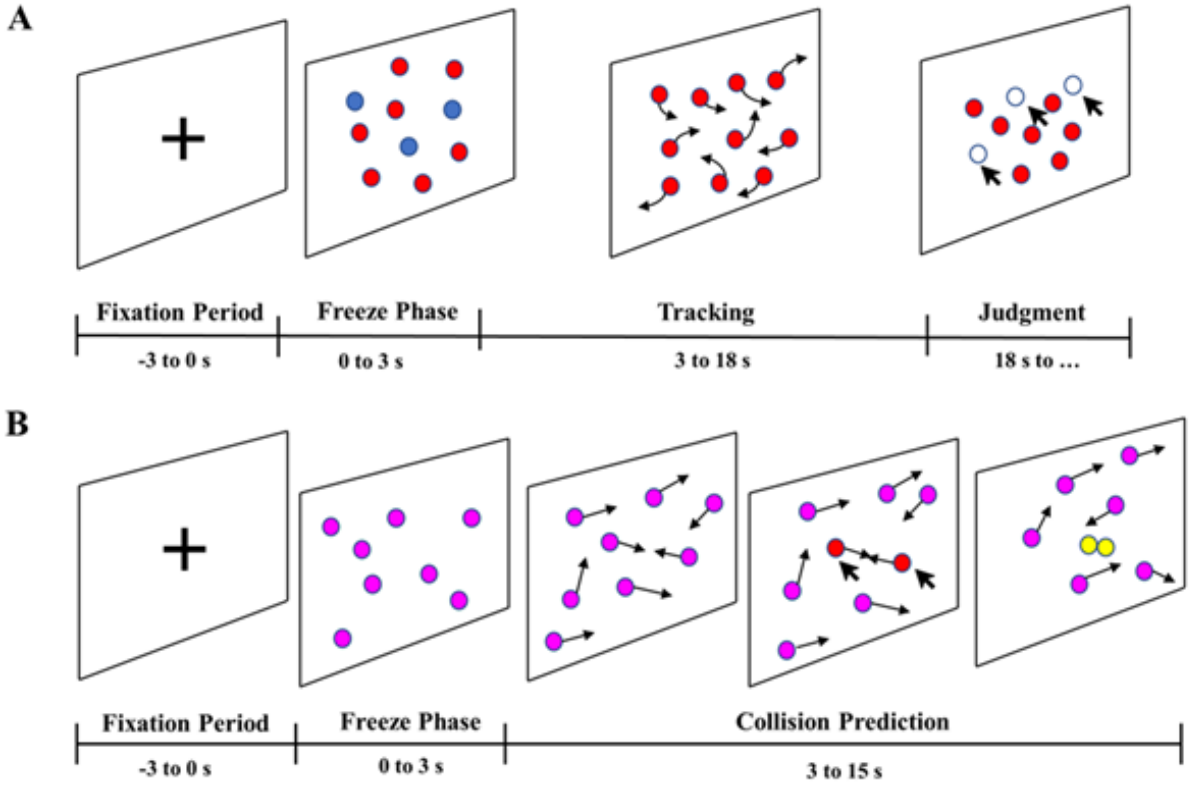

Figure 3.3: The tasks' experimental design: (A) illustrates the design of the tracking task and (B) illustrates the design of the collision prediction task. Note that the number of dots shown in these diagrams is for representation purposes only.

To prevent random guesses, the number of dots that participants could select was limited to two. Once the participant selected a dot, it would change from pink to red. The levels of workload were manipulated by varying the number of dots. As shown in Table 3.2, the low workload condition had six dots, the medium workload condition had 12 dots, and the high workload condition had 18 dots.

Individuals could simultaneously track up to eight objects [285] by focusing on a central point. Therefore, there may not be any need to track each object's motion individually but maintaining the focus at the center of the moving objects could help track the object's trajectory and position. In the tracking task, participants were asked to keep track of the target dots (initially blue) moving randomly among distractor dots for

15 seconds. As participants could simultaneously track the motion of up to eight objects, we tested the experiment design with different numbers of tracking dots. In a pilot study with nine participants, we determined the optimal number of blue dots and the total number of dots to elicit significant variations in the tracking accuracy while ensuring participant engagement.

Further, in the collision prediction task, the participant's task was to predict the trajectory of the dots and identify which pair of dots would collide. Here, the pilot study determined the total number of dots and also the number of collisions to finalize the experiment design. So, in the collision prediction task, the total number of dots and the number of collision for each workload level was determined based on significant changes in time before collision and collision prediction miss rate for the participants in the pilot study. Both tasks were executed by a JavaScript program and could be run on any computer with an Internet Browser, keyboard, and mouse. Both of these tasks were displayed on a 15-inch monitor with a resolution of 1920 x 1080.

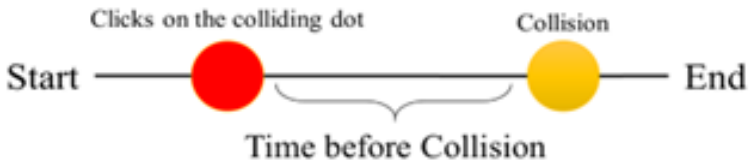

Figure 3.4: A diagram that explains the process of calculating the time before the collision in the collision prediction task.

Each individual had to complete 108 trials of each task, with 36 trials of each workload level. The entire experiment was divided into four blocks, and each block included 27 trials of the tracking task and 27 trials of the collision prediction task. The workload conditions in the trials were randomized within a block to avoid any habituation or expectation effects.

At the end of each trial, participants received feedback on their performance with a message on the screen, for the tracking task it was - "You have correctly tracked x dots out of y dots to track" and for the collision prediction task it was "You have correctly detected this collision" or "You have missed this collision". After reading the performance feedback, they could move to the next trial by pressing the spacebar key. After completing each block, participants were prompted to take a 5-minute break before proceeding to the next block by pressing the spacebar key.

Before starting the experiment, all participants went through a training session that typically lasted around 10 minutes. During this session, they completed six trials of

Table 3.1: Workload Manipulation in Tracking Task

Workload Level	Tracking Dots	Total no.of dots
Low	1	10
Medium	3	12
High	5	15

Table 3.2: Workload Manipulation in Collision Prediction Task

Workload Level	Total no.of dots
Low	6
Medium	12
High	18

each task to get familiar with the tasks and develop strategies for performing them successfully. Participants were encouraged to continue training until they felt comfortable with the tasks. After the training, all participants performed the tasks for about 1.5 hours while we collected EEG, eye activity, and HRV data.

3.2.3 Data Analysis

3.2.3.1 Behavioural and Performance Data Analysis

To evaluate each participant's performance in the tracking task, we measured their tracking accuracy. We determined the tracking accuracy for each trial by calculating the proportion of correctly tracked dots to the total number of dots that needed to be tracked.

$$TA = correct/total \tag{3.1}$$

where TA is the tracking accuracy, correct is the number of dots correctly tracked, and the total is the total number of dots to track.

To measure performance in the collision prediction task, we used two metrics: time before collision and collision miss proportion rate. The time before collision is the duration between when the participant clicks on one of the colliding dots and when the collision occurs (as shown in Figure 3.4). The collision miss proportion rate for a particular workload level of the collision prediction task is the ratio of the number of collision prediction misses to the total number of collisions for that workload level. A collision

miss was recorded when the participant was unable to identify the pair of dots that would collide and thus did not click on either dot before the collision occurred.

$$Rate = missed/total \tag{3.2}$$

where Rate is the collision prediction miss proportion rate, miss is the total number of trials where participant missed correct collision, and total is the total number of collisions.

3.2.3.2 EEG Preprocessing

We preprocessed EEG data using the EEGLAB v2020.0 toolbox in MATLAB R2019a (The Mathworks, Inc., Natick, MA, USA) as shown in Figure 3.5. The data was downsampled to 250 Hz, and a band-pass filter of 2-45 Hz was applied. Channels that had three seconds or more of flat lines were removed using the clean_flatline function. We also removed noisy channels using the clean_channels function in EEGLAB. On average, we removed 3 ± 1 channels, and we restored these channels by interpolating data from neighboring channels using the spherical spline method from the EEGLAB toolbox.

We removed continuous artifactual regions using the EEGLAB function pop_rejcont. We divided the data into 0.5-second epochs with an overlap of 0.25 seconds and considered a frequency threshold of 1 to 100 Hz. Each selected artifactual region consisted of at least four contiguous epochs with high-frequency data (spectrum over 10 dB). Afterward, we cleaned the windows using the clean_windows function in EEGLAB, which calculates the power of each sliding window of one-second length, transforms it to a z-score, and rejects all windows that fall outside of 5 standard deviations. After these artifact removal steps, we extracted two EEG datasets: one for tracking trials and one for collision prediction trials. Each participant had 34 ± 2 high workload, 35 ± 1 medium workload, and 34 ± 1 low workload tracking trials, and 32 ± 2 high workload, 33 ± 2 medium workload and 33 ± 1 low workload collision prediction trials.

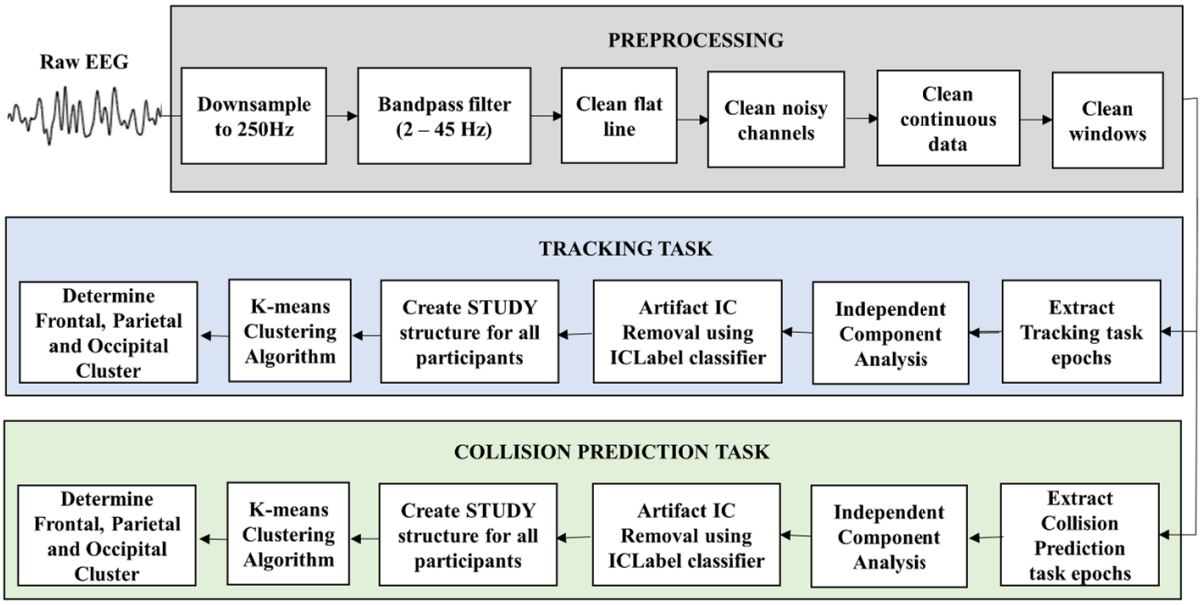

Figure 3.5: The pipeline for preprocessing and processing EEG data for the tracking and collision prediction tasks.

The tracking epochs were 21 seconds long, which included the 3 seconds of the fixation period, followed by 3 seconds of the freeze phase, and then the tracking task began. The collision prediction task epochs were 15 seconds in length, which included the initial 3 seconds of fixation period, followed by 3 seconds of freeze phase, and then, the collision prediction task started. Both the tracking and collision prediction datasets were decomposed using Independent Component Analysis (ICA) performed using EEGLAB's runica algorithm [84]. Finally, we employed ICLabel [276], an automatic IC classifier, to identify components related to brain, heart, line noise, eye, muscle, channel noise, and other activities. This tool was used to generate class labels for each component, and all the components that were not related to brain activity were rejected.

IC Clustering

We used the EEGLAB STUDY structure [85] to manage and process data recorded from multiple participants as it enables component clustering to group similar independent components across participants and allows statistical comparisons of component activities for different workload conditions. We used clustering functions to investigate the contributions of frontal and parietal clusters of independent components (ICs) to the workload dynamics. Research has shown that frontal and parietal brain regions

reflect changes in workload [12, 13, 45, 89, 221, 287, 319], and since both of our tasks also manipulate visual load, we particularly focused on the frontal, parietal and occipital clusters of brain activity.

We created a Study for each task, and each Study had one group with 24 participants, with three conditions corresponding to the three levels of workload. Since the dataset of each participant was recorded in a single session, the resulting independent component maps were the same across all three conditions for each participant. For each participant, we selected only those ICs with a residual variance (RV) less than 15% and inside the brain volume, which we achieved using the Fieldtrip extension [258]. We used the k-means clustering algorithm [143] to cluster independent components across all participants into clusters based on two criteria with equal weight (weight = 1):

1. scalp maps
2. their equivalent dipole model locations, which was performed using DIPFIT routines in EEGLAB [259].

We identified the Talairach coordinates [195] of the fitted dipole sources of these clusters to select frontal, parietal, and occipital clusters.

We subsequently calculated the grand-mean IC event-related spectral power changes (ERSPs) for each condition for each cluster. ERSPs show the relative change in power at components with respect to a baseline period before the stimulus [210]. The three seconds of fixation phase in each tracking and collision prediction epoch was taken as the baseline to see the changes in power spectra during the task. We examined the ERSPs for frontal, parietal, and occipital clusters for both tracking and prediction tasks. To compare the ERSPs of different workload conditions, we used permutation-based statistics implemented in EEGLAB, with a Bonferroni correction and a significance level set to p = .05. Also, for the frontal, parietal, and occipital clusters, we calculated the spectral powers of each IC using EEGLAB's spectopo function, which uses Welch's periodogram method [352] on each 2-s segment using a Hamming window with 25% overlap for a range of frequencies from 2 to 45 Hz. For each IC, we examined the power spectral density (PSD) at different frequency bands to identify the correlates of mental workload.

3.2.3.3 Eye Activity data

The Pupil Core software, Pupil Capture, provides the pupil size for the left and right eye separately, along with the associated confidence value, which represents the quality

of the detection result. The eye activity data was processed as shown in Figure 3.6. We removed all data points where the confidence of the pupil size was less than 0.8 from the data. We low-pass filtered the pupil size data (using a minimum order finite impulse response filter) at 4 Hz [280].

As pupil size is a continuous measurement that varies across participants, we normalized the raw pupil size data using the baseline data (defined as the three seconds of fixation period in each tracking and collision prediction epoch). We also extracted the number of blinks during each trial from the pupil size measurement when the pupil size and confidence of the measurement, reported by the Pupil Capture software, suddenly dropped to zero.

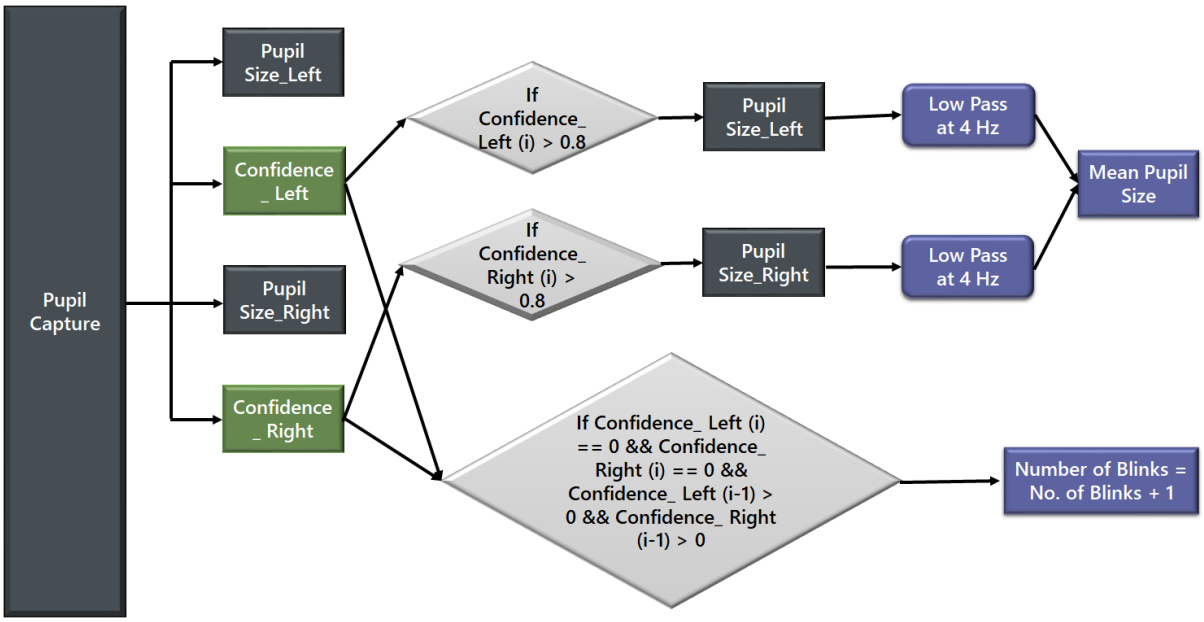

Figure 3.6: The pipeline for Eye activity Processing.

3.2.3.4 Heart Rate Variability

We computed the inter-beat-interval (IBI) time series from the Blood Volume Pulse (BVP) data of each tracking and collision prediction trial. We then calculated the Root Mean Square of the Successive Differences (RMSSD) by using the PeakUtils Python package [244] to detect peaks of the BVP and measure the intervals between adjacent beats, as shown in Figure 3.7.

$$RMSSD = \sqrt{\frac{1}{N}\sum_{1}^{N}(IBI_{i-1} - IBI_i)^2} \tag{3.3}$$

We normalized the RMSSD data by using the three seconds fixation period in each tracking and collision prediction epoch as the baseline.

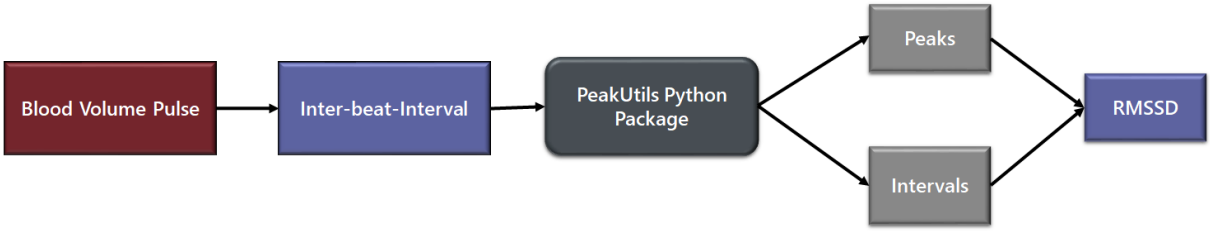

Figure 3.7: The pipeline for Heart Rate Variability Processing.

3.2.4 Statistical Analysis

We used the SPSS statistical tool to conduct statistical analyses. To investigate the differences in performance, EEG, eye activity and HRV parameters across participants in the three workload levels of tracking and collision prediction tasks, we performed a one-way repeated-measures analysis of variance (ANOVA) with workload level (low, medium or high) as the within-subjects factor. We tested for sphericity using Mauchly,Äôs test. If sphericity was not satisfied (p<.05), we applied Greenhouse-Geisser correction. If the main effect of the ANOVA was significant, we then performed post-hoc comparisons to determine the significance of pairwise comparisons, using Bonferroni correction. Finally, we used multiple linear regression to relate EEG, eye activity, and HRV metrics to the performance in the tracking and collision prediction tasks, using EEG power, eye activity, and HRV metrics as predictors, and the performance in the task as the dependent variable.

3.2.5 Mental Workload Classification of Tracking and Collision Prediction Tasks

The EEG data collected for the tracking and collision prediction tasks was used to develop classification models for each task. We employed the cleaned data in developing the workload classifiers for tracking and collision prediction tasks. We used the deep learning

models: EEGNet, Shallow ConvNet, and Deep ConvNet for the workload classification. We performed hyperparameter optimization for choosing parameters for each model.

Each epoch of the tracking task was 15 seconds in length and the tracking workload data was arranged into data of dimensions 64 x 3750 x 2592. Therefore, the total data available for training and testing the workload classifier on 24 participants over four blocks was 64 x 3750 x 2592. The deep learning models were trained on 60% of the data: 64 x 3750 x 1556 and 64 x 3750 x 518, 20% of the data was used for validation. Adam optimizer was used with a batch size of 128, the dropout rate was 50%, and the number of training epochs was 1000. The model was tested on the remaining 20% of the data. The average accuracy, F1-score, and coh-kappa of the tracking classifier are reported in Chapter 4.

Each epoch of the collision prediction task was 12 seconds in length, however, we only considered the last seven seconds of each collision prediction trial. This was selected based on the observation of EEG features in the last seven seconds of the trial. Therefore, the collision prediction workload data was arranged into data of dimensions 64 x 1750 x 2592. Therefore, the total data available for training and testing the collision prediction workload classifier on 24 participants over four blocks was 64 x 1750 x 2592. The deep learning models were trained on 60% of the data: 64 x 3750 x 1556 and 64 x 3750 x 518, 20% of the data was used for validation. Adam optimizer was used with a batch size of 32, the dropout rate was 50%, and the number of training epochs was 1000. The model was tested on the remaining 20% of the data. The average accuracy, F1-score, and coh-kappa of the tracking classifier are reported in Chapter 4.

3.3 pHRC Experiment Methodology

3.3.1 Participants

The study was reviewed and approved by the University of Technology Sydney's Human Research Ethics Committee (ETH21-6346). A total of 24 participants (13 males, and 11 females, with an average age of 23 $\pm$ 4, all of whom were right-handed) took part in the study. The experiment density plot is shown in Figure 3.8.

This data collection was conducted during the second COVID lockdown in Sydney, Australia. We took additional care to follow the additional COVID-19 safe protocol based on New South Wales guidelines (ensured that participants were vaccinated as per the state guidelines), which was further approved by the UTS Ethics Committee and the

University Risk Assessment team. As most of the population was vaccinated by then, we were able to recruit participants comparably easier as compared to the ATC experiment.

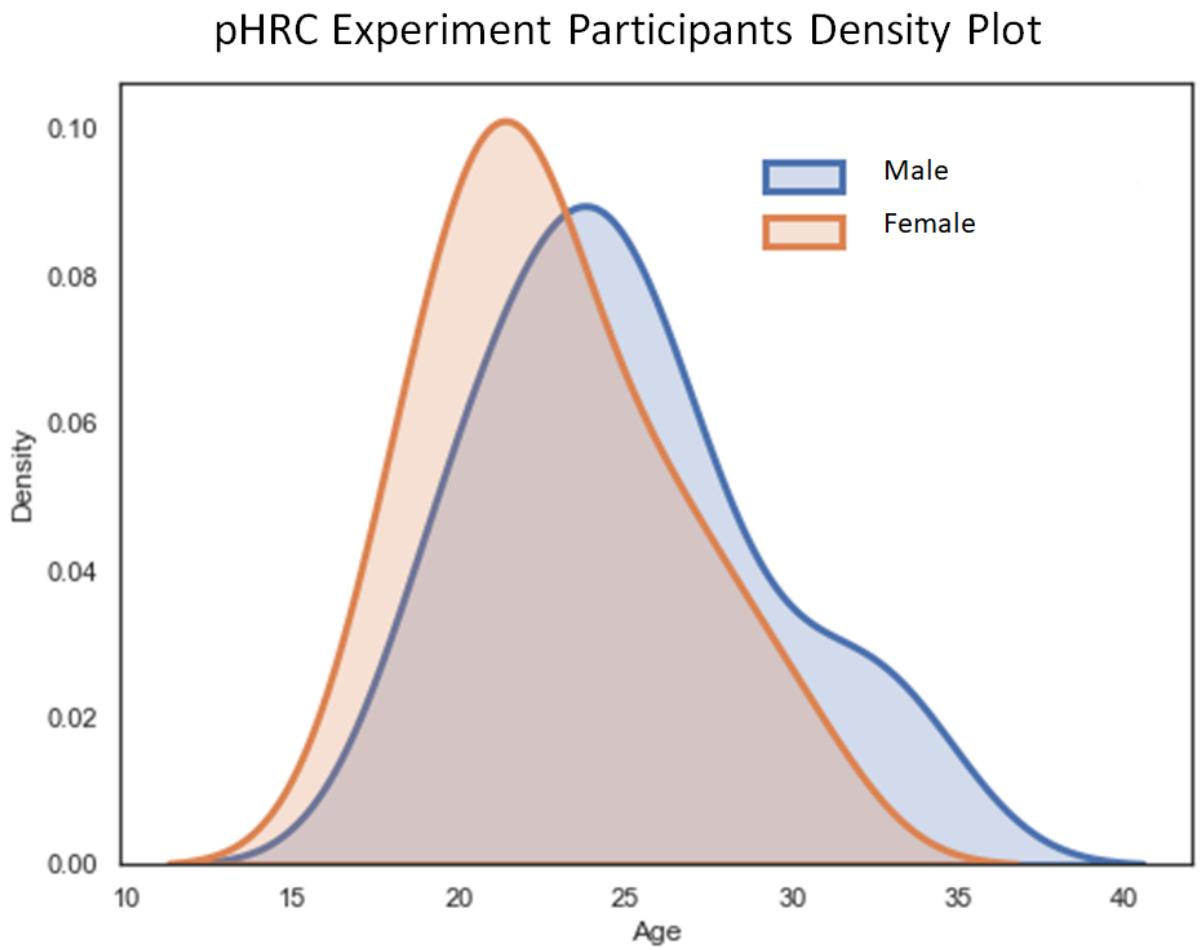

Figure 3.8: pHRC Experiment Participant Distribution.

3.3.2 Experimental Procedures

The EEG data were collected using a wireless MOVE system from Brain Product GmbH in Germany, which featured 29 channels. The electrodes were kept at an impedance of less than 20 kkΩ, and the data were sampled at a rate of 1 kHz. The events from the experiment were synced with the EEG data using Lab Streaming Layer [191].

The ANBOT [52], an intelligent industrial robot for abrasive blasting featuring a UR10 robot arm by Universal Robots, was the collaborative robotic partner in this experiment. ANBOT was specifically designed for collaborative grit-blasting tasks, and featured a six-axis force-torque sensor positioned between the robotic arm and its end-

effector to monitor the forces exerted by the human co-worker. The participants were in contact with the robot arm through a handle mounted on the robot end-effector (see Figure 3.9(A)). The interaction forces and torques are measured and controlled by the ATI mini-45 transducer, placed between the robotic arm and the end-effector.

The control system employs an admittance-based approach to translate the forces applied by the operator on the handles into desired robot movements within the Cartesian space. To ensure safe operation without interfering with the robot's admittance, collision and singularity avoidance strategies were implemented. These strategies limited the task to a specific region within the robot's workspace, ensuring that the manipulator remained clear of potential self-collisions and singular configurations.

During the experiment, the human co-worker continuously gripped the handlebar attached to the robot's end-effector, allowing for direct control. An accompanying monitor displayed the real-time trajectory of the robot's end-effector. Notably, for simplicity, the experiment constrained the robot's motion to two dimensions, enabling the human co-worker to move the robot exclusively in vertical and horizontal directions that aligned with the monitor's orientation.

Prior to commencing the experiment, we provided participants with instructions on how to operate the robot and informed them about the safety precautions in effect. Additionally, participants were given the opportunity to familiarize themselves with the robot's control triggers, although this practice was separate from the actual experimental task.

During the experiment, participants were tasked with completing an interactive game called the 'Clock Game' on a large LED TV. The game was designed to emulate a blasting task and used an arithmetic task to manipulate the participants' workload. The point where the nozzle of the ANBOT robot arm was pointing was represented by a blue dot on the screen, and the robot was limited to 2 degrees of freedom for the experiment. Participants were instructed to use the buttons on the handles of the robot end-effector to move the arm and paint the area where the nozzle was moving, simulating a paint sprayer scenario. This required the use of both hands and kept the participants engaged in the task.

The experiment was divided into three parts, each consisting of 90 trials that were randomly ordered (Figure 3.9). Each section began with a 10-second period where participants focused on a cross on the screen. Following that, a math equation would appear on the screen for them to solve. The equations could involve adding or subtracting one, two, or three-digit numbers.

In the low workload condition, the arithmetic equation involved single-digit numbers, such as the addition or subtraction of a one-digit number (e.g., 9 - 1). In the medium workload condition, the arithmetic equation involved the addition or subtraction of two-digit numbers with no carrying over (e.g., 13 + 24). In the high workload condition, the arithmetic equation was more complex with the addition or subtraction of three-digit numbers, which involved carrying over or borrowing in the ones and tens places (e.g., 135 + 168 or 423 - 187).

The participants were presented with an arithmetic equation and had to solve it within 5 seconds. After that, a large circle with a red dot in the center and 4 green dots along the perimeter appeared on the screen. Each green dot had a number next to it, and the participant's task was to identify the green dot that matched the answer to the equation they had solved. They had 2 seconds to identify the correct target during the "Identify Target" phase.

After identifying the correct target, the participant had to reposition the blue dot back to the center of the big circle, where the red dot was located. During this phase, the participant was given new instructions to move the robot arm to the correct target while simultaneously completing a new arithmetic equation of varying difficulty. The participant had five seconds to complete the task; otherwise, the trial would reset.

To keep the participants engaged, we displayed a score on the screen that was based on their accuracy in identifying the target and the sprayed area. To make sure that the participants were completing the calculations while moving the robot arm, we removed the arithmetic equation from the screen as soon as the arm reached the target.

As the participants were not required to input the correct answer to the calculation, we desired to eliminate any chances of any random guessing in the 'Identify Target' phase. In a pilot study, we realized that participants developed strategies to calculate the correct answer without performing the entire calculation. One of the most common strategies used by the participants involved guessing the correct answer by calculating only the last digit of the correct answer in the medium (two-digit calculations) or high workload (three-digit calculations) conditions. In order to prevent participants from employing this strategy of guessing the correct answer by just figuring out the last digit, we selected the options in a manner that there were multiple options with the same last digit and second digit.

Also, in the pilot study, it was observed that participants continued with the calculation even after they reached the target. In order to avoid this problem, we removed the equation from the screen as soon as the participant reached the target and immediately

followed it with a very short two-second 'Identify Target' phase.

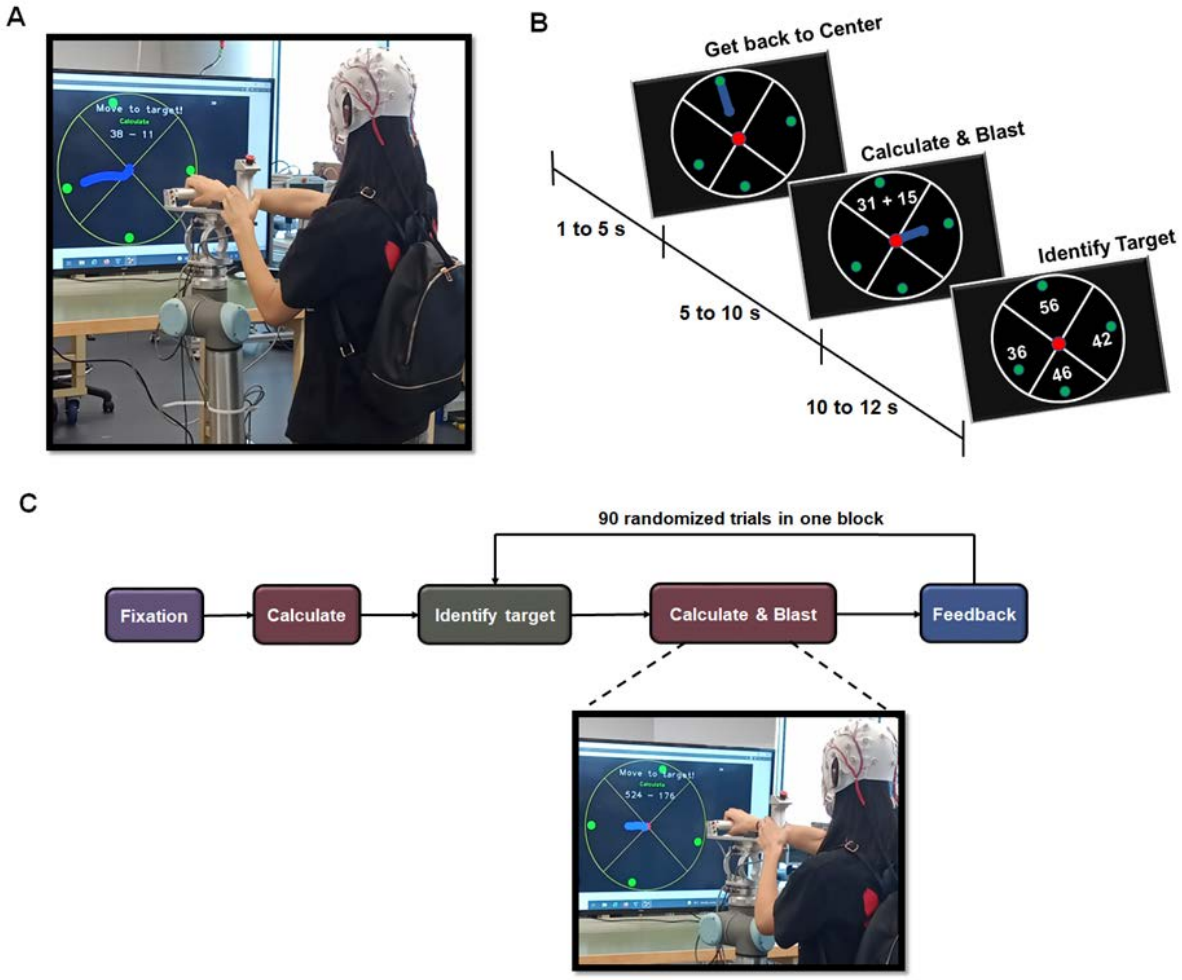

Figure 3.9: The experimental design included (A) a participant wearing an EEG cap with the transmitter in a backpack in front of a screen. The screen displayed the 'Calculate and Blast' portion of the trial, where the participant was moving to a target position while performing a medium workload arithmetic calculation of 38-11, (B) an illustration of an example trial, and (C) the layout of each experimental block.

Each experimental block consisted of 30 trials where the robot arm would stop unexpectedly during the blasting phase, creating a cognitive conflict. This was done by placing an invisible obstacle on the path to the target or green dot. The screen would also turn black when the participant reached the obstacle to indicate an error. To prevent the participant from circumventing the obstacle, it was designed as an annular barrier that restricted the movement of the robot arm. The arm would remain stuck until the blasting phase reset.

The chance of the invisible obstacle appearing in the blasting path was set at one-third to prevent participants from anticipating a sudden stop by the robot. As a result, there were 90 trials in each workload condition, with 60 normal blasting trials and 30 error trials, and the trials with sudden robot stopping were randomly distributed.

After the blast phase, the participant received feedback with messages, such as "Yay! Correct target", "Oops.. Wrong target" or "Sorry.. time is up" in case of the normal trials. They received feedback such as "Aborting trial.. Correct Direction", "Aborting trial.. Wrong Direction" or "Sorry.. time is up" in case of obstacle trials. After the feedback, the participants were asked to identify the target again based on the equation's answer before they were instructed to 'Get back to centre'. During each block, the participants were randomly asked to complete the NASA-TLX questionnaire to record their subjective workload ratings.

The participants were given a training session before the main experiment, where they completed 30 trials of the task to become familiar with the task and the movement of the robot arm. The experiment lasted for approximately 1.5 hours and consisted of three blocks, with a break of 10 minutes between each block for rest.

3.3.3 Data Analysis

3.3.3.1 Subjective Measures Analysis

The participants were requested to evaluate their workload for each condition by means of NASA-TLX scores, which were gathered in random order during each block of the experiment. Subsequently, the mean scores for each subscale were calculated for each condition.

3.3.3.2 Behavioural and Performance Data Analysis

The success of the blasting task was gauged by measuring three factors: target identification accuracy, blasting time, and blasting miss rate. To evaluate target identification accuracy, the proportion of trials where the participant correctly identified the target out of the total number of trials was calculated for each workload level.

$$Accuracy = nCorrect/nTotal \tag{3.4}$$

where nCorrect was the number of trials where participants correctly identified and blasted the target, and nTotal was the total number of trials. The amount of time it took

to move the robot arm from the center of the circle to the designated target was also taken into account for the various workload conditions. Additionally, we evaluated the blasting miss rate for each workload condition, which was determined by the proportion of trials where the participant failed to reach the target within the 5-second time limit out of the total number of trials.

$$BlastingMissRate = nMiss/nTotal \tag{3.5}$$

where nMiss was the number of trials in which participants did not reach the target and nTotal was the total number of trials. To gain additional insights, the average force exerted and the average speed of the end-effector were extracted from the force-torque sensor (ATI Mini-45) located between the handle and the blasting nozzle.

To evaluate the performance of the obstacle trials in the blasting task, two aspects were considered: accuracy in identifying the target and the time required to reach the obstacle. The accuracy of target identification for each workload level was established by calculating the ratio of the number of trials where the participant correctly identified the target, as indicated by the direction of movement of the nozzle, to the total number of trials.

$$accuracyObstacle = nCorrectO/nTotalO \tag{3.6}$$

where nCorrectO was the number of trials where participants correctly identified the target and moved in the correct direction before reaching the obstacle, and O was the total number of obstacle trials. For the various workload conditions, the duration of time taken to move the robot arm from the center of the circle to the invisible obstacle in the blasting path, which is any of the potential targets, was also taken into account. Additionally, the average force exerted and the average speed of the end-effector were obtained from the force-torque sensor (ATI Mini-45) located between the handle and blasting nozzle.

3.3.3.3 EEG Processing

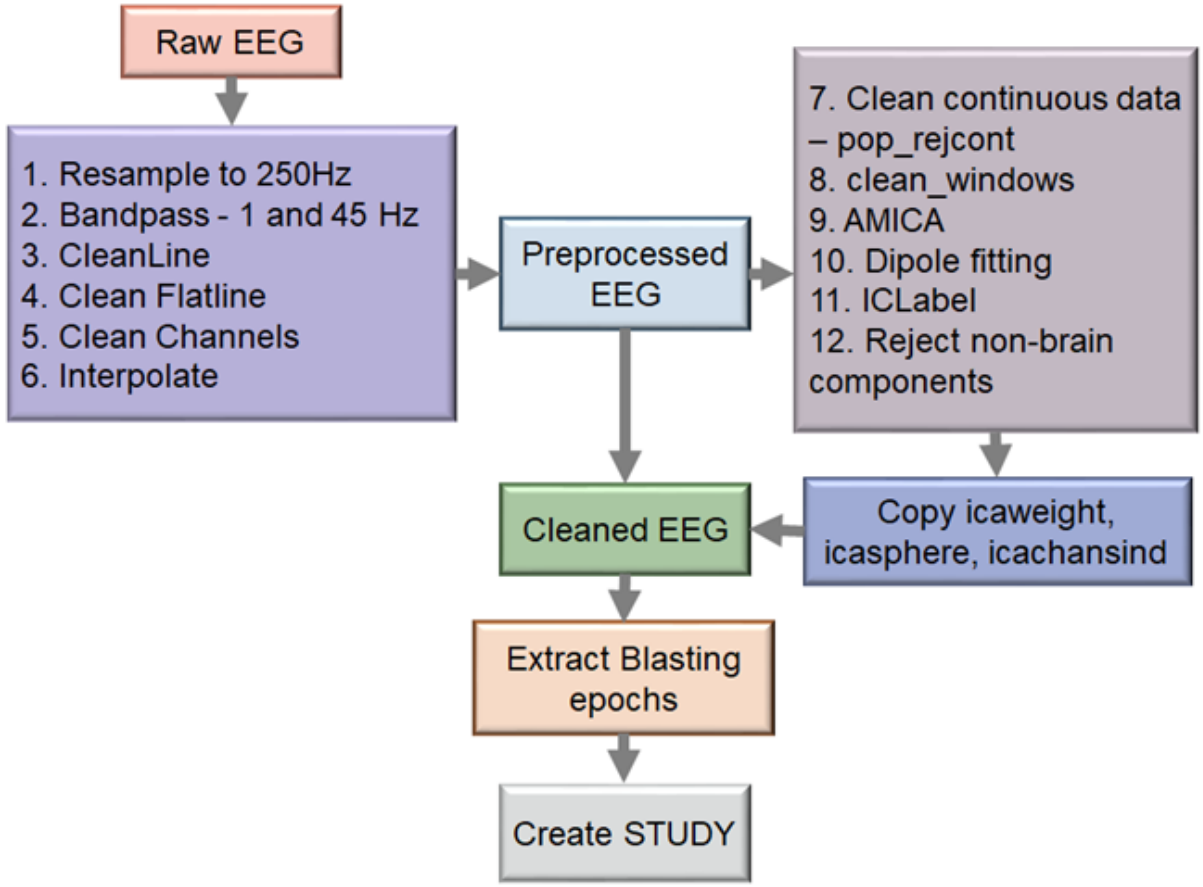

Figure 3.10: EEG data processing.

The EEG data was preprocessed using the EEGLAB v2022.0 toolbox [84] in MATLAB R2020b, as depicted in Figure 3.10. The data was first downsampled to a rate of 250 Hz, then filtered between 1-45 Hz to remove unwanted frequencies. Line noise and channels that showed flat lines for more than three seconds were eliminated through the use of the clean_line and clean_flatline functions. Additionally, noisy channels were identified and removed by the clean_channels function. On average, 2 ± 0.4 channels were removed, which were later restored using the spherical spline method, which interpolates data from neighboring channels. A copy of this dataset was made, and using the pop_rejcont function, continuous artifactual regions were cut off. Then, the clean_windows function was used to perform window cleaning. After these artifact removal steps, ICA was performed using the AMICA algorithm in EEGLAB.

After we used ICA to analyze each participant's data, we used ICLabel to identify

and discard any components that had less than 30% brain activity. Then, we copied the ICA information (icawinv, icasphere, and icaweights) back to the original dataset. From there, we extracted all the relevant epochs and created three separate datasets for each workload level.

IC Clustering

The data from multiple participants was organized and processed using the EEGLAB STUDY structure [85]. A study was created for a group of 24 participants, with three conditions corresponding to the three workload levels. Only the components within the brain, as determined using the Fieldtrip extension [258], and with a residual variance of less than 15% were selected for each participant. The components for all participants were then grouped together using the k-means clustering algorithm [143], based on three criteria that were given equal weight (weight=3):

1. The location of the dipoles,
2. The orientation of the dipoles, and
3. The scalp maps.

To identify the frontal, central, and parietal clusters, Talairach coordinates [195] were used to locate the fitted dipole sources of these clusters. The power spectral density (PSD) of each component was then evaluated for the frontal, central, and parietal clusters, using the spectopo function on 2-second segments with a 25% overlap and a Hamming window. The mean PSD in the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-45 Hz) bands were examined for each independent component (IC) of the clusters to identify any correlation with mental workload. The PSD calculated for each IC was normalized using the fixation period as the baseline to investigate changes in PSD during blasting under various workload conditions.

3.3.3.4 EEG Processing for Normal and Conflict Trials

IC Clustering

Multiple participant data was organized and processed using the EEGLAB STUDY structure [84]. A study was created with a group of 24 participants, with six conditions that corresponded to the three levels of workload and the presence or absence of obstacles. Only the components within the brain, as determined using the Fieldtrip extension [258] and with a residual variance (RV) of less than 15%, were selected for each participant, as

shown in Figure 3.11. The components for all participants were then grouped together using the k-means clustering algorithm [143] based on three criteria, each of which was given equal weight (weight=3):

1. dipole locations,
2. dipole orientation, and
3. scalp maps.

The location of the fitted dipole sources for the frontal, central, and parietal clusters were identified using Talairach coordinates [195], which were then used to determine the location of these clusters.

The power spectral density (PSD) of each component in the frontal, central, and parietal clusters was evaluated using the spectopo function on 1-second segments, with a Hamming window. The mean PSD in the delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-45 Hz) bands were examined for each independent component (IC) of the clusters, specifically for the one second immediately after reaching the target in normal conditions and for the one second immediately after reaching the obstacle in the obstacle trials. The PSD calculated for each IC was normalized using the fixation period as the baseline to compare changes in PSD upon reaching either the target or obstacle under different workload conditions.

We also conducted a simple event-related potential (ERP) study with a group of 24 participants, in which three conditions were created to correspond to the three levels of workload. Only the components within the brain, as determined using the Fieldtrip extension [258] and with a residual variance (RV) of less than 15%, were selected for each participant. The components for all participants were then grouped together using the k-means clustering algorithm [143] based on three criteria, each of which was given equal weight (weight=3):

1. dipole locations,
2. dipole orientation, and
3. scalp maps.

To evaluate the event-related error potential (ERrP), particularly in the theta band [380], under the three workload conditions, a frontal cluster that is located near the anterior cingulate cortex (ACC) was identified [212, 345]. The clustered components were

then used to extract the PEN and Pe from the back-projected and channel-based ERPs. The channel-based ERPs were filtered between 4-8 Hz and smoothed at 10 Hz. PEN was calculated as the minimum error negativity amplitude in the 100-350 ms window after reaching the obstacle, and Pe was calculated as the maximum error positivity in the 350-450 ms window after reaching the obstacle.

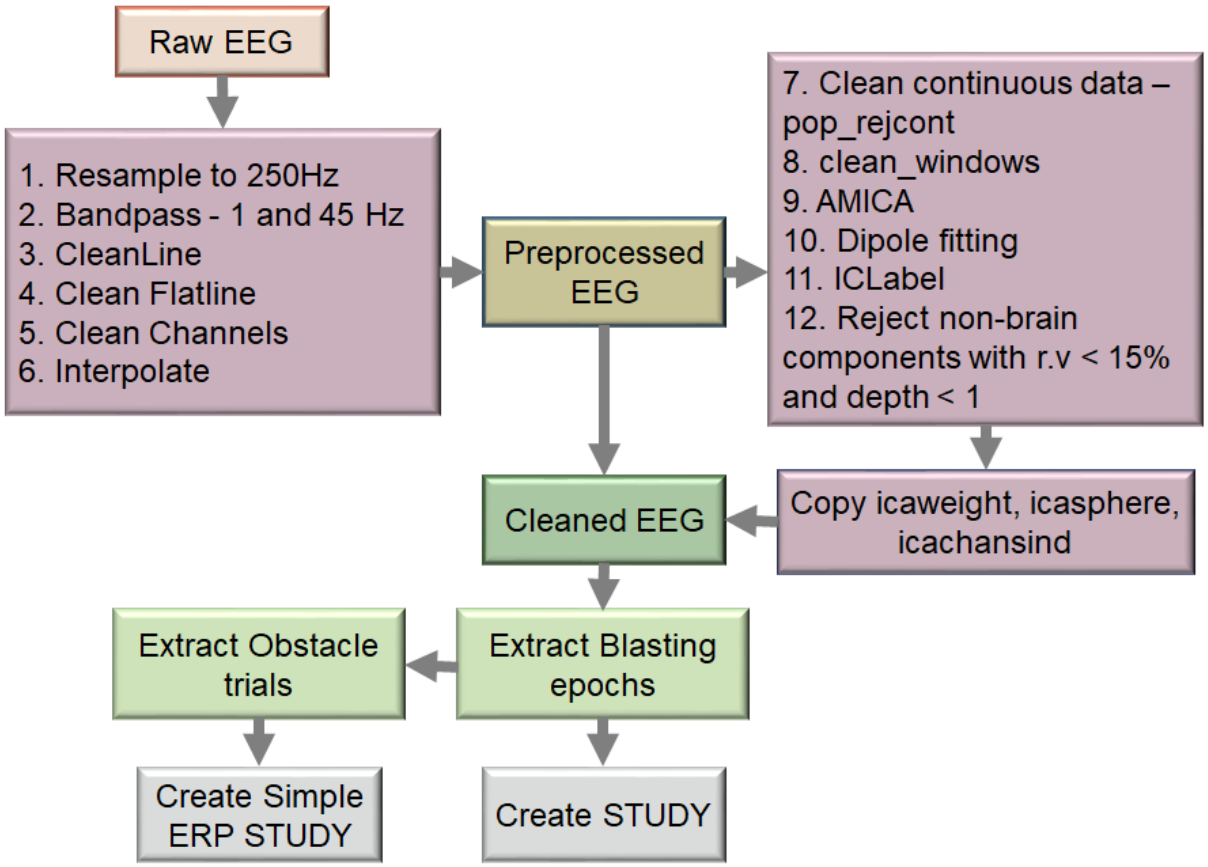

Figure 3.11: EEG data processing for obstacle trials.

3.3.4 Statistical Analysis

All statistical analyses in this study were conducted using the SPSS (IBM SPSS 26.0; Chicago, IL, U.S.A.) tool. The variations of the NASA-TLX scores, performance, and EEG parameters of the three workload levels of the task were investigated by conducting a one-way repeated-measures analysis of variance (ANOVA) with the workload level as the within-subjects factor. Mauchly's test was used to test for sphericity, and if sphericity was not met ($p < .05$), Greenhouse-Geisser correction was applied. Posthoc comparisons

were conducted with Bonferroni corrections (α = 0.05) if the main effect of ANOVA was found to be significant.

Additionally, Pearson correlation analyses were conducted to identify the relationship between the subjective workload measures and the EEG powers. The correlation analyses were also done to examine the association between EEG powers and the applied human force. Multiple linear regression was also applied to predict physical performance from EEG power using the Enter method. Furthermore, Pearson correlation analyses were performed to examine the relationship between the ERrP measures and the normalized applied human force.

3.4 Online Safety Alert tool for Physical Human Robot Collaboration

3.4.1 Online Safety Alert Tool Framework

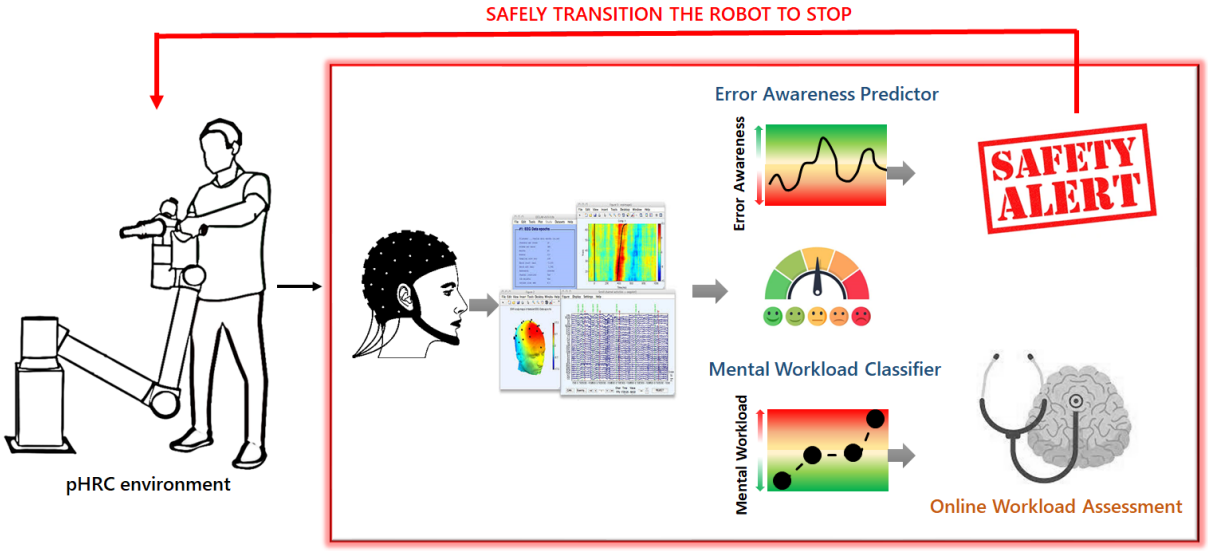

Figure 3.12: Framework for Online Safety Alert tool for pHRC.

We developed a framework for an Online Safety Alert Tool (shown in Figure 3.12) that safely transitions the robot to a stop condition and provides accurate and reliable estimation of error awareness and mental workload of the operator while blasting. For this purpose, we used the data collected in the offline pHRC experiment of the abrasive blasting task combined with a secondary arithmetic task.

The online safety tool comprises of error awareness predictor and workload classifier. The error awareness predictor estimates the error awareness of the operator, represented by prediction error negativity amplitude, error positivity amplitude or normalized power spectral density in the window 0.2-0.9 seconds after the error occurrence in 1-30 Hz frequency bin. For this framework, we selected the threshold for error awareness predictor that predicts PEN as the mean PEN in the lower 30% quartile for all participants under all workload conditions. The threshold for error awareness predictor that predicts Pe amplitude was chosen for testing the online framework as the mean Pe in the lower 30% quartile for all participants under all workload conditions. And the threshold selected for error awareness predictor predicts normalized PSD as the mean PSD in the lower 30% quartile for all participants under all workload conditions.

Another major component of the online safety alert tool is the workload classifier that detects the workload of the human operator into three levels: low, medium, or high. The output of the workload classifier is currently used only for workload diagnosis. However, it could be used to run a closed-loop mental workload adaptive system in the future.

The online safety alert tool was tested using an online simulation of the experiment, and the performance of the tool was validated using within-subjects and cross-subject cross-validation.

3.4.2 Error Awareness Predictor

The simulated offline and online EEG data during the blasting period was preprocessed using the MNE package in Python [129]. The data was down-sampled to 250 Hz using the mne.filter.resample function and bandpass filtered between 1 and 45 Hz using mne.filter.filter_data. The line noise was removed using mne.filter.notch_filter function in the mne package.

In the case of the error awareness predictor, for each window of two seconds in length, the event-related desynchronization and event-related synchronization were computed for frequencies ranging from 1-45 Hz and all channels using mne.time_frequency.tfr_multitaper algorithm. Therefore, each window or epoch of two seconds was transformed to the time-frequency data of dimensions 29 x 44 x 500. We rearranged the data by slicing the data corresponding to each channel over the other, reshaping the data to dimensions 1276 x 500. Therefore, the total data available for training and testing the error awareness predictor models on 24 participants over three blocks was 2160 x 1276 x 500.

The output of the predictor, error awareness, was represented by:

- prediction error negativity (PEN),
- error positivity amplitude,
- mean power spectral density (PSD) between 1-30 Hz at the FCz channel in the window 0.2-0.9 seconds after the error.

Prediction error negativity amplitude was identified as the minimum ERP value in the 0.1-0.3 seconds window after the error, and the error positivity amplitude was identified as the maximum ERP value in the 0.3-0.5 seconds window after the error. The prediction error negativity and error positivity amplitudes were detected online using mne.preprocessing.peak_finder algorithm.

The power spectral density was computed using mne.time_frequency.psd_welch algorithm. We removed the inter-subject variability and workload features from the power spectral density considered immediately after the error by subtracting the mean power spectral density of the two-second EEG data prior to the error.

In the case of the error awareness predictor, for within-subjects testing, we used the first two experimental blocks of all the subjects for training, and the last block was used for testing. Therefore, the models were trained on 1440 x 1276 x 500 with 20% of this data used for validation. The model was tested on the online simulation of the last block of each participant in the within-subjects cross-validation. For cross-subject testing, we trained the error awareness predictor on the data of 23 subjects and tested the entire set of error trials of every left-out subject. So, in this case, the model was trained on data of dimensions of 2070 x 1276 x 500 and validated on 20% of this data. It was tested on the 90 error trials of the test subject. The average mean squared error and mean average error of the within-subjects and cross-subject cross-validation are reported in Chapter 7.

We used the deep learning models: EEGNet, Shallow ConvNet, and Deep ConvNet for error awareness estimation. We performed hyperparameter optimization for choosing parameters for each model. Adam optimizer was used with a batch size of 128, dropout rate of 50%, and the number of training epochs was 1000.

3.4.3 Mental Workload Classifier

The simulated offline and online EEG data during the blasting period was preprocessed using the MNE package in Python [129]. The data was down-sampled to 250 Hz using the mne.filter.resample function and bandpass filtered between 1 and 45 Hz using

mne.filter.filter_data. The line noise was removed using mne.filter.notch_filter function in the mne package.

In the case of the workload, we employed the first three seconds of each blasting trial as input and computed the event-related desynchronization and event-related synchronization for frequencies ranging from 1-45 Hz and all channels using mne.time_frequency.tfr_multitaper algorithm. Therefore, each window or epoch of two seconds was transformed to the time-frequency data of dimensions 29 x 44 x 750. We rearranged the data by slicing the data corresponding to each channel over the other, reshaping the data to dimensions 1276 x 750. Therefore, the total data available for training and testing the workload classifier on 24 participants over three blocks was 4320 x 1276 x 750.

In the case of the workload classifier, for within-subjects testing, we used the first two experimental blocks of all the subjects for training, and the last block was used for testing. Therefore, the models were trained on 2880 x 1276 x 750, with 20% of this data used for validation. The model was tested on the online simulation of the last block of each participant in the within-subjects cross-validation with 60 normal trials. For cross-subject testing, we trained the workload classifier on the data of 23 subjects and tested on the entire set of normal trials of every left-out subject. So, in this case, the model was trained on data of dimensions of 4140 x 1276 x 1250 and tested on the 180 normal trials of the test subject. The average accuracy, F1-score, and coh-kappa of the within-subjects and cross-subject cross-validation are reported in Chapter 7.

We used the deep learning models: EEGNet, Shallow ConvNet, and Deep ConvNet for the workload classification. We performed hyperparameter optimization for choosing parameters for each model. Adam optimizer was used with a batch size of 256, the dropout rate was 50%, and the number of training epochs was 2000.

CHAPTER 4

BIOMARKERS OF MENTAL WORKLOAD VARIATIONS IN BASIC AIR TRAFFIC CONTROL TASKS

4.1 Background

People generally prefer not to engage in activities that require them to push their abilities to their limits as it can be a source of stress and frustration [3]. However, not all job settings allow for such avoidance, making it essential to establish a harmonious relationship between a worker's abilities and the work environment [355]. Even though human workers can adapt to various work environments, perform multiple tasks, and operate different equipment at the same time, poorly designed work environments can lead to an excessive amount of sensory information, resulting in an increased workload.

Air traffic controllers are a prime example, as they work in a highly complex environment to guarantee safe and efficient air traffic flow by organizing aircraft routes to ensure they reach their destinations in an orderly and timely manner. They are responsible for anticipating and preventing conflicts between aircraft by monitoring adherence to International Civil Aviation mandated separation standards [300] and managing the complexity that arises. They routinely handle multiple aircraft, coming from different directions and heading to various destinations at varying speeds and altitudes [131]. However, as air traffic continues to increase, there is a growing need to research the

mental factors that contribute to the effectiveness of air traffic controllers.

Mental workload plays a significant role in determining the effectiveness of air traffic controllers, who work in complex and interactive environments. In recent years, their responsibilities have shifted to a more supervisory level, requiring them to integrate multiple streams of information, which in turn demands more cognitive resources [269], leading to increased workload for operators [187, 196, 248]. Mental workload is defined as the dynamic relationship between the cognitive resources required by a task and an operator's ability to provide those resources [355]. Human operators have limited information processing abilities as they have a finite amount of resources with limited capacity [169, 192].

The theory of limited cognitive resources posits that exposure to demanding task conditions can negatively impact performance due to either resource depletion [261] or impeded access to resources [43]. High levels of mental workload can lead to human error [294], compromising system efficiency and safety [368]. To ensure optimal performance, the mental workload of operators should be neither too high nor too low [41, 137], as performance is known to decline under both overload and underload conditions [48, 138, 341, 371].

According to the dynamic adaptive theory, the brain seeks a balance of cognitive resources and comfort, and extreme task demands can hinder adaptability and performance [139]. By predicting and adjusting mental workload, and modifying task allocation, it is possible to maintain high performance and avoid loss of situational awareness, especially in safety-critical work environments. Accurate and reliable measurement of mental workload is vital in order to create better work environments and human-machine interactions [12, 46, 309].

Researchers have employed various strategies, such as self-assessment, performance measures, and physiological metrics, to evaluate mental workload, each with its own advantages and limitations [254, 362]. The results obtained from these different methods for measuring mental workload can often be dissociated [370], as their sensitivity can vary depending on the operator's workload [82]. Some commonly used subjective measures include the Instantaneous Self-Assessment (ISA) questionnaire [44, 167, 180], NASA Task Load Index [142], and the Subject Workload Assessment Technique [296]. These are used to gauge the workload of the operator.

Assessing mental workload is a complex task and can't be done accurately by using subjective measurements alone [83]. The method of assessment shouldn't impede on the task at hand or alter the operator's mental state, which can be the case with some

subjective assessment techniques using questionnaires. Another widely used method is performance-based workload measurement, which, like the subjective assessment method, provides a retrospective workload assessment. However, performance-based measures can only provide a partial picture as the same performance level can be achieved while experiencing different levels of workload [16].

Physiological metrics have been used over the years to assess workload [54, 57, 221] as it is highly sensitive, diagnostic, and mostly non-intrusive, providing accurate real-time assessment of the operator's workload [266, 380]. Using physiological data such as neurophysiological signals to assess mental workload online, without influencing the task, as there is no explicit output [125, 263, 264]. Neurophysiological measures can also assess changes in mental state that are not discernible in overt task performance [13, 38, 264, 355].

Neurophysiological measures, such as the electroencephalogram (EEG) signal, have been frequently utilized to gauge mental workload as the effects of task demands can be distinctly observed in variations of EEG rhythms [41, 45, 89, 125, 204, 221, 288, 290].

Researchers have also used EEG to accurately forecast performance decline resulting from changes in workload [127, 219], and it has been observed that it is associated with an increase in frontal theta power and alterations in parietal alpha power, which pertains to cognitive and memory performance [119, 124, 126, 274, 330]. Several EEG-based workload indices, such as the ratio of frontal theta to parietal alpha power [118, 154], the ratio of beta to theta and alpha [116], and the theta-beta ratio [235], have been found to accurately reflect workload.

However, EEG features related to mental workload have been found to vary depending on the task, therefore incorporating other modalities like eye activity data and heart rate data can lead to improved outcomes [173, 277]. Pupil size and blink rate have recently gained attention as dependable indicators of workload [216]. Heart rate variability (HRV) is another highly sensitive physiological metric for measuring variations in mental workload [149, 171, 239, 241, 245]. Root mean square of successive differences (RMSSD) is considered the most robust time-domain HRV measure of workload [226].

Once the mental workload of an operator can be accurately measured, it can be utilized to drive a mental workload adaptive system [279, 316]. In such adaptive systems, physiological measures of mental workload can be used to trigger automation that adapts its behavior to the current mental workload of the operator [168, 310]. A mental workload adaptive automation system should be able to adjust to variations in the operator's mental workload without requiring them to explicitly communicate their

needs or activate the automation.

When human operators and automation work together to improve performance and efficiency, the operator expects automation to act like a human colleague [12]. Adaptive automation should be timely, stepping in at the appropriate time, and be cognitively attuned to the operator, helping where needed and taking over tasks that are currently overwhelming the operator.

However, currently, physiological correlates of mental workload are only used to determine when to adapt, not what to adapt, resulting in primitive strategies employed by adaptive automation systems. There is a need to develop intelligent adaptive systems that can identify what form of automation to use depending on the type of mental workload experienced by the operator. However, there is still a lack of evidence that physiological metrics of mental workload can direct the tasks contributing to workload.

4.2 Research Hypothesis

In this study, we investigated whether the use of multimodal physiological metrics of mental workload can provide more information about the tasks that contribute to the workload experienced by an ATC operator. Despite several factors that influence the complexity of ATC tasks [77, 234], such as environmental, display, traffic, and organizational factors, the main functions of an ATC operator are tracking and collision prediction. Therefore, we designed tracking and collision prediction tasks to investigate the physiological effects of workload variations in these basic ATC tasks. The experiment was structured as a classical cognitive paradigm with manipulation of workload (low, medium, high) and repeated stimuli to study whether physiological data such as EEG, eye activity and HRV can accurately assess the mental workload of the operator while they perform these basic tracking and collision prediction tasks. Therefore, we hypothesize the following:

1. The three different levels of workload defined in both tracking and collision prediction tasks will result in a significant performance decline as the workload increases.

2. Variations in workload in tracking and collision prediction tasks can be accurately assessed using EEG, eye activity, and HRV metrics.

3. Performance in tracking and collision prediction tasks can be forecasted based on the measured physiological signals.

4. Physiological responses to workload variations in tracking and collision prediction tasks will differ across tasks.

The chapter organization is shown in Figure 4.1.

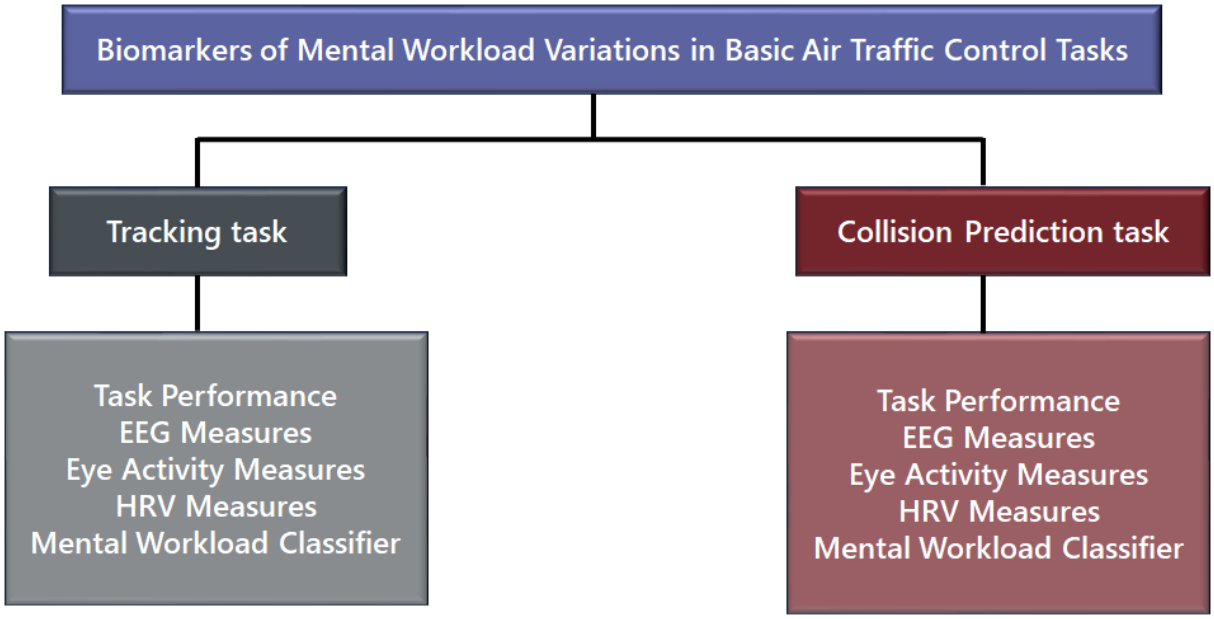

Figure 4.1: Organization of Chapter 4.

4.3 Behavioural and Performance Measures during Workload Variations

In this study, we created two tasks based on Air Traffic Control (ATC) - tracking and collision prediction tasks- as simplified versions. Both tasks are fundamental tasks that ATC operators perform on a regular basis, but we studied them separately to better understand the variations in physiological response to changes in workload in each of these tasks.

In this study, mental workload was evaluated by recording EEG, eye activity, and BVP data while participants completed the tasks. The effects of increasing air traffic on workload were examined by manipulating the mental workload in both tracking and collision prediction tasks by changing the number of dots.

4.3.1 Tracking task

In the tracking task, as the workload increased, tracking accuracy significantly decreased, as shown in Figure 4.2(A). A repeated-measures ANOVA revealed that tracking accuracy varied significantly between workload conditions [F(2, 54) = 239.910, p < .001, η_p^2 = .899]. Tracking accuracy during low workload was significantly higher than in the medium (p < .001) and high workload conditions (p < .001). The tracking accuracy during medium workload was significantly higher than the high workload condition (p < .001). The results showed that performance in the tracking task, as measured by tracking accuracy, decreased significantly as the workload increased.

4.3.2 Collision Prediction task

For the collision prediction task, the time before collision and collision prediction miss proportion rate were taken into account. As the workload increased, the time before collision decreased, as shown in Figure 4.2(B1). A repeated-measures ANOVA was conducted to examine the effect of workload variations on time before collision, and results showed that the time before collision varied significantly between workload conditions [F(1.497, 40.406) = 132.688, p < .001, η_p^2 = .831]. The time before collision decreased significantly in the medium (p < .001) and high (p < .001) workload conditions compared to low workload conditions. The time before collision during medium workload was also significantly greater than during high workload conditions (p = .001).

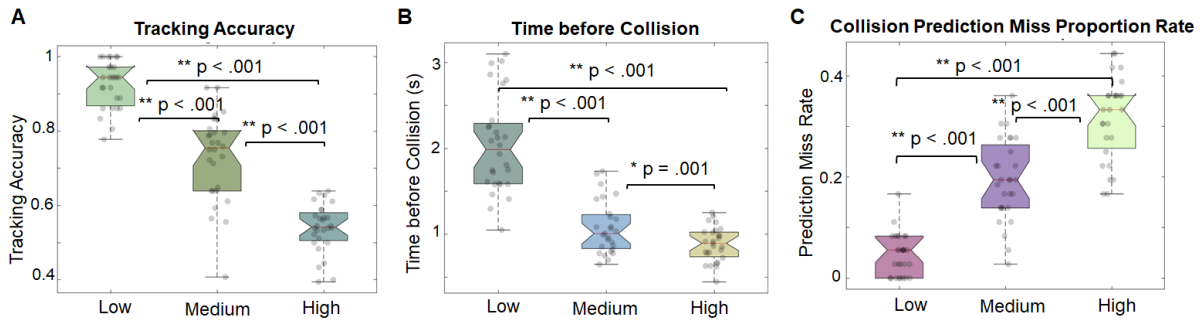

Figure 4.2: A) illustrates the tracking accuracy of all participants in the tracking task for the three levels of workload. (B) illustrates the performance of all participants in the collision prediction task for the three levels of workload. (B1) illustrates the mean time before collision for all participants in the low, medium, and high workload conditions. (B2) illustrates the collision prediction miss proportion rate for the three levels of workload.

As the workload increased, the collision prediction miss proportion rate also increased,

as shown in Figure 4.2(B2). A one-way repeated-measures ANOVA revealed that the collision prediction miss proportion rate varied significantly between workload conditions [$F(1.593, 43.009) = 116.338$, $p < .001$, $\eta_p^2 = .812$]. The prediction miss proportion rate was notably higher in the medium ($p < .001$) and high ($p < .001$) workload conditions compared to the low workload condition. The collision prediction miss proportion rate during high workload was also significantly greater than during medium workload conditions ($p < .001$).

Therefore, for the collision detection task, the time before collision decreased, and the proportion of missed collisions increased significantly as the workload level increased. This suggests that overall performance decreased as the workload level increased in the collision detection task, as participants took longer to identify collisions and were less accurate in identifying collisions when the workload increased.

Therefore, it can be confirmed that the workload manipulation (by varying the number of dots) in both tracking and collision prediction tasks successfully elicited significant variations in performance (H1).

4.4 Independent Brain Source Clusters

The data from all participants were disentangled from the scalp EEG signal through independent component analysis. The frontal, parietal, and occipital clusters were chosen based on the location of fitted dipole sources [259]. For the tracking task (refer Figure 4.2), the Talairach coordinate of the frontal cluster centroid was (-1, 41, 27), the Talairach coordinate of the parietal cluster centroid was (4, -51, 39) and the Talairach coordinate of the occipital cluster centroid was (30, -70, 15).

For the collision prediction task (see Figure 4.3), the Talairach coordinate of the frontal cluster centroid was (-10, 17, 46), the Talairach coordinate of the parietal cluster centroid was (5, -47, 47) and the Talairach coordinate of the occipital cluster centroid was (-3, -69, 20). Significant correlations between mental workload and the spectral powers of frontal, parietal, and occipital clusters were identified.

4.5 EEG Measure of Mental Workload

4.5.1 ERSP Changes with Mental Workload in Tracking Task

Figures 4.3 show the changes in ERSP for the frontal, parietal, and occipital clusters during the tracking task for the three workload conditions: low, medium, and high. A statistical analysis of the ERSP changes of the frontal cluster revealed a significant increase in theta power from the low level to the high level ($p < .05$) and a significant increase in theta power at the frontal cluster during the high workload condition as compared to the medium workload condition. The frontal theta power was significantly greater than the low workload condition, as shown in Figure 4.3.

However, there were no significant variations in spectral power observed at the parietal cluster. The ERSP changes at the occipital cluster revealed a significant decrease in alpha power from the low level to the high level ($p < .05$) and a significant decrease in alpha power at the occipital cluster during the high workload conditions compared to the medium workload condition. The occipital alpha power was significantly lower than the low workload condition.

The tracking task required participants to allocate attentional resources to keep track of one, three, or five dots moving randomly among distractor dots. Working memory load is known to increase with the allocation of attentional resources, which is reflected by an increase in frontal theta power [126, 182, 183]. The tracking task resulted in an increase in frontal theta power, indicating that an increased working memory load was experienced with increasing workload levels.

The task also required working memory mechanisms related to maintaining relevant items, which is reflected by a decrease in alpha power [48, 126, 282, 363]. Alpha power is also known to decrease with an increase in memory load [104, 105, 114, 304, 324], and task difficulty [260, 330]. The study found that occipital alpha power decreased with increasing workload levels in the tracking task, supporting the presence of this working memory mechanism.

4.5.2 ERSP Changes with Mental Workload in the Collision Prediction task

The changes in ERSP (Event-Related Spectral Perturbation) at the frontal, parietal, and occipital clusters during different workload conditions in the collision prediction task are illustrated in Figure 4.4. A statistical analysis of the ERSP changes at the

frontal cluster revealed a significant increase in theta power during high workload conditions compared to low workload conditions. Additionally, the frontal power during high workload conditions was found to be significantly greater than that during medium workload conditions.

The ERSP changes at the parietal cluster showed a significant increase in theta power during high workload conditions compared to low workload conditions ($p < .05$) and a significant decrease in alpha power ($p < .05$). There was also a significant increase in theta power and a decrease in alpha power at the parietal cluster during high workload conditions compared to medium workload conditions. In the medium workload condition, the parietal theta power was found to be significantly higher than in the low workload condition, while the parietal alpha power was significantly lower than that in the low workload condition.

The results of the ERSP analysis on the occipital cluster in the collision prediction task revealed a significant increase in delta and theta power during high workload conditions when compared to low workload conditions ($p < .05$). This is illustrated in Figure 4.4, which also shows that there was a significant increase in delta and theta power at the occipital cluster during high workload conditions as compared to medium workload conditions. Additionally, it was found that in medium workload conditions, the delta and theta power at the occipital cluster were significantly higher than in low workload conditions.

In the collision prediction task, participants were required to anticipate the trajectory of the dots and predict whether they would collide, which demands attention and internal focus. Delta power is a marker of attention or internal focus in mental tasks, and it is known to increase with an increase in workload [140, 330, 363]. The results of the study showed an increase in delta power at the occipital sites, indicating that the participants allocated more attentional resources with increasing workload levels in the collision prediction task.

Additionally, keeping track of the trajectory of six, 12, or 18 dots added to the memory load of the participants. Studies have shown that theta power is associated with memory load [160, 163] and working memory capacity [182, 183, 308]. The study found a significant increase in theta power at the frontal, parietal, and occipital clusters in the collision prediction task, confirming an increase in memory load with increasing workload levels.

In addition, our results show that as the workload level increases, there is an increase in parietal alpha power. This increase in alpha power with increasing workload levels is

related to maintaining relevant items in working memory [51, 126, 282, 330, 363] and is known to decrease with an increase in memory load [104, 105, 114, 304, 324] and task difficulty [260, 330].

However, in the collision prediction task, the greatest decrease in parietal alpha power was observed just before the collision, which may be related to an increase in the time pressure [321] experienced by the participants as they try to identify and click on the colliding pair of dots before the collision occurs.

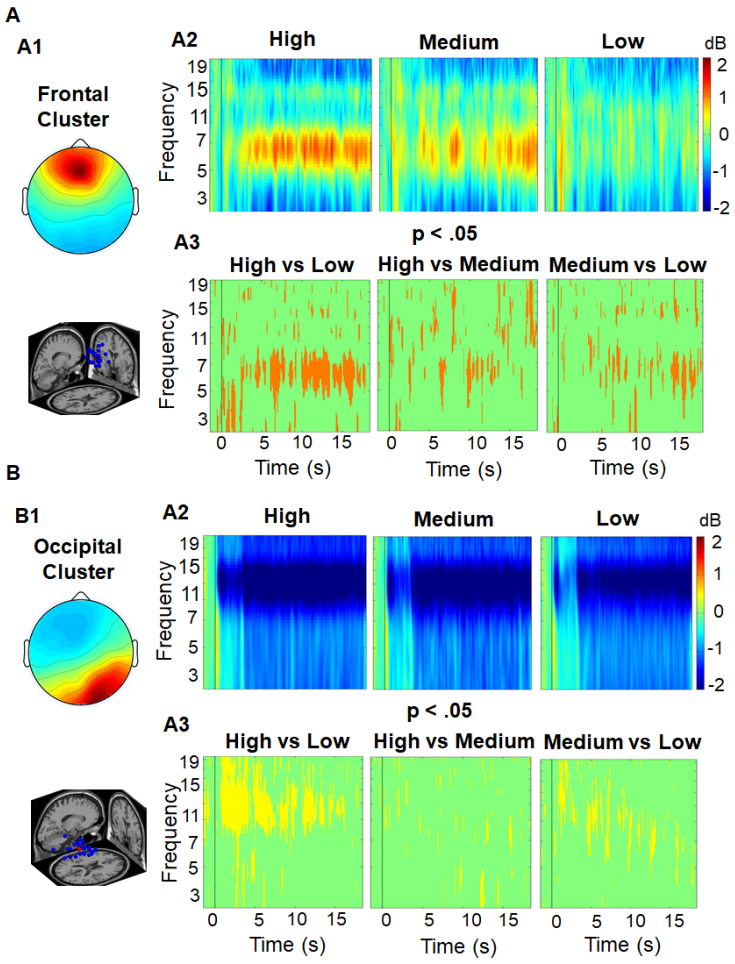

Figure 4.3: The results of the ERSP for the tracking task show the changes in the spectral power of the frontal [Talairach coordinate: (-1, 41, 27)] and occipital [Talairach coordinate: (30, -70, 15)] clusters selected for the tracking task, including spatial scalp maps and dipole source locations, for the high, medium, and low workload conditions of the task.

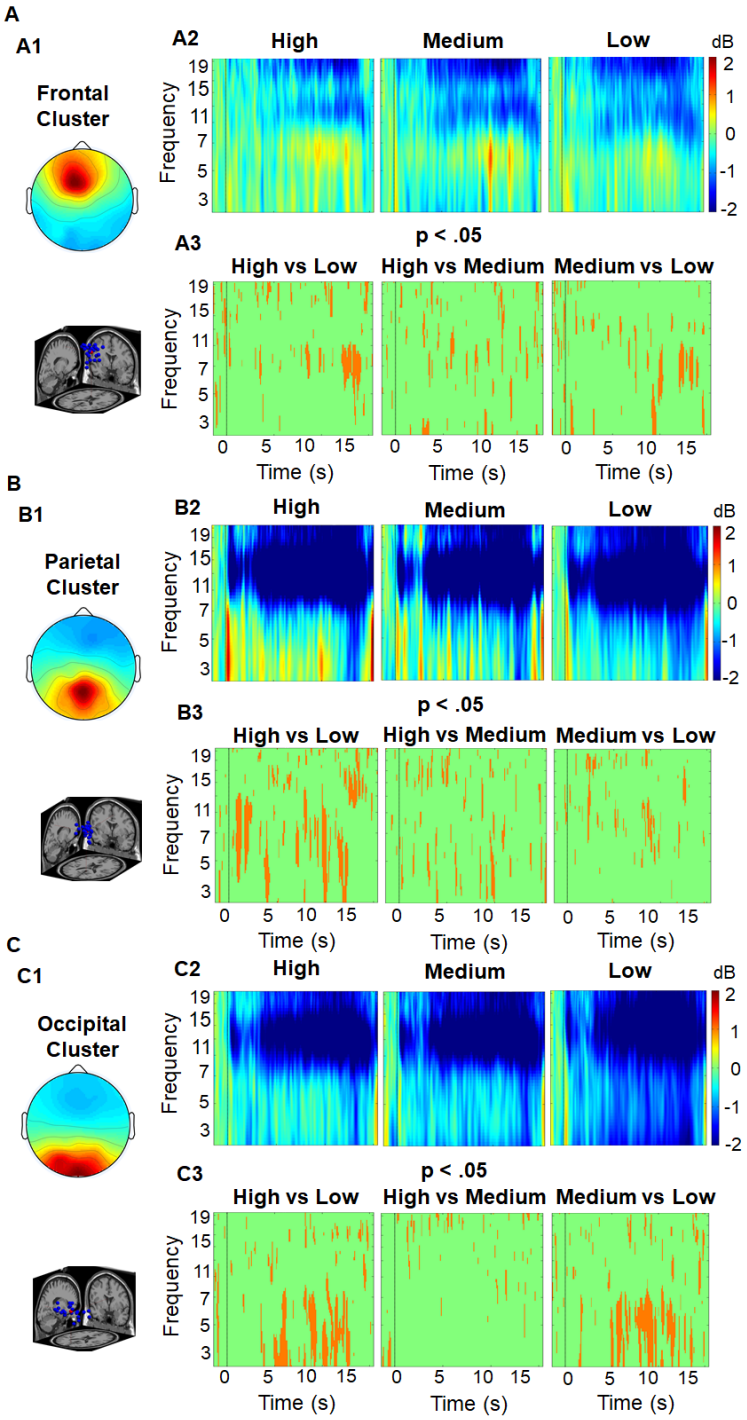

Figure 4.4: The results of the ERSP for the collision prediction task illustrate the changes in the spectral power of the frontal [Talairach coordinate: (-10, 17, 46)], parietal [Talairach coordinate: (5, -47, 47)], and occipital [Talairach Coordinate: (-3, -69, 20)] clusters selected for the task, including spatial scalp maps and dipole source locations, for the high, medium, and low workload conditions of the task.

4.5.3 Power Spectral Density Changes with Mental Workload

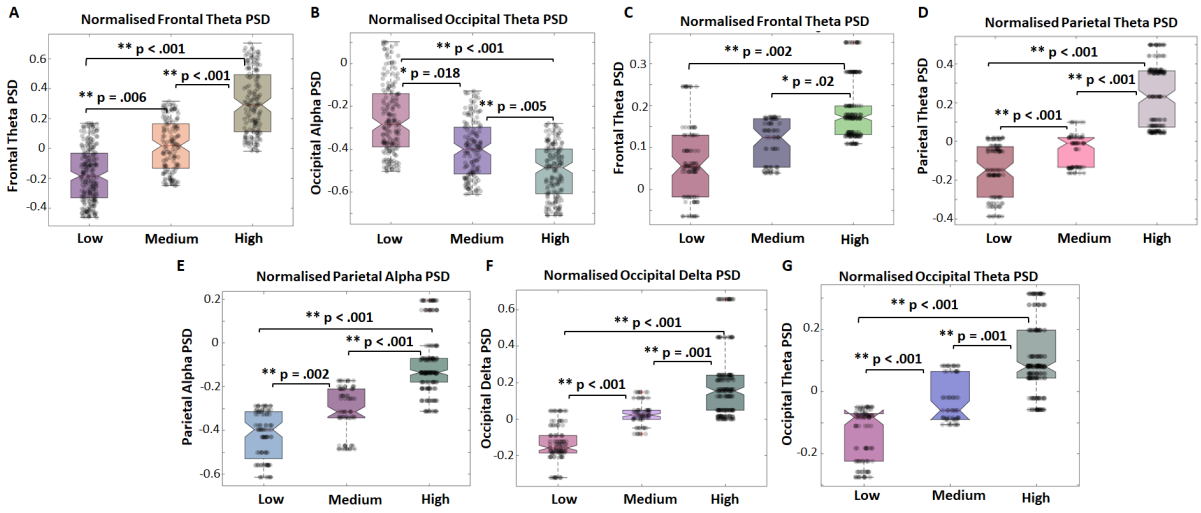

Figure 4.5: Figure demonstrates the power spectral density during tracking and collision prediction tasks. (A) shows the normalized frontal theta PSD for low, medium, and high workload conditions. (B) illustrates the normalized occipital alpha PSD for the tracking task, separated by workload level. (C) displays the average frontal theta PSD across low, medium, and high workload conditions of the collision prediction task. (D) shows the average parietal theta PSD for the three workload levels. (E) illustrates the average parietal alpha power for different workload conditions, and (F) displays the average occipital delta PSD for low, medium, and high workload conditions. (G) presents the average occipital theta PSD for the three workload levels in the collision prediction task.

Figure 4.5 shows the changes in the spectral power of the frontal cluster's ICs during the tracking task for different workload levels. The results of the statistical analysis indicate that there was a significant difference in the frontal theta power across the three workload conditions [F(2, 46) = 50.931, p < .001, η_p^2 = .822]. Specifically, the frontal theta power was found to be higher during high workload as compared to low (p < .001) and medium workloads (p < .001), and higher during medium workload as compared to low workload (p = .006).

A one-way repeated-measures ANOVA revealed a significant difference in occipital alpha power across various workload conditions [F(2, 46) = 24.780, p < .001, η_p^2 = .693]. The analysis indicated that occipital alpha power spectral density was found to be significantly lower during high workload conditions in comparison to both low workload (p < .001) and medium workload (p = .005). Additionally, medium workload conditions

displayed a statistically significant decrease in occipital alpha power spectral density as compared to low workload conditions (p = .018).

The frontal cluster's ICs demonstrated varying theta power levels in response to different workload conditions, as determined by a one-way repeated-measures ANOVA [F(2, 46) = 8.570, p = .001, η_p^2 = .271]. Specifically, theta power in the frontal cluster was found to be higher during high workload compared to low workload (p = .002), but no statistically significant difference was observed between high and medium workload (p = .051). Additionally, there was no statistically significant change in theta power in the frontal cluster between the medium and low workload conditions (p = .336). On the other hand, the parietal cluster's ICs exhibited a significant increase in theta power and a significant decrease in alpha power across all workload conditions (low, medium, and high).

The results of a one-way repeated-measures ANOVA revealed that there were significant variations in parietal theta [F(2, 46) = 47.764, p < .001, η_p^2 = .675] and alpha [F(2, 46) = 38.639, p < .001, η_p^2 = .627] power levels across different workload conditions. Specifically, the parietal theta power during low workload was significantly lower than the medium (p < .001) and high workload conditions (p < .001). Additionally, the medium workload had a lower parietal theta power level than the high workload condition (p < .001). Furthermore, there was a significant decrease in parietal alpha power during the medium (p = .002) and high workload conditions (p < .001) compared to the low workload condition. The parietal alpha power level was also found to be significantly lower during the high workload condition when compared to the medium workload (p < .001).

The results of a one-way repeated-measures ANOVA indicated that there were significant variations in occipital delta [F(1.563, 35.951) = 35.321, p < .001, η_p^2 = .606] and theta [F(2, 46) = 39.101, p < .001, η_p^2 = .630] power levels across different workload conditions. Specifically, the occipital delta power was significantly higher during the medium (p < .001) and high workload (p < .001) conditions when compared to the low workload. Additionally, the occipital delta power during the high workload condition was found to be significantly higher than during the medium workload condition (p = .001). Furthermore, the occipital theta power during the low workload was significantly lower than during the medium (p < .001) and high workload conditions (p < .001), and the medium workload had a lower occipital theta power level than the high workload condition (p = .001).

4.6 Eye Activity changes with Mental Workload

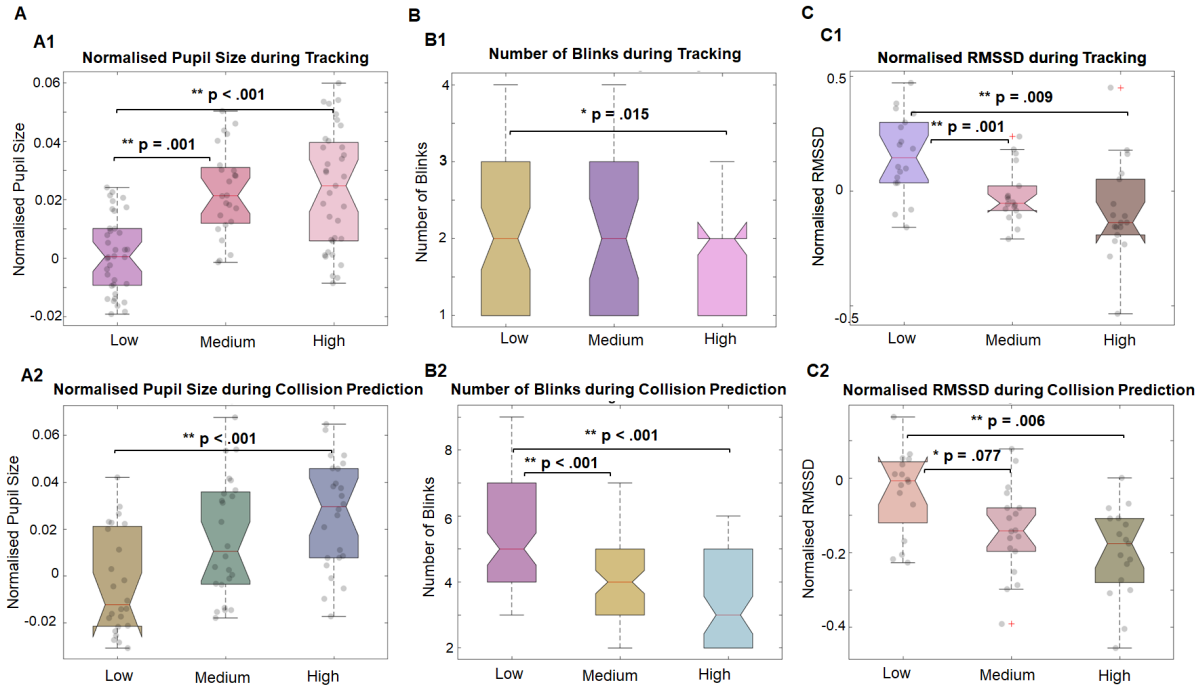

Figure 4.6: The eye activity and HRV of participants observed to change during both tracking and collision prediction tasks. (A) illustrates the normalized pupil size of all participants, which shows a positive trend with increasing workload. (A1) shows the normalized pupil size during the low, medium, and high workload conditions of the tracking task, and (A2) shows the normalized pupil size during the low, medium, and high workload conditions of the collision prediction task. (B) displays the negative trend in the number of blinks as workload increases. (B1) shows the number of blinks during different workload conditions of the tracking task, and (B2) shows the number of blinks during the collision prediction task, which decreases as the workload level increases. (C) illustrates the declining trend in the normalized RMSSD of all participants as workload increases. (C1) shows the normalized RMSSD of all participants during the low, medium, and high workload conditions of the tracking task, and (C2) shows the normalized RMSSD during the collision prediction task for the three levels of workload.

We also examined eye-related metrics during changes in workload. Eye activity data was transformed into information on pupil size and blink rate. As displayed in Figure 4.6(A), the size of the pupils increased as the workload increased for both tracking and collision prediction tasks. For the tracking task, a one-way repeated-measures ANOVA [$F(2, 38) = 13.205$, $p < .001$, $\eta_p^2 = .410$] revealed that there were significant changes in pupil size for different workload conditions.

Specifically, there was a significant increase in pupil size for the medium (p = .0001) and high workload conditions (p = .001) when compared to the low workload condition. However, the increase in pupil size for the high workload condition was not found to be statistically significant when compared to the medium workload condition (p = .313) in the tracking task.

The data from a one-way repeated measures ANOVA demonstrates that during the collision prediction task, there were significant variations in pupil size across different workload conditions [F(2, 46) = 9.276, p < .001, η_p^2 = .287]. Specifically, the pupil size was found to be significantly larger for the medium (p = .011) and high workload conditions (p < .001) when compared to the low workload condition. However, no significant difference in pupil size was observed between the high and medium workload conditions (p = .180).

The number of blinks during tracking and collision prediction tasks decreased with the increasing workload, as shown in Figure 4.6(B). One-way repeated-measure ANOVA was conducted to study the effect of workload variations on the number of blinks, which revealed significant variations in the number of blinks during the tracking task for different workload conditions [F(2, 46) = 3.624, p = .035, η_p^2 = .136]. The number of blinks in the low workload condition of the tracking task was significantly greater than that of the high workload condition (p = .015) but not significantly greater than that of the medium workload (p = .328). There was no significant decrease in the number of blinks in the high workload condition as compared to the medium workload (p = .106).

The effect of workload on the number of blinks in the collision prediction task was analyzed using one-way repeated-measure ANOVA. It showed a significant variation in the number of blinks [F(2, 46) = 18.586, p < .001, η_p^2 = .447]. In the low workload condition, the number of blinks was significantly greater in the medium (p < .001) and the high workload conditions (p < .001). However, the number of blinks in the medium workload condition was not significantly higher when compared to the high workload condition (p = .604).

Pupil size significantly increased with the increasing workload in both tracking and collision prediction tasks, and the number of blinks also decreased significantly with the increasing workload in both tasks. Pupil size is a reliable measure of workload [28, 213, 216, 217] as it dilates with increasing workload [26, 33, 174, 189, 215, 338, 351]. Research has shown that blink inhibition occurs during high workload [295], and blink rate is inversely related to attentional levels, and workload experienced by the operator [41, 45, 344, 350, 361, 363]. The results are summarized in Table 4.1.

Table 4.1: PSD Changes for Tracking and Collision Prediction Task

Brain Region	Tracking Task	Collision Prediction Task
Frontal	theta PSD ↑ with workload	theta PSD ↑ with workload
Parietal		theta and alpha PSD ↑ with workload
Occipital	alpha PSD ↓ with workload	delta and theta PSD ↑ with workload

4.7 Heart Rate Variability (RMSSD) changes with Mental Workload

We also examined HRV metrics during changes in workload. Figure 4.6(C) illustrates the variation in the RMSSD for various workload conditions in tasks related to tracking and collision prediction. Analysis using one-way repeated-measures ANOVA revealed a significant change in the RMSSD for different workload conditions in the tracking task [F(2, 34) = 10.171, p < .001, η_p^2 = .374], with a decrease observed in the RMSSD for medium (p = .001) and high workload conditions (p = .009) compared to the low workload condition. However, there was no significant difference in the RMSSD for medium and high workload conditions in the tracking task (p = .440).

Similarly, in the collision prediction task, the RMSSD for low workload condition was found to be greater than the medium (p = .001) and high workload conditions (p = .009), as shown by the one-way repeated-measures ANOVA [F(2, 44) = 4.279, p = .022, η_p^2 = .201]. No significant variation in the RMSSD for medium and high workload conditions (p = .326) was observed in the collision prediction task.

The results showed that RMSSD was negatively correlated with the mental workload in both tasks. This decrease in RMSSD with the increasing workload is a well-established finding in the literature [66, 67, 108, 144, 226, 268, 335]. This behavior, eye activity, and HRV results for tracking and collision prediction tasks are summarized in Table **??**.

4.8 Multiple Regression to Predict Task Performance

4.8.1 Tracking Task

A multiple regression analysis was conducted to determine if EEG, eye activity, and HRV metrics of workload could effectively predict performance in the tracking task. The

Table 4.2: Task Performance, Eye Activity and HRV for Tracking and Collision Prediction Task

	Tracking Task	Collision Prediction Task
Task Performance	↓ with workload	↓ with workload
Pupil Size	↑ with workload	↑ with workload
Blink Rate	↓ with workload	↓ with workload
RMSSD	↓ with workload	↓ with workload

results of the regression analysis showed that the model explained 54.3% of the variance and that the model was a significant predictor of tracking performance, as determined by F(3, 67) = 26.543, p < .001. The analysis found that while EEG metrics (B = .067, p = .001) and eye activity (B = -.089, p < .001) had a significant impact on the model, HRV metrics did not (B = -.152, p = .125). The final predictive model is:

$$P\_t = 0.725 - 0.067 * EEG - 0.089 * Eye - 0.152 * HRV \tag{4.1}$$

where P_t is the tracking performance, EEG is the EEG metrics, Eye is the eye metrics and HRV is HRV measures.

4.8.2 Collision Prediction

Multiple regression analysis was performed to investigate if EEG, eye activity, and HRV metrics could effectively predict performance in the collision prediction task. The results of the analysis showed that the model explained 61.7% of the variance and that the model was a significant predictor of performance in the collision prediction task, as determined by F(3, 68) = 24.324, p < .001. The analysis found that while eye activity (B = -.276, p = .02) and EEG metrics (B = -.532, p < .001) had a significant impact on the model, HRV metrics did not (B = .444, p = .443). The final predictive model is:

$$P\_cp = 0.055 - 0.532 * EEG - 0.276 * Eye + 0.444 * HRV \tag{4.2}$$

where P_cp is the performance in the collision prediction task, EEG is the EEG metrics, Eye is the eye metrics and HRV is HRV measures.

Our findings indicate that EEG power spectra in the frontal, parietal, and occipital regions, as well as eye activity and HRV metrics, can effectively and accurately evaluate the mental workload of the participants in both tasks. Therefore, our second hypothesis

Table 4.3: Tracking Task Workload Classification

Model	Accuracy	F1-score	coh-kappa
EEGNet	62.8	0.53	0.34
Shallow ConvNet	60.1	0.50	0.28
DeepConvNet	61.04	0.52	0.31

(H2) is confirmed for both tracking and collision prediction tasks. Our third hypothesis (H3) is also supported, as the results of the multiple regression analysis showed that performance in the tracking and collision prediction tasks can be predicted using EEG, eye-related, and HRV metrics.

Although EEG, eye activity, and HRV measures were able to distinguish between low and high levels of workload effectively, some of these measures were not able to accurately differentiate medium workload from low/high workload conditions. Two possible reasons for this inconsistency reported in the literature are issues with experimental design [192] or inter-individual differences [27, 29, 340]. In our experimental design, the medium workload condition might have required nearly similar cognitive resources and, thus, did not show a significant variation from the low/high workload condition. However, our results still showed a significant decrease in performance with increasing workload levels in both the tracking and collision prediction tasks.

Therefore, it is more likely that the inconsistency is caused by inter-individual differences. It is known that the correlation between workload and task demand is not straightforward [19, 59]. Sperandio [326] argues that the relationship can be better understood by exploring the strategies that human operators use to manage their cognitive resources and workload, and many researchers concur with this perspective [19, 20, 76, 146, 209]. Different participants might reflect workload variations differently based on their cognitive resources and the strategies that they use to perform the task.

4.9 Workload Classification Results

The workload classification results for the tracking task are shown in Table 4.3.

The results demonstrate that EEGNet was most efficient in classifying workload in the tracking task. Shallow ConvNet and Deep ConvNet demonstrated similar performance.

Table 4.4: Collision Prediction Task Workload Classification

Model	Accuracy	F1-score	coh-kappa
EEGNet	61.4	0.49	0.31
Shallow ConvNet	67.9	0.56	0.36
DeepConvNet	68.13	0.56	0.38

The workload classification results for the collision prediction task are shown in Table 4.4.

The results demonstrate that Deep ConvNet was most efficient in classifying workload in collision prediction task. Shallow ConvNet also demonstrated comparable performance in the classification.

4.10 Implications for Future Studies

Our results also suggest that even though eye activity and HRV metrics are sensitive to variations in task load, they may not provide any useful information about the specific task that causes the workload variations. However, EEG measures were found to not only be sensitive to workload variations but also to the type of task. The increased workload in the tracking task was reflected by an increase in frontal theta power and a decrease in occipital alpha power. No significant changes were observed in the parietal theta, alpha, occipital delta, or theta power with the increasing workload in the tracking task. In the collision prediction task, the increase in workload was correlated with increases in frontal theta, parietal theta, occipital delta, and theta power and a decrease in parietal alpha power. No significant variation was observed in occipital alpha power during the collision prediction task.

The neurometrics correlated with the variations in workload for the tracking and collision prediction tasks are different, which supports our fourth hypothesis (H4) is true. This means that neurometrics can help identify the specific task that is contributing to an increase in workload in complex ATC environments at a given moment and can inform strategies that can be employed by the workload adaptive system to reduce this increase. These findings provide evidence that the use of EEG measures in a closed-loop adaptive system can not only aid in determining "when" to deploy automation but also "what" type of automation to use to mitigate workload variations for operators. Therefore, the results presented here contribute to the development of adaptive strategies that are

crucial for the design of intelligent closed-loop mental workload adaptive ATC systems.

4.11 Limitations

While we have examined the effects of workload variations systematically, the main limitation of this study is that different variables were controlled to highlight the impact of workload variations based on variations in traffic load in basic ATC tasks. Our experimental scenario was not entirely realistic as several environmental factors contribute to the workload experienced by ATC operators even while performing these basic tasks. Additionally, a prior gaming experience can influence the strategies employed by participants and thus, significantly affect the experienced workload. We did not study inter-individual differences between participants in this study.

4.12 Summary

The performance and efficiency of a system can be enhanced by keeping the operator's workload within an optimal range. In order to understand the impact of variations in basic task load that make up the variations in complex ATC tasks, we separately designed two basic ATC tasks: tracking and collision prediction tasks. We successfully uncovered EEG spectral power, eye, and HRV correlates of mental workload variations for tracking and collision prediction tasks of air traffic controllers, providing a comprehensive understanding of the workload demands in ATC tasks.

Our results show that EEG, eye and HRV metrics can offer a sensitive and reliable way to predict the mental workload and performance of the operator. The neural response differences to increased workload in the tracking and collision prediction tasks indicate that these neural measures are sensitive to variations and types of mental workload and have the potential to aid in not just determining 'when' to adapt, but also 'what' to adapt, helping to create intelligent closed-loop mental workload aware systems. The investigation of physiological indicators of workload variation in basic ATC tasks has practical applications for designing future adaptive systems that incorporate neurometrics in determining the type of automation to use to reduce variations in workload in complex ATC systems.

CHAPTER 5

NEURAL CORRELATES OF WORKLOAD IN PHYSICAL HUMAN ROBOT COLLABORATION

5.1 Background

The emergence of the fifth industrial revolution has led to a growing presence of robots in human environments. These technological advancements have enabled physical interactions between human users and robots. Despite significant progress in robotics, robots must be equipped with the ability to understand human states in order to truly work in partnership with human users.

Mental workload is a key factor in determining the state of human users as they interact closely with robots, moving from supervisory to collaborative roles. Challenging tasks can not only restrict access to cognitive resources [43] but can also completely drain resources [261] required for task completion, thus affecting performance. Keeping human users' mental workload at an optimal level is ideal, avoiding underload conditions that lead to boredom and overwhelming overload conditions [41, 48, 138]. These negative effects of workload on human performance and situational awareness can raise concerns about safety in physical interactions with robots, negatively impacting the efficiency of the interaction [368].

Providing robots with a means of communication that accurately predicts a human's level of mental workload can aid in adapting strategies to enhance human productivity

and prevent loss of awareness [12, 46, 309]. To gauge mental workload, researchers commonly use self-evaluation, performance-based, or physiological methods. As mental workload is a complex, individual concept reflecting cognitive resources, many researchers use a combination of measures as each has its own strengths and limitations [254].

Physiological measures offer a precise and dependable way to predict mental workload, as well as a non-invasive assessment with high sensitivity and diagnostic ability [266, 380]. They have been used in rehabilitation robotics to appropriately challenge users and aid in their recovery [22, 132, 184, 250, 251]. However, it should be noted that these measures can also be affected by physical workload. In order to accurately measure mental workload during physical interaction with robots, a reliable physiological measure should not be impacted by physical activity [251]. One potential solution is to focus on brain dynamics, as the brain integrates various sensory and motor inputs to perceive the environment and perform tasks, thus potentially isolating and extracting mental workload correlations without any influence of physical activity.

Electroencephalography (EEG) is an effective method for measuring the ongoing electrical activity of the brain without restricting movement, due to its portability and high temporal resolution [211]. EEG can provide insight into a user's overall workload level, as well as the variation of workload during a task, not just after its completion. In recent years, researchers have used EEG measures to study the correlation between motor control difficulty and perceived workload in haptic interactions with robots [228, 229]. However, it's important to note that haptic interactions with robots may not involve significant physical activity, and the relationship between mental workload and physical activity in these interactions may differ from other types of physical interactions with robots.

5.2 Research Hypothesis

Physical activity that is demanding can not only decrease the quality of the EEG signal but also make it difficult to extract biomarkers for mental workload, as physical activity also requires mental effort. This work investigates Mobile Brain/Body Imaging methods [211] to analyze EEG features that reflect workload variation and are resistant to movement artifacts. Data-driven approaches, such as Independent Component Analysis [31], have shown the ability to separate EEG data into brain and artifact sources [211]. Researchers have found that theta power in the frontal region of the brain is related to working memory capacity [166, 182, 183, 308] and memory load [160], while alpha

power in the parietal region indicates task difficulty [330] and memory load [105].

Additionally, sensorimotor processing is reflected in the alpha and beta power in the central region of the brain [7, 134, 274]. However, the effects of physical activity on mental workload are a result of the additional mental resources required for sensorimotor integration and workload [153, 190, 225, 232, 312]. Despite this, there is no agreement on how mental workload affects physical performance [79, 153, 212, 345]. An increase in mental workload may negatively impact the physical performance of the human partner in interaction and compete with sensorimotor resources.

Previous studies have generally evaluated workload in laboratory settings with limited physical activity during the experiment, which does not reflect real-world conditions. To address this, we designed our task as a clock game with a blasting task that mimics the operation of an abrasive robot. We varied the mental workload by using arithmetic tasks with varying difficulty levels (low, medium, and high) while keeping the physical load constant to investigate the impact of mental workload on physical performance.

The experiment was designed as an engaging clock game to assess the reliability of EEG signals in measuring an operator's mental workload during active physical interaction with a robot. To further validate the results, we included a subjective measure of workload, the NASA Task Load Index (NASA TLX) questionnaire [141]. We hypothesize that as the workload increases during the blasting task, there will be:

1. an increase in frontal theta power and a decrease in parietal alpha power,
2. changes in alpha and beta power at the central region of the brain and
3. a decline in physical performance despite keeping physical load consistent across the three levels of mental workload.

The chapter organization is shown in Figure 5.1.

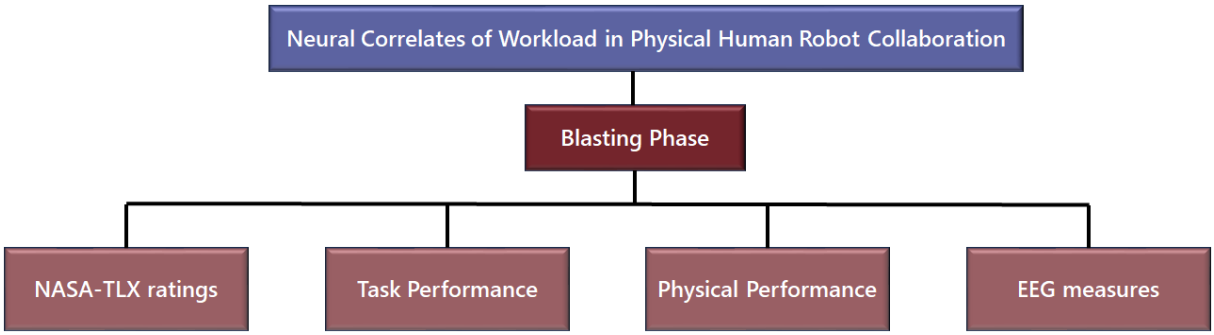

Figure 5.1: Organization of Chapter 5.

5.3 Subjective Workload Measures

The ratings for subjective workload increased as the workload in the clock game task increased, as shown in Figure 5.2. The scores for mental demand also significantly increased with the workload, as shown by statistical analysis [F(2, 29) = 1754.972, p < .01, η_p^2 = .992]. There was a significant increase between the low and medium (p < .01), low and high (p < .01), and medium and high (p < .01) conditions. Interestingly, the scores for physical demand also increased with the workload [F(2, 29) = 42.370, p < .01, η_p^2 = .745]. There was a significant increase between low and high (p < .01) and medium and high (p < .01) conditions, but no significant difference between low and medium (p = .573) conditions.

The scores for temporal demand also significantly increased with workload [F(2, 29) = 734.805, p < .01, η_p^2 = .981]. The temporal load scores showed a significant increase for low-medium (p < .01), low and high (p < .01), and medium and high (p < .01) conditions. Additionally, the scores for performance decreased significantly with workload [F(2, 29) = 1392.381, p < .01, η_p^2 = .990], with a significant decrease in high-low (p < .01), high-medium (p < .01) and medium-low conditions (p < .01).

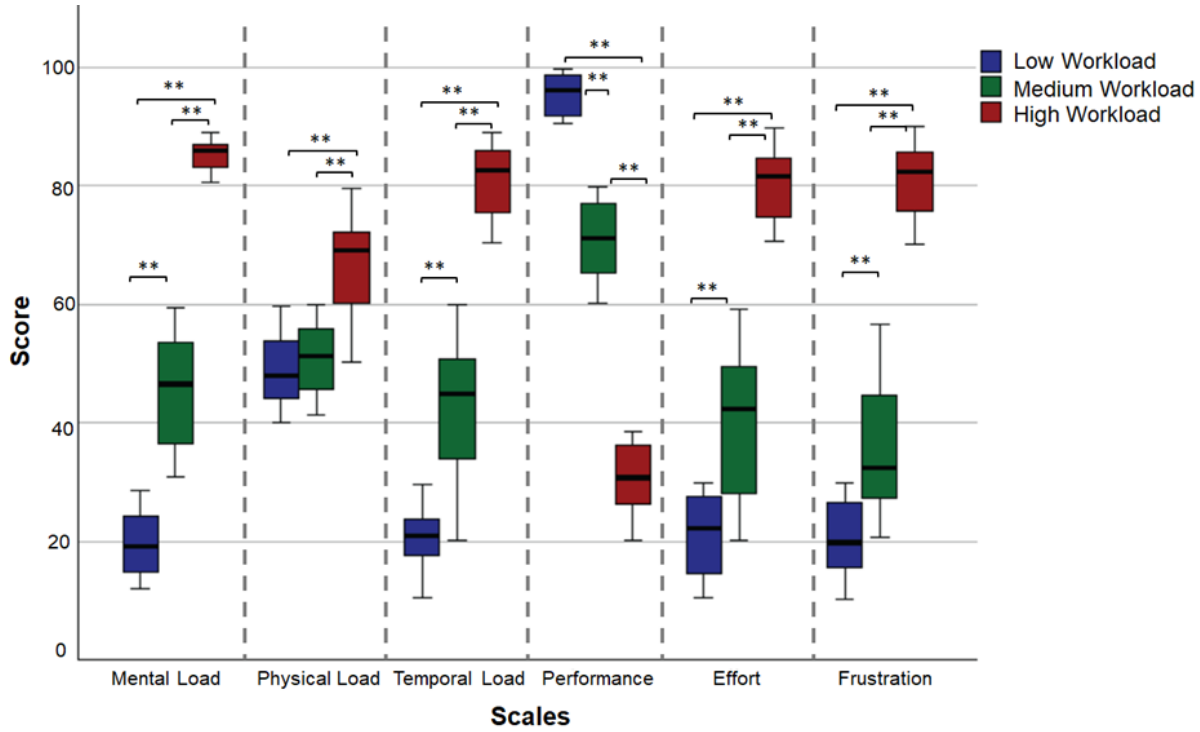

Figure 5.2: The NASA-TLX scores for all participants during the blasting task for the three workload levels (** denotes p < .01).

Table 5.1: NASA-TLX ratings for the Blasting task

Parameter	Trend with Increasing Workload
Mental Load	Increased
Physical Load	Increased
Temporal Load	Increased
Performance	Decreased
Effort	Increased
Frustration	Increased

The scores for effort increased as the workload increased, [$F(2, 29) = 783.922$, $p < .01$, $\eta_p^2 = .982$], with significant increases for low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p < .01$) conditions. Additionally, the scores for frustration also significantly increased with the workload, [$F(2, 29) = 1382.903$, $p < .01$, $\eta_p^2 = .990$], with significant increases for low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p = .01$) conditions. The results are also summarized in Table 5.1.

5.4 Arithmetic and Blasting task Performance

The mental workload in the blasting task was manipulated by a dual arithmetic task of varying difficulty.

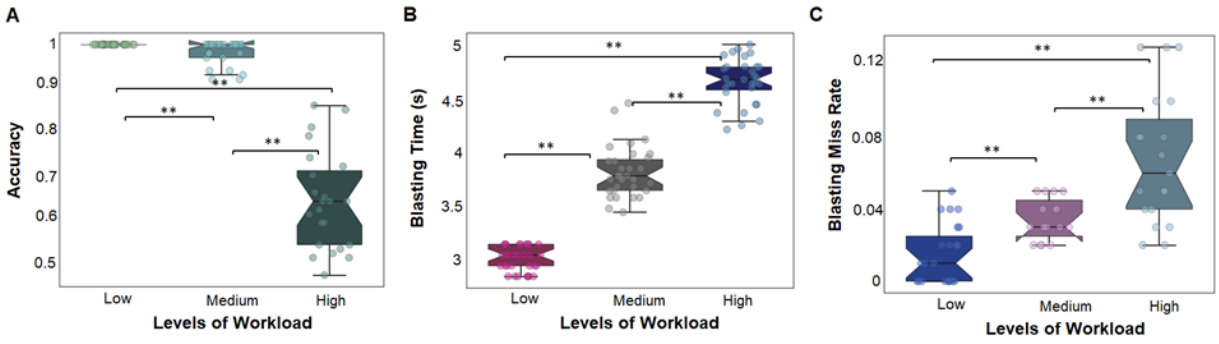

Figure 5.3: Blasting task performance results (A) illustrates the accuracy of identifying the target, (B) illustrates the time taken for blasting, and (C) illustrates the rate of missed blasting for the three workload levels. (** denotes $p < .01$).

As the workload increased, the target identification accuracy decreased significantly, [$F(2, 24) = 8.023$, $p = .002$, $\eta_p^2 = .799$], as demonstrated by Figure 5.3(A). This decrease was significant for all conditions, including low and medium ($p = .007$), low and high ($p < .01$), and medium and high ($p < .01$). Additionally, the blasting time also increased with

increasing workload [F(2, 29) = 58.349, p < .01, η_p^2 = .801], as shown in Figure 5.3(B), and this increase was significant for all conditions (p < .01). Furthermore, the blasting miss rate also increased with increasing workload, as shown in Figure 5.3(C), and this increase was significant for all conditions.

Performance in the blasting task, assessed by the target identification accuracy, blasting time and blasting miss rate, degraded notably with the rising workload. Therefore, the workload manipulation during blasting successfully evoked substantial performance variations.

5.5 Physical Performance

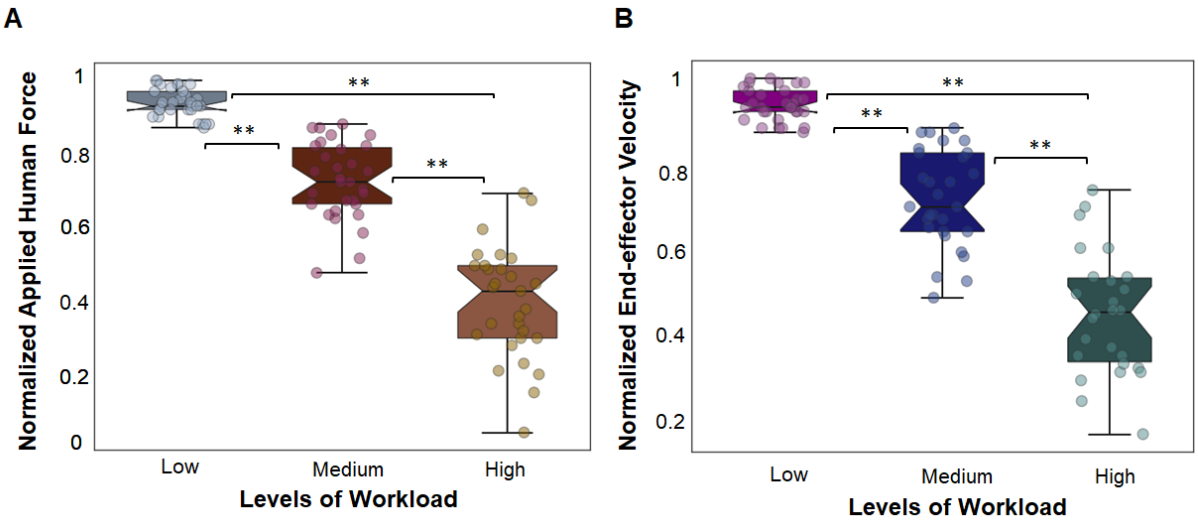

Figure 5.4: (A) illustrates the participants' applied human force and (B) the velocity of the end-effector for the three workload conditions of low, medium, and high. ** denotes p < .01.

As the workload increased, the amount of human force applied decreased significantly, [F(2, 27) = 199.359, p < .01, η_p^2 = .937], as shown in Figure 5.4(A). This decrease in applied force was significant for all conditions, including low and medium (p < .01), low and high (p < .01), and medium and high (p < .01). Additionally, the velocity of the end-effector also decreased with increasing workload [F(2, 24) = 123.490, p < .01, η_p^2 = .911], as demonstrated by Figure 5.4(B), and this decrease was significant for all conditions (p < .01).

The physical performance assessed by the applied human force and the resultant end-effector velocity decreased significantly with increasing workload levels. Hence, the increasing mental workload resulted in significant physical performance deterioration, proving our third hypothesis (H3).

5.6 EEG Measures of Workload during Physical Collaboration with ANBOT

In order to investigate the EEG correlates of mental workload variations during blasting, the independent components were extricated from the scalp EEG data, and significant correlations between the task load variations and the spectral powers at frontal, central and parietal clusters were observed.

5.6.1 Independent Source Clusters

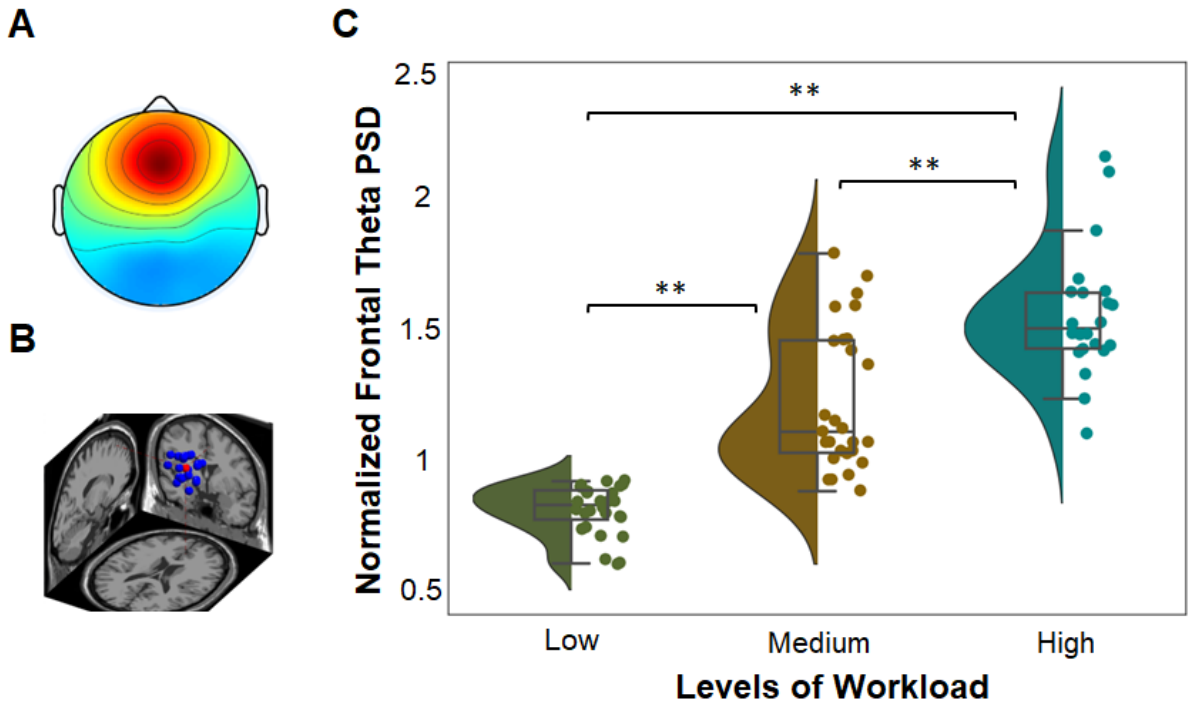

Figure 5.5: Frontal Cluster, (A) shows a scalp map, (B) displays the locations of the dipoles for the components. The centroid of the cluster in the Talairach Coordinate system is located at (4, 38, 40), and (C) presents the normalized power spectral density of theta at the ICs selected in the Frontal cluster during blasting (** denotes $p < .01$).

The clusters in the frontal, central, and parietal regions were identified based on the locations of the dipoles [259]. The centroid of the frontal cluster in the Talairach Coordinate system was located at (4, 38, 40) and consisted of 19 components from 19 subjects. The centroid of the central cluster was located at (3, -5, 53) and had 18 components from 18 subjects. The centroid of the parietal cluster was located at (11, -49, 47) and included 24 components from 24 subjects.

5.6.2 Frontal Cluster

Figure 5.5 shows the power spectral density of the ICs in the frontal cluster for different workload conditions during blasting. The theta power at the ICs in the frontal cluster increased with increasing workload, as demonstrated by [F(2, 20) = 100.170, p < .01, η_p^2 = .909], and this increase was significant for all conditions, including low and high (p < .01), low and medium (p < .01), and medium and high (p < .01). However, there were no significant differences observed in the delta, alpha, beta, or gamma power spectral density of the ICs in the frontal cluster.

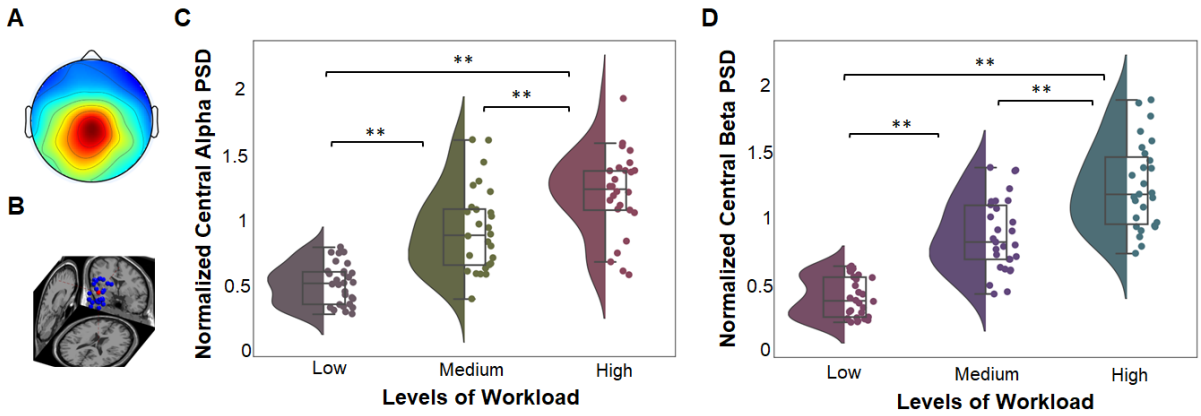

Figure 5.6: Central Cluster, (A) illustrates a scalp map, (B) shows the locations of the dipoles for the components. The centroid of the cluster in the Talairach Coordinate system is located at (3, -5, 53), (C) presents the normalized power spectral density of alpha, and (D) displays the normalized power spectral density of beta at the ICs selected in the Central cluster during blasting (** denotes p < .01).

During blasting, the participants performed a dual arithmetic task, manipulating one-digit, two-digit and three-digit numbers, adding to their memory load. Our results revealed a considerable increase in the frontal theta power with the increasing mental

workload during blasting. These results suggest the increase in memory load [160] with the increasing workload levels, also substantiated by the correlation of frontal theta power with the perceived mental demand.

5.6.3 Central Cluster

Figure 5.6 displays the power spectral density of the ICs in the central cluster for different workload conditions during blasting. The power of alpha in the central cluster increased with increasing workload, as demonstrated by [$F(2, 24) = 124.806$, $p < .01$, $\eta_p^2 = .912$], and this increase was significant for all conditions, including low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p = .02$). Similarly, the power of beta in the central cluster also increased with increasing workload, as demonstrated by [$F(2, 25) = 88.441$, $p < .01$, $\eta_p^2 = .876$], and this increase was also significant for all conditions, including low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p = .03$). However, there were no significant changes observed in the other frequency bands across the workload conditions.

The alpha and beta power at the central region are indicative of sensorimotor processing [7, 134, 274], and desynchronization of alpha and beta power at the central region during dynamic movements of the upper limbs is well-known in the literature [193, 200, 317]. In our experiment, moving the robot arm involved the strenuous movement of both upper limbs, which would explain the desynchronization of alpha and beta power in the central region. However, as the blasting tasks were carried out under three mental workload conditions, we observed a decrease in the alpha and beta power suppressions at the central region with increasing workload. This might be due to competition in resources for sensorimotor integration and memory load, proving our hypothesis.

5.6.4 Parietal Cluster

Figure 5.7 illustrates the power spectral density of the ICs in the parietal cluster for the three workload levels in the blasting task. The power of alpha in the parietal cluster decreased significantly with increasing workload, as demonstrated by [$F(2, 26) = 144.9429$, $p < .01$, $\eta_p^2 = .917$], and this decrease was significant for all conditions, including low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p < .01$). Conversely, the power of beta in the parietal cluster increased with increasing workload, as demonstrated by [$F(2, 27) = 68.117$, $p < .01$, $\eta_p^2 = .835$], and this increase was also significant for all conditions, including low and medium ($p < .01$), low and high

Table 5.2: PSD Changes for Blasting Task

Brain Region	Blasting Task
Frontal	theta PSD ↑ with workload
Central	alpha and beta PSD ↑ with workload
Parietal	alpha PSD ↓ and beta PSD ↑ with workload

($p < .01$), and medium and high ($p < .01$). However, there were no significant changes observed in the delta, theta, or gamma power spectral density of the ICs in the parietal cluster. All these findings are summarized in Table 5.2.

Figure 5.7: Parietal Cluster, (A) illustrates a scalp map, (B) shows the locations of the dipoles for the components. The centroid of the cluster in the Talairach Coordinate system is located at (11, -49, 47), (C) presents the normalized power spectral density of alpha, and (D) displays the normalized power spectral density of beta at the ICs selected in the Parietal cluster during blasting (** denotes $p < .01$).

Further, several studies have noted a desynchronization of alpha power with increased item maintenance in the working memory [51, 126, 282, 330] and memory load [105]. Also, this decrease in alpha power could be associated with the perceived escalation of the time pressure [321]. Our results further demonstrate a decreased parietal alpha power with the increasing workload during blasting. This might be from the increased task difficulty and increased temporal demand as the participants attempt to perform arithmetic operations of varying difficulty while moving the robot arm to the target within the fixed five seconds. These findings were further evidenced by the negative correlation of parietal alpha power with the perceived mental and temporal load scores.

The frontal theta and parietal alpha PSD could reliably determine the mental workload during blasting, proving our first hypothesis (H1).

5.7 Correlation with EEG Measures of Workload

5.7.1 NASA-TLX Scores and EEG Measures of Workload

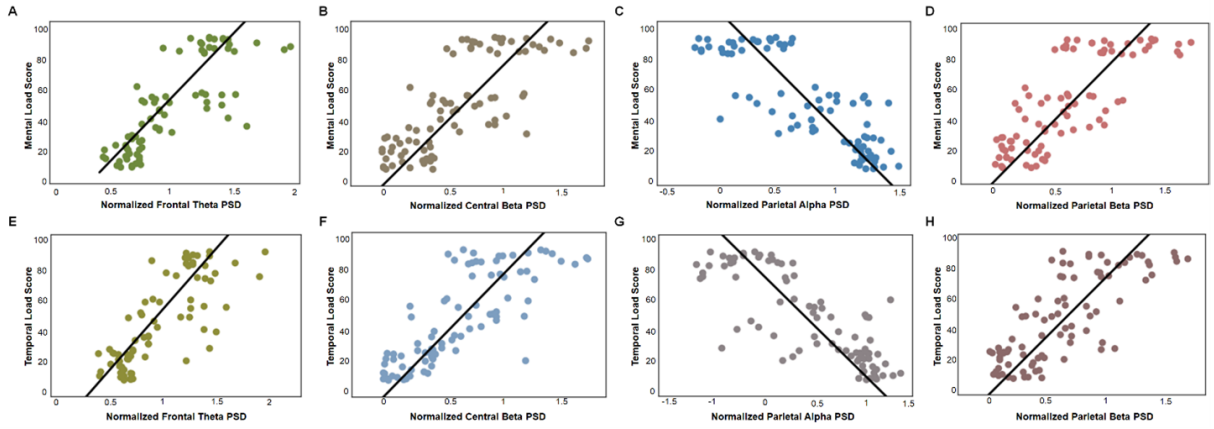

Figure 5.8: Correlations of NASA-TLX scores - Mental Load Scores with the (A) Normalized Frontal Theta, (B) Normalized Central Beta, (C) Normalized Parietal Alpha and (D) Normalized Parietal Beta Powers; Temporal Load Scores with (E) Normalized Frontal Theta, (F) Normalized Central Beta, (G) Normalized Parietal Alpha and (H) Normalized Parietal Beta Powers.

A positive correlation was found between subjective mental demand scores and brain activity in the frontal theta, $r(77) = .797$, $p < .01$, central beta, $r(85) = .764$, $p < .01$ and parietal beta regions, $r(87) = .780$, $p < .01$, as well as a negative correlation with activity in the parietal alpha region, $r(85) = -.816$, $p < .01$. This suggests that an increase in perceived mental load is associated with an increase in activity in the frontal theta, central beta, and parietal beta regions, and a decrease in activity in the parietal alpha region, as shown in Figure 5.8.

Our research found a positive correlation between subjective scores of temporal demand and measures of brain activity, specifically in the frontal theta, $r(77) = .780$, $p < .01$, central beta, $r(85) = .788$, $p < .01$, and parietal beta regions, $r(85) = .769$, $p < .01$ as well as a negative correlation with activity in the parietal alpha region $r(85) = -.793$, $p < .01$. These results suggest that an increase in perceived temporal load is linked to changes in brain activity.

The perceived performance scores were negatively correlated with the frontal theta power, $r(77) = -.768$, $p < .01$, with central alpha, $r(83) = -.717$, $p < .01$ and beta power, $r(85) = -.749$, $p < .01$, and parietal beta power, $r(87) = -.786$, $p < .01$ and positively correlated with parietal alpha power, $r(93) = .839$, $p < .01$. Therefore, the decreases in the perceived performance were correlated with increases in frontal theta, central alpha and beta, and parietal beta power and also, with the decreases in the parietal alpha power.

The subjective effort scores were positively correlated with frontal theta power, $r(77) = .735$, $p < .01$, with central beta, $r(85) = .769$, $p < .01$ and with parietal beta power, $r(87) = .737$, and negatively correlated with parietal alpha power, $r(85) = -.780$, $p < .01$. So, the increases in the perceived temporal load were associated with the increases in frontal theta, central beta and parietal beta power and also, with the decreases in the parietal alpha power.

The subjective frustration scores were found to be positively correlated with frontal theta power, $r(77) = .762$, $p < .01$, with the central beta, $r(85) = .712$, $p < .01$ and with parietal beta power, $r(87) = .798$, and negatively correlated with parietal alpha power, $r(85) = -.804$, $p < .01$. So, the increases in the perceived temporal load were associated with the increases in frontal theta, central beta and parietal beta power and also, with the decreases in the parietal alpha power.

Furthermore, despite maintaining a uniform physical load across trials, surprisingly, the NASA-TLX scores show an increased perceived physical demand with an increasing mental workload. This was accompanied by an increased alpha and beta power in the central region and an increase in beta power in the parietal region, proving our second hypothesis (H2).

5.7.2 Applied Human Force and EEG Measures of Workload

A negative correlation was observed between the alpha power observed in the central cluster and the applied human force, $r(81) = -.770$, $p < .01$. The central beta power had a negative correlation with the applied human force, $r(81) = -.560$, $p < .01$. The increases in central alpha and beta powers were correlated with decreases in applied human force.

Furthermore, the frontal theta power was found to correlate with the applied human force negatively, $r(81) = -.671$, $p < .01$, while the parietal alpha power was found to have a strong positive correlation with the applied human force, $r(81) = .738$, $p < .01$. Therefore, the increases in frontal theta power and decreases in parietal alpha were correlated with decreases in applied human force as shown in Figure 5.9.

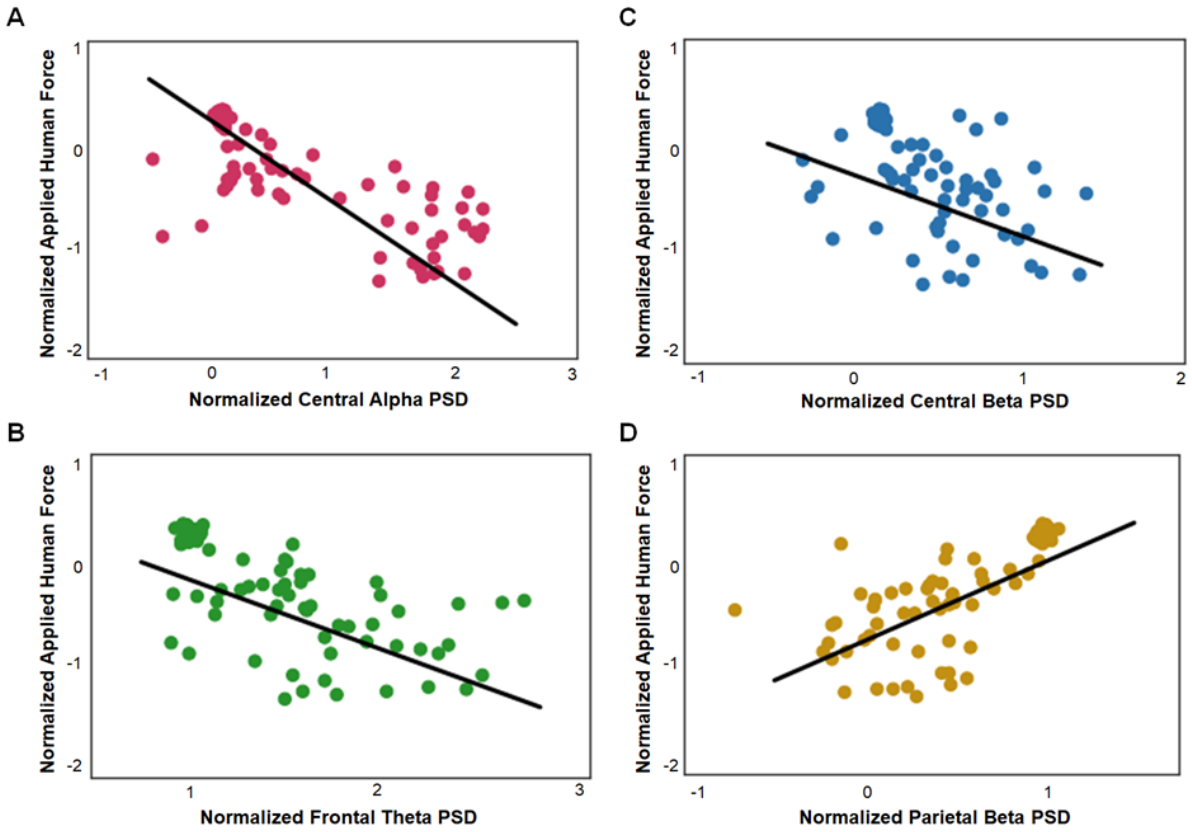

Figure 5.9: Correlations of Applied Force with the (A) Central Alpha, (B) Central Beta, (C) Frontal Theta, and (D) Parietal Alpha Powers.

Nevertheless, previous studies indicate that the neural resources for cognitive processing are entirely independent of those employed for motor processing [91, 128, 358]. Our results indicate a strong correlation not just between physical performance and mental workload but also between central alpha and beta power with the mental workload level. This might also be due to the characteristics of the experiment design, which was put in place to ensure that arithmetic calculations were performed while moving the robot arm. Our experiment design removed the arithmetic equation from the screen as soon as the participant reached the target and had the participant identify the next target soon after. This design aspect of limiting the arithmetic calculation to while the robot arm was being moved might have inadvertently resulted in participants prioritizing the arithmetic calculation over reaching the target and hence, deterioration of physical performance.

5.8 Multiple Regression to Predict Physical Performance

Multiple regression was employed to investigate whether EEG measures (normalized frontal theta power, normalized central alpha and beta power, normalized parietal alpha and beta power) could predict the physical performance as measured using the normalized applied human force. The model explained 71.7% of the variance (R^2 = .858) and could reliably predict physical performance, F(5, 71) = 39.555, p < .01. The normalized frontal theta power (B = -.355, p = .009), normalized parietal alpha power (B = .272, p = .01) and normalized parietal beta power (B = -.314, p = .003) contributed significantly to the model. The contributions of the normalized central alpha power (B = -.022, p = .869) and normalized central beta power (B = -.165, p = .716) were not significant. Therefore, the model for physical performance was:

(5.1)

$$Physical\_Performance = 0.272*P\alpha - 0.314*P\beta - 0.355*F\theta - 0.022*C\alpha - 0.165*C\beta + 0.227$$

where $P\alpha$ is the normalized parietal alpha power, $P\beta$ is the normalized parietal beta power, $F\theta$ is the normalized frontal theta power, $C\alpha$ is the normalized central alpha power, and $C\beta$ is the normalized central beta power.

5.9 Summary

This work presented a systematic investigation into the validity of EEG measures in assessing mental workload during physical human-robot collaboration. Twenty-four participants' EEG signals and perceived workload were examined while they moved the robot arm and performed arithmetic calculations of varying difficulty. Task performance, physical performance and subjective measures revealed the effectiveness of workload manipulation in the task. Task and physical performance degraded while the perceived workload increased with increasing workload levels. Furthermore, the theta power in the frontal region increased along with decreased parietal alpha power with increasing workload levels during blasting. This was accompanied by increased central alpha and beta power with increasing workload. This work demonstrates how EEG measures can be employed for mental workload assessment in physical human-robot collaborations with strenuous physical activity. As our study emulated uncontrolled, real-world physical interactions with the ANBOT, our results are encouraging.

Additionally, increased perceived physical load and deterioration of physical performance were observed with the increasing mental workload in physical human-robot collaboration. Also, the changes in physical performance (evaluated by applied human force) can be reliably estimated from the EEG measures of mental workload. The results of this study demonstrate to the designers the relevance of maintaining the users' mental workload to ensure the quality of physical human-robot collaboration.

CHAPTER 6

EFFECTS OF MENTAL WORKLOAD VARIATIONS ON ERROR AWARENESS IN A PHYSICAL HUMAN-ROBOT COLLABORATION

6.1 Background

Mishaps and accidents occur in overloaded conditions due to time pressure or hecticness and also in underloaded situations due to boredom or carelessness [151]. Most work environments are risk-assessed to identify and mitigate these factors. In work environments involving human-machine interactions, safety is ensured by maintaining spatial separation between the human and robot. However, these classical safety measures may be insufficient for a modern work environment, especially a collaborative environment where humans and machines work in close proximity, often with physical contact.

BCI technology plays a prominent role in enabling occupational safety in physical human robot collaboration by providing a direct channel for communicating the cognitive state of the user. With BCI, the robotic partner gets a direct glimpse of the user,Äôs brain activity affording an alternate communication pathway rather than relying on traditional communication routes which might distract the human partner from the task at hand. Embracing passive BCI technology in pHRC affords the ability to optimally adapt the robot behaviour for each individual user using real-time brain activity, enriching the

quality of the interaction [332, 377].

EEG represent the electrical activity of the brain as measured from the scalp [246] and EEG has been employed on BCI technology to decipher the cognitive state of the user [60, 178] and adapt the interaction accordingly.

In a pHRC, the robot might perform complex motion sequences, unpredicted positions, and unexpected changes in velocity [379]. Any unpredicted or erroneous robot behaviour or a faulty interaction can be measured from EEG as error-related potential naturally generated by the brain during cognitive conflict when the human partner perceives an unexpected behaviour or error occurrence [30, 106, 123]. This ERrP consists of a time-locked negative peak or error-related negativity around 200 ms after the error occurrence [107, 339]. It is mostly followed by a positive peak around 300 ms after the error onset [106, 107], and this error positivity is related to the user,Äôs error awareness [353].

As this error-related potential occurs at short latency, when detected online using machine learning methods [62, 175, 176], can be used as implicit human feedback to optimize robot behaviour in the environment [177]. Based on the context of interaction, ERrP detection can aid the automatic detection of higher-level user intention and achieve personalized rule-based robot control [159] for a safe and efficient pHRC.

ERrPs detection algorithms have been employed to train a reinforcement learning algorithm to train robots to mirror the gestures of the human user [157, 177]. ERrP was also used in a closed-loop system to switch the trajectory of a robot arm between two targets [306]. Further, Ehrlich and Cheng employed ERrP signals in a closed-loop system to adapt the robot gaze based on the intention of the user [97] and later compared the ERrP detection performance in case of unexpected movement of a robot head or cursor [98]. However, in these studies, human users were merely passive observers of the error unlike in real-world applications.

Prediction error negativity (PEN) is the ERrP in mobile brain-body imaging (MoBI) setting [211] that appears during the task, prior to any external error indication, and it was first observed in a cognitive conflict study with active physical movement [320]. In 2019, Aldini and colleagues observed PEN during an unexpected robot behaviour in pHRC [5]. They evaluated the intuitiveness in a pHRC by comparing the modulations in the amplitude of PEN under different degrees of mechanical resistance in the transition to an unexpected robot stopping. Recently, they evaluated the performance of six classifiers in PEN detection, and this work is a step in the direction of intuitive pHRC [6]. However, the integration of BCI technology with PEN for robot control demands

failure-free integration of highly uncertain EEG signals, particularly influenced by the physical and cognitive state of the user.

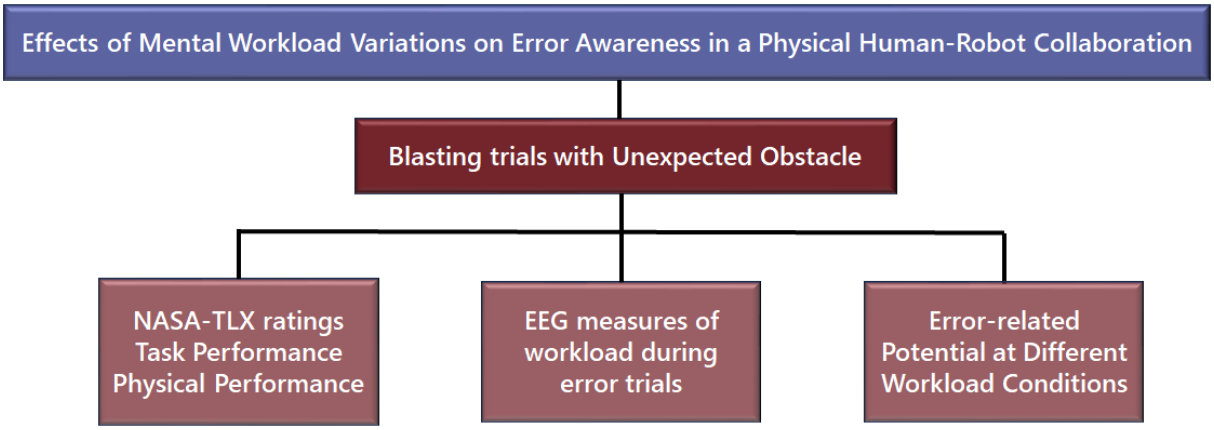

Figure 6.1: Organization of Chapter 6.

6.2 Research Hypothesis

Mentally demanding tasks might not only compromise access to cognitive resources [43] but might completely deplete resources [261] needed for successfully executing the task. Mental workload reflects the tuning between the task demand and the cognitive capacity of the user [137]. As task demands and the user's cognitive capacity continuously vary over time, the mental workload experienced by a user at any point in time could be unpredictable, and therefore, a one-size-fits-all adaptation approach may not accommodate the inter-subject variability. Further, there has been no systematic study on how mental workload variations impact cognitive conflict response. The detrimental effects of workload on human performance and situational awareness are well-known in literature; however, its influence on error awareness might also raise safety concerns in physical human robot collaboration.

We hypothesize that with increasing workload, there will be a diminished error awareness, reflected in reduced error-related potential. We employ MoBI methods to study the impact of workload variations on error awareness or cognitive conflict. We designed our experimental paradigm as an interactive clock game with blasting that emulates the operation of the abrasive robot. The mental workload variations were achieved using arithmetic tasks with low, medium, and high task difficulty. The physical load was constant across these conditions to investigate how mental workload variations influence physical performance.

An error occurrence was achieved by an unexpected robot stopping at an invisible obstacle during the blasting operation. The NASA Task Load Index (NASA TLX) questionnaire [141] was also employed in this investigation to validate the mental workload variations by incorporating a subjective measure of workload. We hypothesize that an erroneous robot behavior from an unexpected robot stopping during the blasting operation will result in a diminished PEN and Pe amplitude with increasing workload. The chapter organization is shown in Figure 6.1.

6.3 Subjective Workload Measures during Error Conditions under Different Workload Conditions

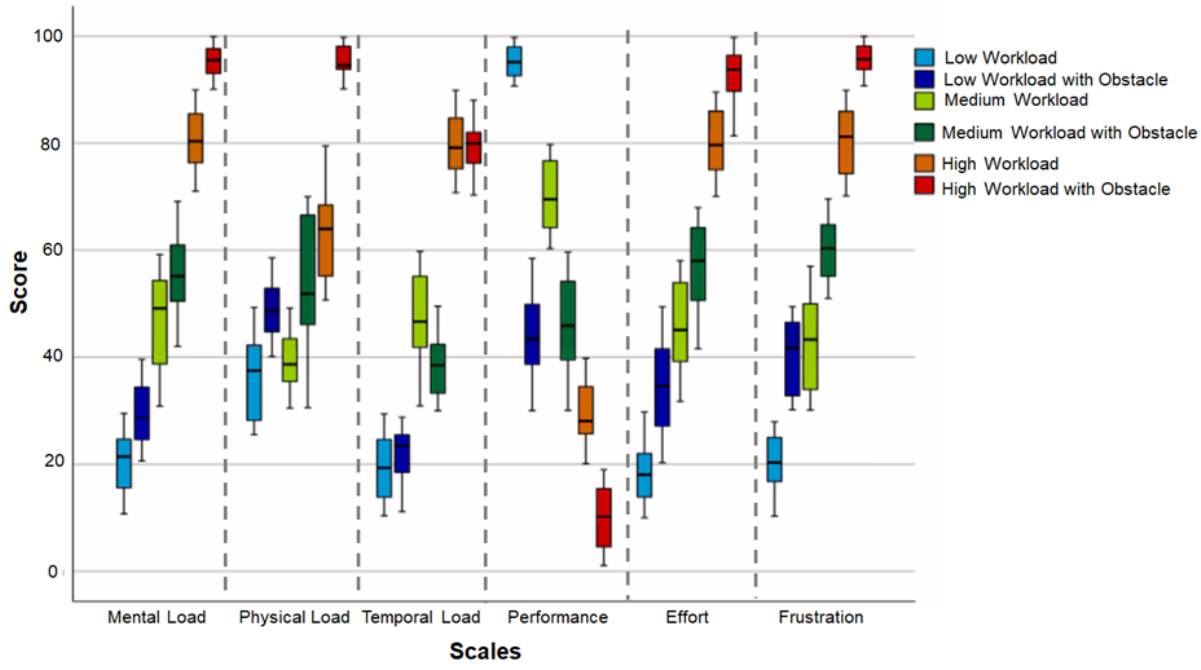

Figure 6.2: NASA-TLX scores during the normal and obstacle conditions for the three workload levels (** denotes p < .01).

6.3.1 Mental Demand Scores

The subjective workload ratings increased with the workload in the clock game task with unexpected robot stopping, as shown in Figure 6.2. The results of the two-way repeated measures ANOVA revealed that there was a significant main effect of workload on participants' mental demand scores (F(2, 60) = 1499.764, p < .01, η_p^2 = .980). Also,

the ANOVA revealed that mental demand scores during trials with unexpected robot stopping were significant ($F(1, 30) = 163.748$, $p < .01$, $\eta_p^2 = .845$). Further, there was a significant interaction between workload and unexpected robot stopping ($F(2, 60) = 3.191$, $p = .048$, $\eta_p^2 = .096$) such that participants reported increasing mental demand scores with the increasing workload during conditions with unexpected robot stopping than the normal conditions.

6.3.2 Physical Demand Scores

There was a significant main effect of workload on participants' physical demand scores ($F(2, 60) = 403.112$, $p < .01$, $\eta_p^2 = .931$). Also, the ANOVA revealed that physical demand scores during trials with unexpected robot stopping was significant ($F(1, 30) = 408.133$, $p < .01$, $\eta_p^2 = .932$). Further, there was a significant interaction between workload and unexpected robot stopping ($F(2, 60) = 26.118$, $p < .01$, $\eta_p^2 = .465$) such that participants reported greater physical demand scores with the increasing workload during conditions with unexpected robot stopping than the normal conditions.

Furthermore, despite maintaining a uniform physical load across trials, the NASA-TLX scores show an increased perceived physical demand with an increasing mental workload during the obstacle condition.

6.3.3 Temporal Demand Scores

The results of the two-way repeated measures ANOVA revealed that there was a significant main effect of workload on participants' temporal demand scores ($F(2, 60) = 1177.726$, $p < .01$, $\eta_p^2 = .975$). Also, the ANOVA revealed that temporal demand scores during trials with unexpected robot stopping was significant ($F(1, 30) = 12.462$, $p = .001$, $\eta_p^2 = .293$). Further, there was a significant interaction between workload and unexpected robot stopping ($F(2, 60) = 15.432$, $p < .01$, $\eta_p^2 = .340$) such that participants reported increasing temporal demand scores with the increasing workload during conditions with unexpected robot stopping than the normal conditions. There was a significant main effect of workload on participants' performance scores ($F(2, 60) = 1016.185$, $p < .01$, $\eta_p^2 = .971$).

6.3.4 Performance Scores

ANOVA revealed that performance scores during trials with unexpected robot stopping were significant ($F(1, 30) = 925.325$, $p < .01$, $\eta_p^2 = .969$). Further, there was a significant

interaction between workload and unexpected robot stopping ($F(2, 60) = 100.151$, $p < .01$, $\eta_p^2 = .769$) such that participants reported higher performance scores with the increasing workload during conditions with unexpected robot stopping than the normal conditions.

6.3.5 Effort Scores

There was a significant main effect of workload on participants' effort scores ($F(2, 60) = 1091.580$, $p < .01$, $\eta_p^2 = .973$). Also, the ANOVA revealed that effort scores during trials with unexpected robot stopping were significant ($F(1, 30) = 202.924$, $p = .001$, $\eta_p^2 = .871$). However, there was no significant interaction between workload and unexpected robot stopping as the participants reported similar increasing effort scores with the increasing workload during unexpected robot stopping and normal conditions.

6.3.6 Frustration Scores

There was a significant main effect of workload on participants' frustration scores ($F(2, 60) = 1368.842$ $p < .01$, $\eta_p^2 = .979$). Also, the ANOVA revealed that frustration scores during trials with unexpected robot stopping were significant ($F(1, 30) = 334.303$, $p < .01$, $\eta_p^2 = .918$). Also, there was no significant interaction between workload and unexpected robot stopping, as the participants reported similarly increasing frustration scores with the increasing workload during unexpected robot stopping and normal conditions.

6.4 Arithmetic task Performance during Error trials

The target identification accuracy deteriorated significantly with the rising workload during the unexpected robot stopping conditions, [$F(2, 42) = 228.770$, $p < .001$, $\eta_p^2 = .916$], as shown in Figure 6.3(A). Target identification accuracy significantly decreased for low and medium ($p = .004$), low and high ($p < .01$), and medium and high ($p < .01$) conditions. The time to reach the obstacle increased with the increasing workload in the unexpected robot stopping conditions, [$F(2, 60) = 113.896$, $p < .01$, $\eta_p^2 = .792$], as shown in Figure 6.3(B). The increase was significant for low and medium ($p < .01$), low and high ($p < .01$), and medium and high conditions ($p < .01$).

Performance in the blasting task, assessed by the target identification accuracy and time needed to reach the obstacle, degraded notably with the rising workload. Therefore, the workload manipulation during blasting in error conditions successfully evoked substantial performance variations.

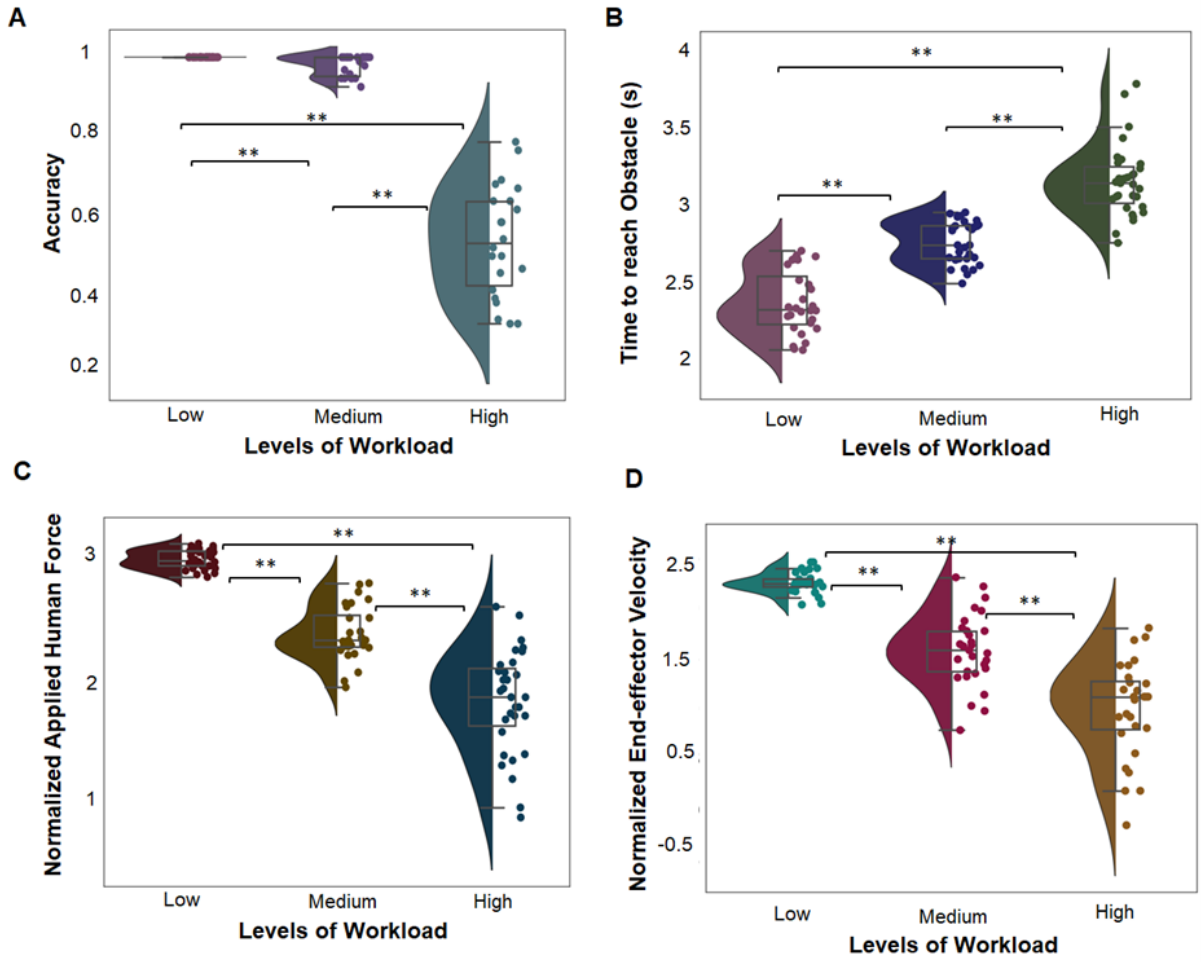

Figure 6.3: (A) shows the target identification accuracy, (B) shows the time to reach the obstacle, (C) shows the normalized applied force, and (D) shows the normalized end-effector velocity for the three workload levels with unexpected robot stopping condition due to obstacle in the blasting path. ** denotes $p < .01$.

6.5 Physical Performance during Error trials

The applied human force decreased significantly with the increasing workload during the unexpected robot stopping conditions, [$F(2, 54) = 130.340$, $p < .01$, $\eta_p^2 = .828$], as shown in Figure 6.3(C). Applied human force decreased significantly for low and medium ($p < .01$), low and high ($p < .01$), and medium and high conditions ($p < .01$). The end-effector velocity decreased with the increasing workload during the unexpected robot stopping conditions, [$F(2, 54) = 86.647$, $p < .01$, $\eta_p^2 = .762$], with significant decreases for low and medium ($p < .01$), low and high ($p < .01$), and medium and high ($p < .01$), as shown in

Figure 6.3(D).

The physical performance assessed by the applied human force and the resultant end-effector velocity decreased significantly with increasing workload levels prior to reaching the obstacle. Hence, the increasing mental workload resulted in significant physical performance deterioration.

6.6 EEG Measures of Workload during Error Processing

In order to investigate the neural correlates of error processing under three levels of mental workload variations, the independent components were extricated from the scalp EEG data, and significant correlations between the task load variations and the spectral powers at frontal, central, and parietal clusters were observed.

6.6.1 Independent Source Clusters

The frontal, central, and parietal clusters were identified based on the dipole locations [259]. The Talairach coordinate of the frontal cluster centroid was at (4, 38, 39) and included 19 components from the 19 subjects. The Talairach coordinate of the central cluster centroid was at (3, -5, 51) and had 18 components from 18 subjects. The Talairach coordinate of the parietal cluster centroid was at (11, -49, 45), including 24 subjects from 24 subjects.

6.6.2 Frontal Cluster

The frontal cluster's ICs' PSD for the different workload conditions during blasting is shown in Figure 6.4. There was a significant main effect of workload on frontal theta PSD ($F(2, 42) = 37.894$, $p < .01$, $\eta_p^2 = .643$) during unexpected robot stopping conditions. Frontal theta PSD increased significantly in the low-medium ($p < .01$), medium-high ($p = .003$) and low-high ($p < .01$) conditions shortly after the unexpected robot stopping. Further, the frontal theta power was significantly higher after reaching the obstacle as compared to reaching the target for the low ($p < .01$), medium ($p = .029$), and high ($p < .01$) workload conditions. However, there were no significant differences in the frontal cluster's delta, alpha, beta, or gamma PSD.

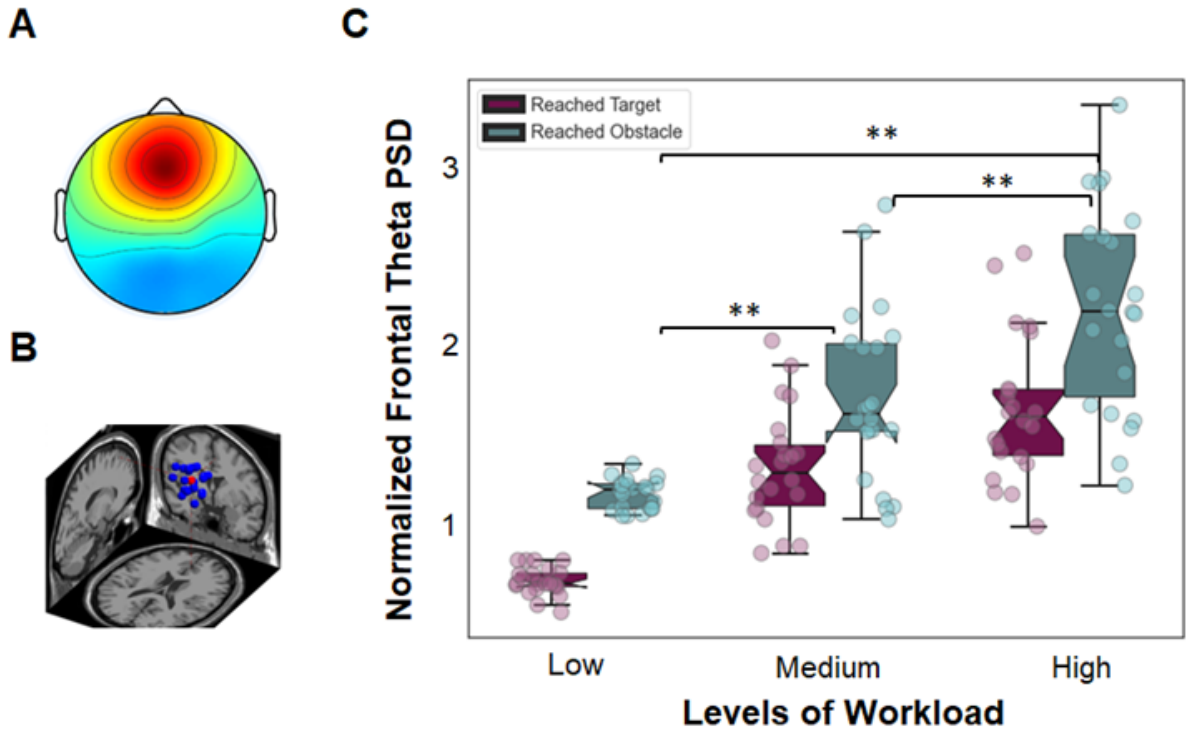

Figure 6.4: Frontal Cluster - (A) scalp map, (B) dipole locations of the components, the Talairach Coordinate of the cluster centroid was at (4, 38, 39) and (C) normalized theta PSD for the one second immediately after reaching the target or the obstacle at the ICs selected in the Frontal cluster during blasting. ** denotes $p < .01$.

During blasting, the participants performed a dual arithmetic task, manipulating one-digit, two-digit, and three-digit numbers, adding to their memory load. Our results revealed a considerable increase in the frontal theta power with the increasing mental workload immediately after reaching an obstacle or unexpected stopping compared to reaching the target or expected stopping. These results suggest the increase in memory load [160] with the increasing workload levels after reaching the obstacle. This increase in the memory load might also be due to the unexpected removal of the arithmetic equation from the screen when the participants reach the obstacle, requiring them to rely on their working memory to continue the calculations. This was further evidenced by an increase in the mental demands score of the NASA-TLX in the obstacle condition compared to the normal blasting condition.

6.6.3 Central Cluster

The central cluster's ICs' spectral power for different workload conditions soon after reaching the obstacle in the unexpected robot stopping condition and after reaching the target in the normal conditions is shown in Figure 6.4. There was a significant main effect of workload on central alpha PSD ($F(2, 26) = 15.274$, $p < .01$, $\eta_p^2 = .540$) during unexpected robot stopping condition. Central alpha PSD increased significantly in the medium-high ($p = .038$) and low-high ($p < .01$) conditions shortly after the unexpected robot stopping. Further, the central alpha power varied significantly only for the low workload condition ($p < .01$) upon reaching the obstacle compared to reaching the target. The central beta PSD varied significantly after reaching the obstacle across the three workload levels ($F(2, 30) = 18.453$, $p < .01$, $\eta_p^2 = .552$). Central beta PSD increased significantly in the medium-high ($p = .002$) and low-high ($p = .001$) conditions shortly after the unexpected robot stopping. Further, the central beta power increased significantly for the low ($p < .01$) and medium ($p < .01$) workload conditions upon reaching the obstacle as compared to reaching the target. No significant changes were noted in the other frequency bands across the workload conditions.

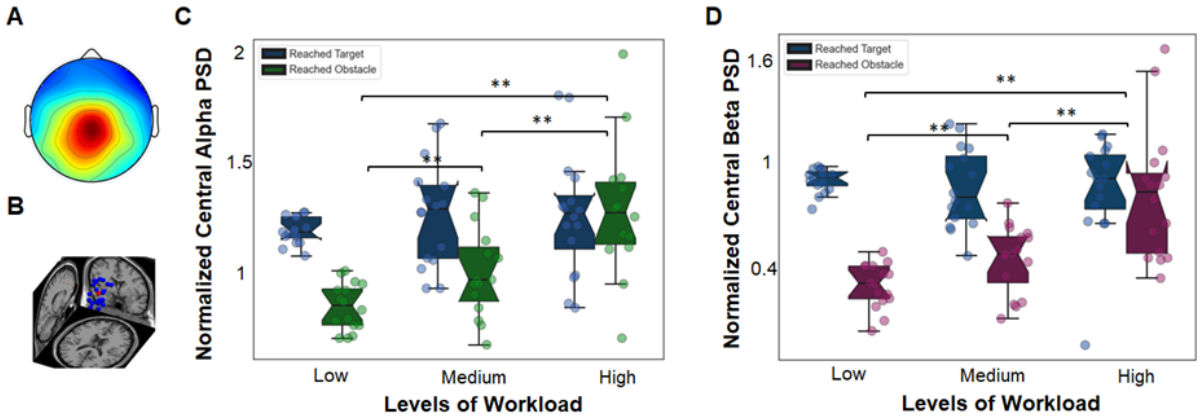

Figure 6.5: Central Cluster - (A) scalp map, (B) dipole locations of the components, the Talairach Coordinate of the cluster centroid was at (3, -5, 51), (C) normalized alpha PSD, and (D) normalized beta PSD for the one second immediately after reaching the target or obstacle at the ICs selected in the Central cluster (** denotes $p < .01$).

In our experiment, once the participants reach either the target or the obstacle, the robot arm stops and cannot be moved until the end of the blasting period of five seconds. We saw an increase in the alpha and beta power at the central region with increasing workload immediately after reaching the obstacle as opposed to the similar

alpha and beta power across workload conditions after reaching the target in the normal condition. The alpha and beta power at the central region is indicative of sensorimotor processing [7, 134, 274], and desynchronization of alpha and beta power at the central region during dynamic movements of the upper limbs is well-known in the literature [193, 200, 312]. The participants might have attempted to move the robot arm to the target after the unexpected robot stopping resulting in a decrease in the alpha and beta power suppressions at the central region with increasing workload.

6.6.4 Parietal Cluster

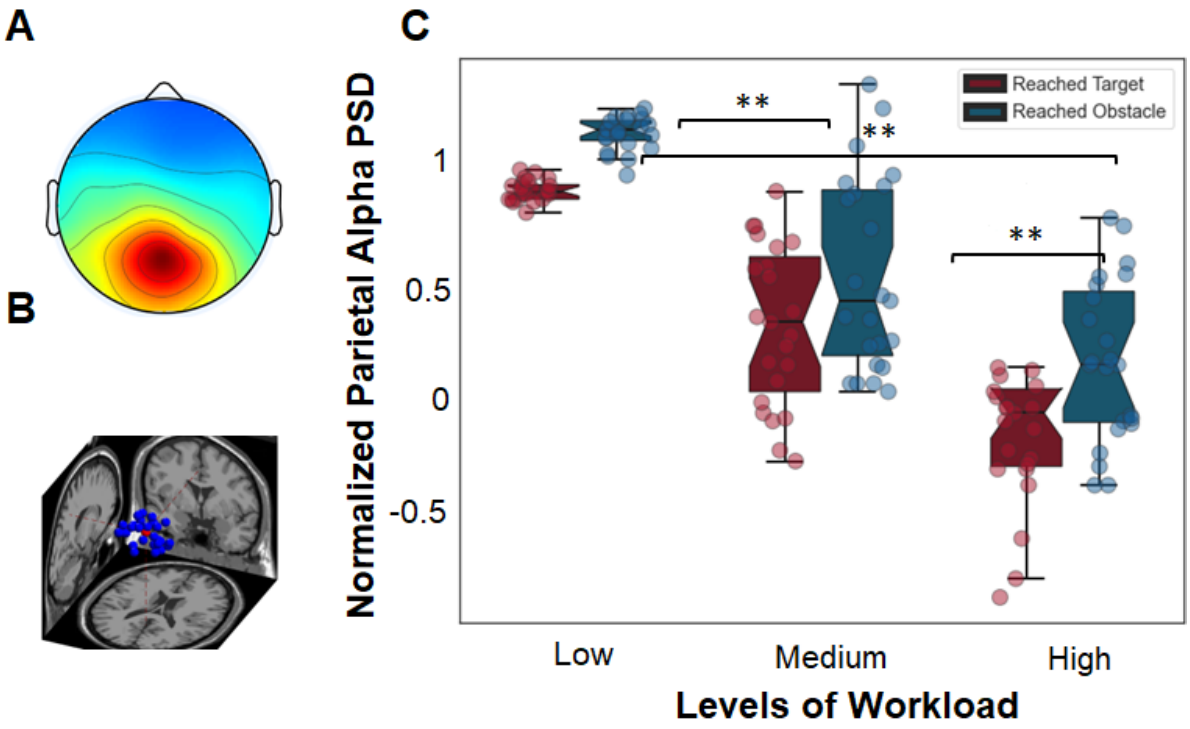

Figure 6.6: Parietal Cluster - (A) scalp map, (B) dipole locations of the components, the Talairach Coordinate of the cluster centroid was at (11, -49, 45) and (C) normalized alpha PSD for the one second after reaching the target or the obstacle at the ICs selected in the Parietal cluster (** denotes $p < .01$).

The parietal cluster's ICs' spectral power for the three workload levels upon reaching the target in the unexpected robot stopping condition or reaching the obstacle in the normal condition is illustrated in Figure 6.6. There was a significant main effect of workload on parietal alpha PSD ($F(2, 44) = 59.458$, $p < .01$, $\eta_p^2 = .730$) during unexpected robot

stopping condition. Parietal alpha PSD decreased significantly in the low-medium ($p < .01$), medium-high ($p = .003$), and low-high ($p < .01$) conditions shortly after unexpected robot stopping. Further, the parietal alpha power varied significantly only for the low workload condition ($p < .01$) upon reaching the obstacle compared to reaching the target. No significant changes were noted in delta, theta, beta, or gamma PSD at the parietal cluster.

Our results further demonstrate an increased parietal alpha power with the increasing workload soon after reaching the obstacle compared to reaching the target. The decrease in the parietal alpha power with increasing workload might be related to the desynchronization of alpha power with increased item maintenance in the working memory [51, 126, 282, 330] and memory load [105]. Our claims were also validated by the increased mental and temporal demand score with the increasing workload in the obstacle condition compared to the normal condition.

6.7 Error-related Potential at Different Workload Conditions

The ErRP at the Fz electrode showed significant changes in the amplitudes of the PEN and Pe, as shown in Figure 6.7. Further, the ANOVA results showed a significant increase in the Prediction Error Negativity with increasing workload ($F(2, 60) = 55.276$, $p < .01$, $\eta_p^2 = .648$). PEN increased significantly for low-medium ($p < .01$), medium-high ($p < .01$) and low-high ($p < .01$) workload conditions, as shown in Figure 6.7(A) and 6.7(B). The error positivity (Pe) also varied significantly across the workload levels ($F(2, 60) = 415.107$, $p < .01$, $\eta_p^2 = .933$). Pe decreased significantly for low-medium ($p < .01$), medium-high ($p < .01$) and low-high ($p << .01$) workload conditions, as shown in Figure 6.7(A) and 6.7(C). Further, the normalized applied force was negatively correlated with the PEN, $r(90) = -.644$, $p < .01$ and positively correlated with Pe, $r(90) = .785$, $p < .01$.

This work also investigated the neural correlates of error processing under three workload levels. We detected PEN and Pe after the unexpected robot stopped despite the load on the cognitive resources induced by performing arithmetic calculations. PEN and Pe amplitudes significantly decreased with increasing workload (proving our hypothesis). This indicates a shortage of cognitive resources for error processing in increased workload conditions despite similar error types and severity. Earlier studies have observed a reduction in P300 amplitude at high workload or fatigued conditions [178, 186]. The changes in the central alpha and beta power immediately after reaching the obstacle

or at the error occurrence due to the participant's continued efforts to move the robot arm might be a potential confounding factor. However, these changes were observed in the alpha and beta power at the central region while our ERrP processing was mainly focused on the theta band at the frontal region, known to be the principal frequency band and region of interest [55, 62, 155].

As PEN and Pe have low amplitudes of few microvolts, they are more prone to distortion from artifacts introduced in the measured EEG signals during active physical interaction with the robot [281]. In higher workload conditions, our results demonstrate that PEN and Pe amplitudes were further diminished, raising potential safety concerns when ERrP signals are relied upon for ensuring a safe and efficient pHRC environment. Therefore, at high workload conditions, the user‚Äôs error processing capacity becomes compromised, possibly allowing a safety breach that might escape the attention of the human user.

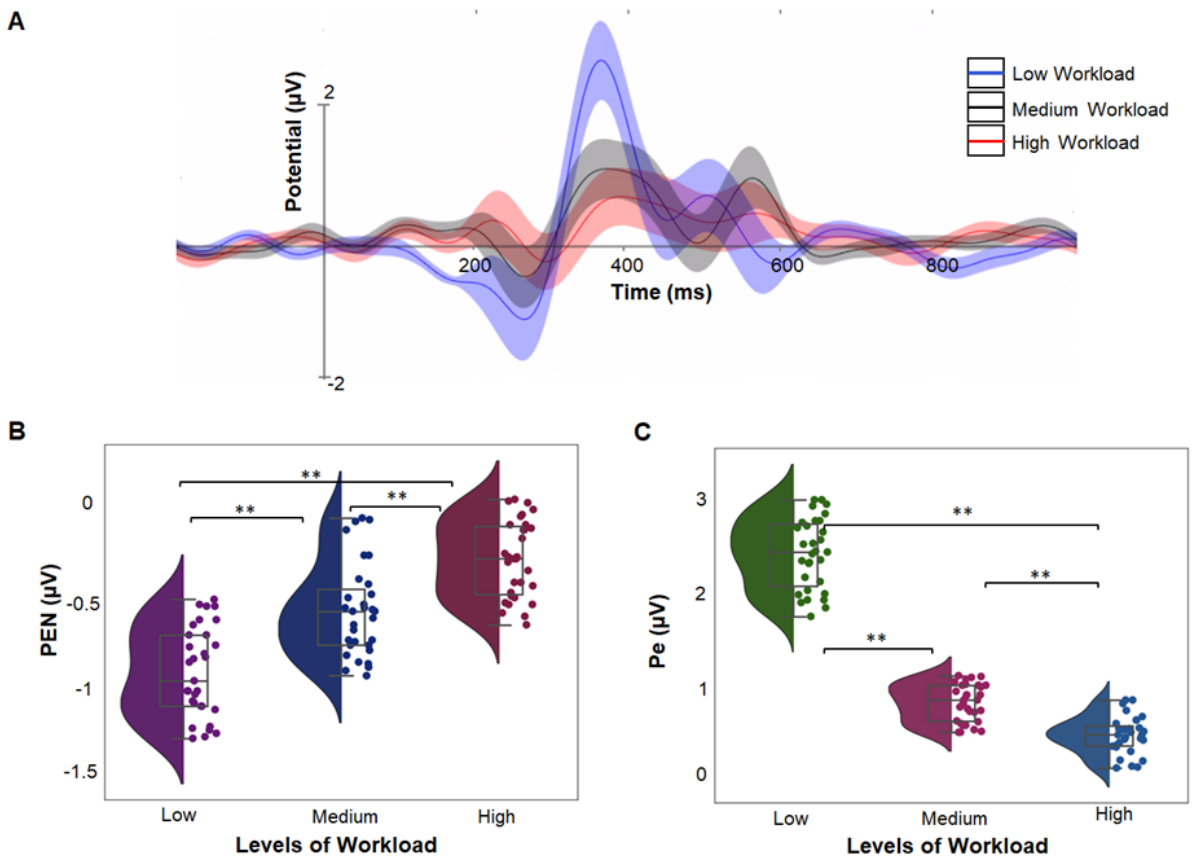

Figure 6.7: A) ErRP at the Fz channel, (B) PEN, and (C) Pe upon unexpected robot stopping at the obstacle for the three workload levels. (** denotes $p < .01$).

6.8 Summary

This work presented a systematic investigation into the impact of workload variations on cognitive conflict in a physical human-robot collaboration. Twenty-four participants' EEG signals, subjective workload rating, task, and physical performance measures were examined while they moved the robot arm and performed arithmetic calculations of varying difficulty. Task performance, physical performance, and subjective measures revealed the effectiveness of workload manipulation in the task. Task and physical performance degraded while the perceived workload increased with increasing workload levels. Furthermore, the perceived workload was greater in the error trials compared to the normal trials. We observed a decrease in the amplitude of the error-related potential with increasing workload levels. Typical deep learning models could predict the PEN and Pe amplitude from the EEG data prior to the error, allowing an opportunity to adapt the system dynamics to lower the workload of the human user. Additionally, we observed significant increases in the frontal theta, central alpha, and beta power soon after the error as compared to the normal condition. This work exposes the significance of measuring and maintaining the human user's workload at an optimal level to ensure the safe and reliable use of BCI technology for intuitive physical human-robot collaboration.

CHAPTER 7

INTELLIGENT ONLINE WORKLOAD ASSESSMENT AND SAFETY ALERT TOOL FOR PHYSICAL HUMAN ROBOT COLLABORATION

7.1 Background

Accidents can happen in situations where there is too much pressure or stress, as well as in situations where there is not enough to do, such as boredom or carelessness [151]. Employers typically assess the risks in the workplace to identify and address these issues. In workplaces where people interact with machines, safety is maintained by keeping people and machines separated. However, these traditional safety measures may not be enough in more modern work environments, especially when humans and machines are working together closely and sometimes in physical contact.

Assessing the level of safety in physical human robot collaboration requires various techniques, including direct system input, behavioral or physiological data, and questionnaires. However, while using multiple methods is ideal for a more accurate evaluation [34], some studies have chosen to only utilize questionnaires. This is likely because of their cost-effectiveness compared to other methods, such as physiological measurements. Another common approach is to use questionnaires in conjunction with behavioral assessment. Physiological measurements are usually paired with questionnaires as well, likely because of their cost-effectiveness. However, the use of direct input devices is less

frequent.

Safety in physical human robot collaboration can be interpreted in a variety of ways. Psychological safety refers to people's perceptions of the potential consequences of taking interpersonal risks in specific contexts, such as a workplace [95]. It can also be seen as a shared belief among a group that allows them to take risks [2]. In the context of pHRC, maintaining psychological safety involves making sure the human feels safe interacting with the robot, and that the interaction does not cause any psychological discomfort or stress because of the robot's attribute [197].

In psychology, mental safety is sometimes used as an alternative term for psychological safety [262, 313]. In the context of physical human robot collaboration (pHRC), mental safety is described as being related to the mental stress and anxiety caused by close interactions with robots [346], or as the state in which humans do not feel fear or surprise towards the robot [305]. Subjective safety can be considered as a general measure that reflects a person's perception of the security of a particular location [270, 325]. It is not limited to personalized safety systems that account for the unique characteristics of human beings [337], which is also sometimes referred to as subjective safety.

Especially since the robot may perform complex motion sequences, take unexpected positions, and change velocity in a physical human robot collaboration, it would raise safety concerns for human partners [379]. Any unpredicted or incorrect robot behavior or a faulty interaction can be detected from EEG as an error-related potential generated by the brain during cognitive conflict when the human partner perceives an unexpected behavior or error occurrence [30, 106, 123]. This error-related potential (ERrP) consists of a time-locked negative peak or error-related negativity around 200 ms after the error occurrence [107, 339]. It is generally followed by a positive peak around 300 ms after the error onset [106, 107], and this error positivity is related to the user's error awareness [353].

Since this error-related potential occurs with a short latency, when detected online using machine learning methods [62, 175, 176], it can serve as implicit human feedback to optimize the robot's behavior in the environment [177]. Depending on the context of the interaction, detecting ERrP can aid in automatically detecting the user's higher-level intentions and achieve personalized, rule-based robot control [159] for a safe and efficient pHRC.

Scientists have used ERrP detection algorithms to train robots to mirror human gestures using reinforcement learning [157, 177]. They also employed ERrP detection in closed-loop systems to change the trajectory of a robot arm between two targets [306].

Additionally, Ehrlich and Cheng used ERrP signals in a closed-loop system to adjust the robot's gaze based on the user's intention [97] and later compared the ERrP detection performance when the robot's head or cursor movement was unexpected [98]. However, in these studies, human users were simply passive observers of the errors, unlike in real-world applications.

The Prediction error negativity, or PEN, is a type of brain activity that is seen in mobile brain imaging studies [211] when a task is being performed. It is observed before any external indication of an error and was first discovered during a study on cognitive conflict and physical movement [320].

7.2 Research Hypothesis

Brain-computer interface (BCI) technology plays a crucial role in improving occupational safety in physical human robot collaboration by providing a direct means of communicating the cognitive state of the user. With BCI, the robotic partner can directly monitor the user's brain activity, offering an alternative communication method instead of relying on traditional methods, which can distract the human partner from the task at hand. Adopting passive BCI technology in pHRC allows the ability to adapt the robot's behavior to each individual user using real-time brain activity, enhancing the quality of the interaction [332, 377].

However, using PEN in combination with brain-computer interface technology for controlling robots can be challenging because it requires the seamless integration of EEG signals that can be greatly affected by the user's physical and mental state.

Mentally demanding tasks might not only compromise access to cognitive resources [43] but might completely deplete resources [261] needed for successfully executing the task. Mental workload reflects the tuning between the task demand and the cognitive capacity of the user [137]. As task demands and the user's cognitive capacity continuously vary over time, the mental workload experienced by a user at any point in time could be unpredictable, and therefore, a one-size fits-all adaptation approach may not accommodate the inter-subject variability.

Therefore, this research aims to develop an intelligent workload assessment and safety alert for pHRC. we hypothesize that deep learning models will reliably classify workload in physical human robot collaboration. We also hypothesize that the error awareness of the operator can be accurately predicted based on their brain activity prior to the error occurrence. We employ an online workload classifier and error awareness

predictor for designing the framework for a safety alert tool in physical human robot collaboration.

Our experiment was set up as a game where participants controlled a clock while also operating an abrasive robot. We manipulated the mental demands of the task by including arithmetic problems of varying difficulty. However, the physical demands of the task remained the same in order to see how changes in mental workload affect physical performance. We introduced an error into the task by having the robot unexpectedly stop at an unseen obstacle during the blasting operation.

We used common deep learning techniques, like ShallowConvNet [314], DeepConvNet [314], and EEGNet [198] to analyze the EEG data before the obstacle or unexpected robot stopping, in order to estimate the PEN and Pe values. Additionally, we utilized the same methods to classify the mental workload of the operator. Our hypothesis is that by analyzing the EEG dynamics before the obstacle or unexpected robot stops, we can accurately predict the PEN and Pe amplitude with high precision. The chapter organization is shown in Figure 7.1.

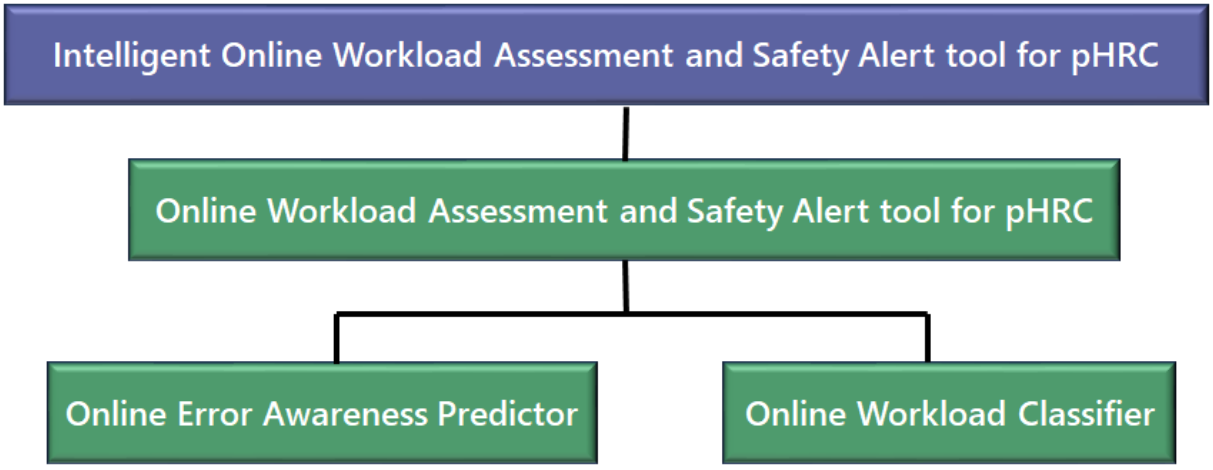

Figure 7.1: Organization of Chapter 7.

7.3 Online Workload Assessment and Safety Alert tool for pHRC

In this study, we presented a new safety alert tool for physical human-robot collaboration that incorporates an error awareness predictor and a workload classifier. The tool was tested on 24 healthy individuals while they were operating ANBOT in scenarios that

required continuous physical contact. The test task involved a blasting task under different levels of workload, which were induced by a secondary arithmetic task. In 33% of the trials, the robot would stop unexpectedly or make an error, resulting in an error-related potential, which was evaluated using prediction error negativity, error positivity, or power spectral density.

For this research, we developed a clock game that simulated an abrasive blasting task with an unexpected robot stopping, and we manipulated the mental workload level by using a dual arithmetic task of varying difficulty. We incorporated an online safety alert tool that included both a workload classifier and an error awareness predictor. The tool would send a signal to the ROS system controlling ANBOT to safely stop the robot when the predicted error awareness was below a certain threshold, which was determined by the average of the training data. The error awareness predictor triggered the safety alert tool to intervene and stop the robot when the human operator's error awareness decreased. The workload classifier served as a diagnostic tool that provided real-time information on the detected workload level.

So, we designed a safety alert tool that raises the alarm and stops the functioning of the robot when the error awareness of the human operator diminishes beyond detection. To evaluate the accuracy of the error awareness predictor and workload classifier, we used a within-subject cross-validation method. We trained the models using data from the first two blocks of the experiment and then tested them on the data from the final block. Additionally, we used a cross-subject cross-validation method to assess the performance of the overall online safety alert system. We trained the models using data from 23 participants and then tested them using data from one additional participant.

7.3.1 Online Error Awareness Predictor

The online error awareness predictor drives our developed framework for the online safety alert tool for pHRC. Three types of error awareness prediction models were developed based on the parameter being predicted:

- prediction error negativity (PEN) predictor,
- error positivity predictor and
- mean power spectral density (PSD) predictor.

Table 7.1: Within-Subject error awareness predictor - PEN Predictor

Model	mse	mae
EEGNet	0.137 ± 0.034	0.271 ± 0.096
Shallow ConvNet	0.383 ± 0.09	0.365 ± 0.023
DeepConvNet	0.301 ± 0.064	0.339 ± 0.072

Table 7.2: Within-Subject error awareness predictor - Pe Predictor

Model	mse	mae
EEGNet	0.137 ± 0.011	0.218 ± 0.035
Shallow ConvNet	0.383 ± 0.029	0.296 ± 0.073
DeepConvNet	0.301 ± 0.018	0.232 ± 0.060

Table 7.3: Within-Subject error awareness predictor - PSD Predictor

Model	mse	mae
EEGNet	0.108 ± 0.027	0.252 ± 0.021
Shallow ConvNet	0.168 ± 0.039	0.278 ± 0.036
DeepConvNet	0.189 ± 0.049	0.306 ± 0.043

7.3.1.1 Within-Subject Cross Validation

The within-subject cross-validation results for the online error awareness predictor are shown in Tables 7.1, 7.2, and 7.3. EEGNet demonstrated significantly high performance for PEN and Pe prediction in within-subject and cross-subject cross-validation. Our results show that both Shallow ConvNet and Deep ConvNet do not perform comparably well for both predictors. However, in the case of PSD predictor, we observed equally good performance by EEGNet, Shallow ConvNet, and Deep ConvNet. Therefore, it might be reliable to design our online error awareness predictor based on the prediction of power spectral density after the error.

7.3.1.2 Cross-Subject Cross Validation

The cross-subject cross-validation results are shown in Tables 7.4, 7.5 and 7.6.

Our results also demonstrate the performance of the error awareness models did not deteriorate significantly in the cross-subject cross-validation. Therefore, our online error awareness tool might be able to predict error awareness of the new subject with reasonable certainty.

Table 7.4: Cross-Subject error awareness predictor - PEN Predictor

Model	mse	mae
EEGNet	0.149 ± 0.081	0.188 ± 0.098
Shallow ConvNet	0.387 ± 0.094	0.433 ± 0.149
DeepConvNet	0.349 ± 0.110	0.391 ± 0.106

Table 7.5: Cross-Subject error awareness predictor - Pe Predictor

Model	mse	mae
EEGNet	0.166 ± 0.065	0.255 ± 0.082
Shallow ConvNet	0.390 ± 0.118	0.476 ± 0.121
DeepConvNet	0.332 ± 0.098	0.376 ± 0.111

Table 7.6: Cross-Subject error awareness predictor - PSD Predictor

Model	mse	mae
EEGNet	0.128 ± 0.081	0.263 ± 0.088
Shallow ConvNet	0.197 ± 0.067	0.302 ± 0.121
DeepConvNet	0.188 ± 0.059	0.342 ± 0.119

Measuring PEN and Pe can be challenging due to their low amplitudes of a few microvolts and the potential for distortion from artifacts introduced during active physical interaction with the robot [281]. Our results demonstrate that a simple prediction model with EEGNet could predict the amplitudes of PEN and Pe from the EEG data prior to the error. This prediction model could forewarn the system and operators of the diminished error awareness of the user, alluding to a potential safety breach in the ERrP-based BCI system for pHRC. Therefore, our work paves the way for embracing BCI technology in pHRC to optimally adapt the robot behavior for personalized user experience using real-time brain activity, enriching the quality of the interaction.

7.3.2 Online Workload Classifier

The within-subject cross-validation results for the online workload classifier are shown in Table 7.7, and the cross-subject cross-validation results are shown in Table 7.8.

Table 7.7: Within-Subject Workload Classifier

Model	Accuracy	F1-score	coh-kappa
EEGNet	73.4 ± 0.61	0.722 ± 0.005	0.571 ± 0.007
Shallow ConvNet	80.9 ± 0.421	0.801 ± 0.005	0.688 ± 0.004
DeepConvNet	81.97 ± 0.304	0.817 ± 0.002	0.682 ± 0.004

Table 7.8: Cross-Subject Workload Classifier

Model	Accuracy	F1-score	coh-kappa
EEGNet	71.2 ± 0.49	0.709 ± 0.004	0.565 ± 0.005
Shallow ConvNet	78.8 ± 0.68	0.78 ± 0.006	0.639 ± 0.007
DeepConvNet	81.3 ± 0.602	0.804 ± 0.007	0.671 ± 0.005

7.3.2.1 Within-Subject Cross Validation

The results demonstrate that Deep ConvNet was most efficient in classifying workload in physical human robot collaboration. Shallow ConvNet also demonstrates similar performance.

7.3.2.2 Cross-Subject Cross Validation

Our results also demonstrate the performance of the workload classifier did not deteriorate significantly in the cross-subject cross-validation. Therefore, our online workload classifier might be able to assess the workload of the new subject performing the blasting task with reasonable certainty.

The results of our study showed that the workload classifier had a high ability to differentiate between the three workload levels. Additionally, the error awareness predictor had a high degree of accuracy for both within-subject and cross-subject cross-validation. These findings indicate that the workload level of the operator can be accurately assessed and, when combined with an estimation of error awareness, can be used to safely interrupt the collaborative task and halt the robot's movement until the operator's cognitive state improves.

Our tool demonstrates that it is possible to use EEG data to evaluate mental workload and predict error awareness in users, and to raise safety concerns when predicted ERrP amplitudes fall below a certain threshold, making it difficult to detect the ERrP. This is critical for ensuring a safe and efficient environment for physical human-robot

collaboration. This tool has far-reaching implications, not just in robotics, but in any field that uses error-related potential brain-computer interfaces to continually monitor the user's state and determine the reliability of error-related potential amplitudes.

Therefore, this safety alert tool can be used to proactively detect when a human user's capacity to process errors has been impaired, potentially preventing a safety incident that the human user may not have noticed. By measuring PEN and Pe values and detecting workload, the tool provides an opportunity to adjust the environment and reduce the user's workload, paving the way for more effective deployment of BCI in physical human-robot collaboration scenarios.

One potential drawback of our study is that the error awareness predictor and workload classifier may be affected by perceptual motor confounds, as the models were trained using data that included motor movements such as arm and body movements used to control the robot arm. As a result, these models may not be suitable for use in classifying workload or predicting error awareness across different tasks.

7.4 Summary

In this research, we developed a framework for an online safety alert tool that can be utilized in situations where humans and robots work together physically. The tool is designed to accurately detect the operator's workload and error awareness, enabling the robot to safely stop its movement when necessary. We have demonstrated that through the use of minimal data processing and traditional deep learning models, it is possible to distinguish between different levels of workload and that EEG data can be effectively used to predict the operator's error awareness. These findings are promising, especially as the research replicates real-world scenarios of physical human-robot collaboration. This tool has the potential to enhance the safety and efficiency of error-related BCI deployment in physical human-robot collaboration by accurately detecting the operator's workload and error awareness. The cognitive workload and conflict model developed can be utilized to develop a closed-loop system that adjusts for workload responses in the human-robot collaboration environment.

Further, similar safety alert systems that identify suboptimal workload conditions and decreased error awareness could have implications for not just safety-critical work environments but also error-related potential-based BCI systems for rehabilitation. Maintaining optimal human performance is paramount in safety-critical work environments such as the military, medicine, or aviation. Therefore, deploying an online safety alert in

demanding work environments such as surgical operations, defense operations, or air traffic control, where an error could have dire consequences is crucial.

CHAPTER 8

CONCLUSION AND FUTURE WORK

8.1 Introduction

In the third chapter, we explained the reasoning and process behind our experiments. This information is important so that other scientists can replicate our work and either confirm or challenge our results. In that chapter, we discovered that both the ATC and physical human robot collaboration experiments were successful in creating different levels of workload. This was evident through the changes in participants' responses on questionnaires, their performance, and physiological measures such as EEG, eye activity, and HRV. Overall, our new approach to ATC and pHRC experiments produced consistent results when it comes to inducing variations in workload.

In Chapter 4, we elaborate on our further investigation into physiological biomarkers of workload in tracking and collision prediction tasks. Our work considered tracking and collision prediction tasks of ATC separately and untangled the impact of workload variations experienced by operators working in a complex ATC environment. The mental workload in the tracking task was found to be positively correlated with the frontal theta power and negatively correlated with the occipital alpha power. In contrast, for the collision prediction task, the frontal theta, parietal theta, occipital delta, and theta power were positively correlated, and parietal alpha power was negatively correlated with the increases in mental workload. The pupil size, number of blinks and HRV metric, root mean square of successive difference (RMSSD), also varied significantly with the mental

workload in both these tasks in a similar manner. According to our study, variations in task demands can be observed in physiological signals such as EEG, eye movement, and heart rate variability in tasks related to air traffic control. Furthermore, the distinct brain measurements of workload changes in tasks involving tracking and predicting collisions indicate that these measurements can give us an understanding of the type of mental effort required. These findings have potential applications in the design of systems that can adapt to changes in mental workload by not only detecting when it changes but also determining what type of adaptation is necessary. Our research provides strong evidence for the feasibility of creating intelligent systems that can adjust workload for enhanced efficiency and safety in air traffic control and other fields.

In Chapter 5, we consider biomarkers of the human partner's mental workload in a physical human collaboration. We investigated human brain dynamics as an objective measure of mental workload in physical interaction with a collaborative robot to examine how mental workload impacts physical performance. We found that the arithmetic task and the physical performance deteriorated with the increasing workload during physical interaction with the robot. Furthermore, frontal theta and parietal alpha power were found to be effective in reliably assessing mental workload variations during strenuous physical interactions with the robot arm, covarying with subjective workload scores. Additionally, an increase in the perceived physical load and an increase in the alpha and beta power at central sites of the brain accompanied the increases in workload during blasting, albeit maintaining a uniform physical load across all conditions. Therefore, this chapter sheds light on how mental workload impacts physical performance in the physical human-robot collaboration.

In Chapter 6, we explore how mental workload variation influences cognitive conflict or error awareness as a higher workload on the user compromises their access to cognitive resources needed for error awareness. We observed an increased frontal theta and parietal alpha power and a decreased central alpha and beta power upon reaching the invisible obstacle (error condition) as compared to reaching the target (normal condition). We also observed diminished amplitude for prediction error negativity (PEN) and error positivity (Pe) in the ERrP, indicating reduced error awareness with increasing workload. Our results demonstrate that the diminished error awareness of the user at high workload conditions might allude to a potential safety breach in an ERrP-based BCI system for pHRC. Therefore, this chapter impresses on the need to maintain optimal workload conditions while using ERrP-based BCI technology in pHRC to optimally adapt the robot behavior for personalized user experience using real-time brain activity,

enriching the quality of the interaction.

In Chapter 7, we describe the design of a safety alert tool that raises an alarm and stops the functioning of the robot when the error awareness of the human operator diminishes beyond detection. Our results demonstrate that an error awareness prediction model could predict the amplitudes of PEN and Pe from the EEG data prior to the error. This chapter demonstrates that an online safety alert tool with the error awareness prediction model and workload classifier can provide an accurate assessment of the operator's state, which could forewarn the system and operators of the diminished error awareness of the user, alluding to a potential safety breach in the ERrP-based BCI system for pHRC. Therefore, our work paves the way for embracing BCI technology in pHRC to optimally adapt the robot behavior for personalized user experience using real-time brain activity, enriching the quality of the interaction.

Therefore, the effects of workload variations were examined in a basic ATC-based environment, which was a seated, controlled setting aimed at reducing noise. The effects of workload variations in a pHRC environment were also examined. The key feature of the pHRC environment is that it simulates a real-world working environment. This dataset would be valuable to the broader BCI community as there are few experiments conducted in a controlled, real-world simulation environment.

The goal of this thesis was to gain a deeper understanding of mental workload in safety-critical work environments and to create models and algorithms for workload classification and error awareness prediction that facilitate seamless physical collaboration between humans and robots when performing labor-intensive tasks in complex, unstructured environments.

The outcomes of this thesis will bring a BCI device a step closer to the goal of becoming a mainstream product. This thesis will have a significant and immediate impact on the field of assistive robotics and its real-world applications. The research will pave the way for the creation of assistive robots that can work alongside humans in industrial settings where human workers are still needed to perform physically demanding tasks, withstand large forces for extended periods, adopt uncomfortable body positions, and operate in noisy and dusty conditions, for example in manufacturing, construction, mining, and health. The practical applications of this research are expected to have a major positive economic impact on society, including increased productivity, lower costs associated with workplace injuries, support for the workforce, and new business opportunities.

8.2 Key Contributions

- We presented two paradigms (one stationary and another dynamic) that successfully elicited workload variations, which can be used in the future for more translatable BCI research and classifier training.
- We found that the EEG measures of workload variation in tracking and collision predictions are unique, enabling the development of future mental workload adaptive systems that would know not just "when" to adapt but also "what" to adapt.
- We also found a reliable biomarker for workload variations during blasting operations involving strenuous physical activity.
- We determined that the physical performance during the blasting operation can be predicted from EEG measures of mental workload during blasting.
- We found a diminished error awareness (decreased prediction error negativity and error positivity amplitudes) with the increasing workload during blasting operation in a physical human robot collaboration task.
- We also developed a safety alert tool for pHRC work environment based on our findings that the EEG data prior to an error could possibly predict the error processing ability of the human partner.

8.3 Summary and Future Works

In this thesis, we have provided an overview of our research, detailing the methods employed in our studies and discussing the outcomes and potential future applications. We have introduced two paradigms, one stationary and one dynamic, which effectively elicit variations in workload, and can be utilized in the development of classifier training and closed-loop mental workload adaptive systems for safety-critical work environments. Our findings indicate that EEG measures of workload variation in tracking and collision prediction tasks are distinct, allowing for the creation of mental workload adaptive systems that can not only identify when to adapt but also what to adapt. These results were also used to design a system for alerting air traffic controllers of potential performance drops.

In this research, a dependable biomarker for variations in workload during human-robot collaboration was identified. Our findings revealed that as workload increases during a blasting operation task, there is a reduction in error awareness (as evidenced by decreased amplitudes in prediction error negativity and error positivity). Additionally, based on our results, we developed a safety alert system for human-robot collaboration work environments. Our data suggests that the EEG readings taken prior to an error can potentially predict the human partner's ability to process the error.

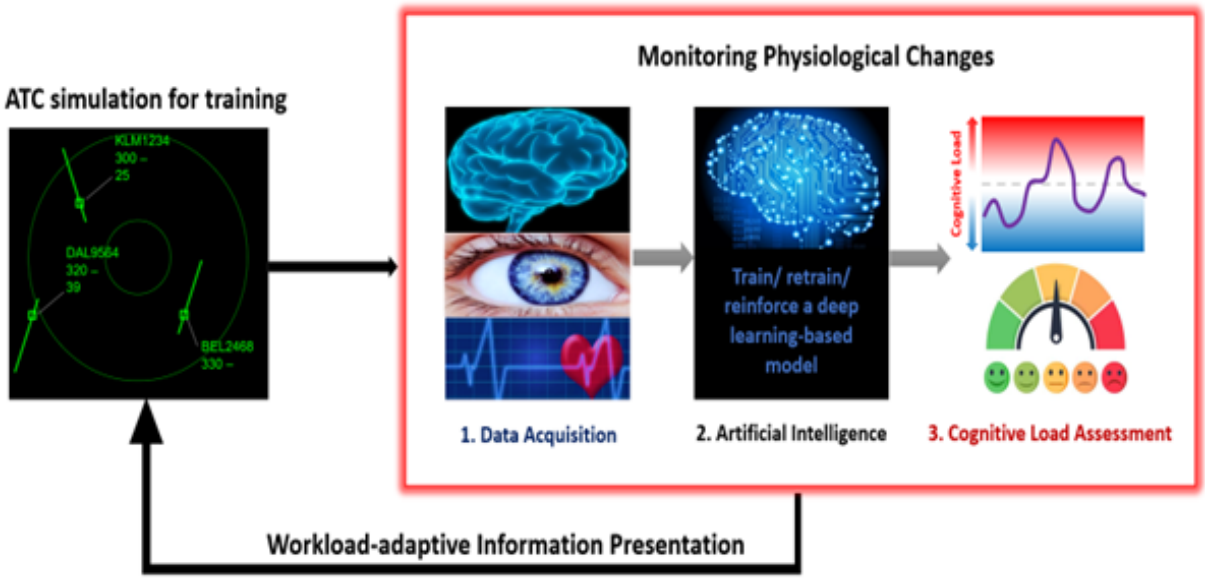

Figure 8.1: A schematic diagram showing the closed-loop mental workload adaptive system for ATC tasks.

We anticipate that future research on the correlation between task type and mental workload biomarkers will produce noteworthy discoveries about the efficiency of various mental workload adaptive approaches. While previous studies have extensively examined specific paradigms, there is a scarcity of research that compares the different paradigms and evaluates the success of each in inducing mental workload variations. In this thesis work, we have not succeeded in developing mental workload models or safety systems that are applicable across platforms. We urge future researchers to develop workload and error models that are task-independent.

In the realm of air traffic control, we can apply EEG measures to recognize when an individual is at risk of poor performance and develop an adaptive brain-computer interface that modifies task demands, allocation, or environmental settings to maintain high performance in air traffic control (as illustrated in Figure 8.1).

Another potential application is a mental workload adaptive system for human-robot collaboration, building upon the safety alert system outlined in Chapter 7. A visual representation of a mental workload-adaptive brain-robot interface for pHRC is presented in Figure 8.2. The cognitive workload and conflict model developed can be utilized to develop a closed-loop system that adjusts for workload responses in the human-robot collaboration environment. This model will have three main objectives: to identify suboptimal workload conditions, decrease error awareness, and perform prediction error negativity detection, in order to adapt the robot's actions in the environment to optimize the performance of the human-robot collaboration system.

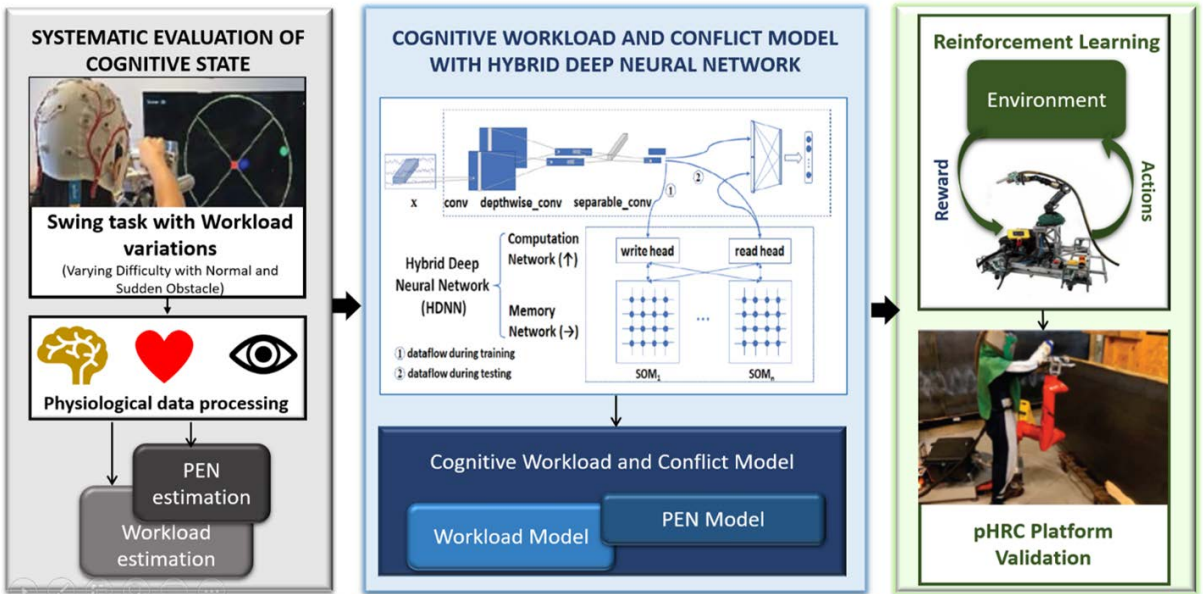

Figure 8.2: A schematic diagram showing the closed-loop mental workload adaptive system for pHRC.

As a result, the closed-loop, real-time mental workload adaptive system for human-robot collaboration will involve adapting the robot's actions in the environment to match the workload and intentions of the human operator (as illustrated in Figure 8.2). The system's efficacy can be evaluated by observing the precision of workload classification, prediction of error awareness, and detection of error-related potentials and determining whether the adaptive strategies implemented by the robot are successful in restoring workload to an optimal level and resolving cognitive conflict experienced by the operator.

We suggest that future brain-computer interface research utilizing error-related potentials could sustain optimal workload conditions to effectively employ error processing to adjust system or robot behavior. The findings from the air traffic control and

human-robot collaboration experiments can be further utilized in adaptive or closed-loop BCI systems to make real-time adjustments to the EEG signal by either modulating the signal or the user's cognitive state. In general, there are numerous possibilities for future research by integrating and combining the findings presented in this thesis.

BIBLIOGRAPHY

[1] H. A. ABBASS, J. TANG, M. ELLEJMI, AND S. KIRBY, *Visual and auditory reaction time for air traffic controllers using quantitative electroencephalograph (qeeg) data*, Brain informatics, 1 (2014), pp. 39–45.

[2] A. ABROR AND D. PATRISIA, *Psychological safety and organisational performance: a systematic literature review*, Personality and Social Psychology Review, 16 (2020), pp. 7–21.

[3] U. AHLSTROM, *An eye for the air traffic controller workload*, in Journal of the Transportation Research Forum, vol. 46, 2010.

[4] T. AIDA, H. NOZAKI, AND H. KOBAYASHI, *Development of muscle suit and application to factory laborers*, in 2009 International Conference on Mechatronics and Automation, IEEE, 2009, pp. 1027–1032.

[5] S. ALDINI, A. AKELLA, A. K. SINGH, Y.-K. WANG, M. CARMICHAEL, D. LIU, AND C.-T. LIN, *Effect of mechanical resistance on cognitive conflict in physical human-robot collaboration*, in 2019 international conference on robotics and automation (ICRA), IEEE, 2019, pp. 6137–6143.

[6] S. ALDINI, A. K. SINGH, D. LEONG, Y.-K. WANG, M. G. CARMICHAEL, D. LIU, AND C.-T. LIN, *Detection and estimation of cognitive conflict during physical human-robot collaboration*, IEEE Transactions on Cognitive and Developmental Systems, (2022).

[7] D. P. ALLEN AND C. D. MACKINNON, *Time–frequency analysis of movement-related spectral power in eeg during repetitive movements: A comparison of methods*, Journal of neuroscience methods, 186 (2010), pp. 107–115.

[8] F. ALOISE, F. SCHETTINI, P. ARICÒ, L. BIANCHI, A. RICCIO, M. MECELLA, F. BABILONI, D. MATTIA, AND F. CINCOTTI, *Advanced brain computer interface for*

communication and control, in Proceedings of the International Conference on Advanced Visual Interfaces, 2010, pp. 399–400.

[9] C. W. ANDERSON, E. A. STOLZ, AND S. SHAMSUNDER, *Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks*, IEEE Transactions on Biomedical Engineering, 45 (1998), pp. 277–286.

[10] E. M. ARGYLE, A. MARINESCU, M. L. WILSON, G. LAWSON, AND S. SHARPLES, *Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments*, International Journal of Human-Computer Studies, 145 (2021), p. 102522.

[11] P. ARICÒ, G. BORGHINI, G. DI FLUMERI, AND F. BABILONI, *Metodo per stimare uno stato mentale, in particolare un carico di lavoro, e relativo apparato (a method for the estimation of mental state, in particular of the mental workload and its device)*, P1108IT00, (2015).

[12] P. ARICÒ, G. BORGHINI, G. DI FLUMERI, S. BONELLI, A. GOLFETTI, I. GRAZIANI, S. POZZI, J.-P. IMBERT, G. GRANGER, R. BENHACENE, ET AL., *Human factors and neurophysiological metrics in air traffic control: a critical review*, IEEE reviews in biomedical engineering, 10 (2017), pp. 250–263.

[13] P. ARICÒ, G. BORGHINI, G. DI FLUMERI, A. COLOSIMO, S. BONELLI, A. GOLFETTI, S. POZZI, J.-P. IMBERT, G. GRANGER, R. BENHACENE, ET AL., *Adaptive automation triggered by eeg-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment*, Frontiers in human neuroscience, 10 (2016), p. 539.

[14] P. ARICÒ, G. BORGHINI, G. DI FLUMERI, A. COLOSIMO, S. POZZI, AND F. BABILONI, *A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks*, Progress in brain research, 228 (2016), pp. 295–328.

[15] P. ARICÒ, G. BORGHINI, I. GRAZIANI, J.-P. IMBERT, G. GRANGER, R. BENHACENE, S. POZZI, L. NAPOLETANO, G. DI FLUMERI, A. COLOSIMO, ET AL., *Air-traffic-controllers (atco): neurophysiological analysis of training and workload*, Italian Journal of Aerospace Medicine, (2015), p. 35.

[16] P. ARICÒ, M. REYNAL, G. DI FLUMERI, G. BORGHINI, N. SCIARAFFA, J.-P. IMBERT, C. HURTER, M. TERENZI, A. FERREIRA, S. POZZI, ET AL., *How neurophysiological measures can be used to enhance the evaluation of remote tower solutions*, Frontiers in Human Neuroscience, 13 (2019), p. 303.

[17] X. ARTUSI, I. K. NIAZI, M.-F. LUCAS, AND D. FARINA, *Performance of a simulated adaptive bci based on experimental classification of movement-related and error potentials*, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 1 (2011), pp. 480–488.

[18] I. ASIMOV, *Runaround*, Astounding science fiction, 29 (1942), pp. 94–103.

[19] S. ATHÈNES, P. AVERTY, S. PUECHMOREL, D. DELAHAYE, AND C. COLLET, *Atc complexity and controller workload: Trying to bridge the gap*, in Proceedings of the International Conference on HCI in Aeronautics, AAAI Press Cambridge, MA, 2002, pp. 56–60.

[20] P. AVERTY, S. ATHENES, C. COLLET, AND A. DITTMAR, *Evaluating a new index of mental workload in real atc situation using psychophysiological measures*, in Proceedings. The 21st Digital Avionics Systems Conference, vol. 2, IEEE, 2002, pp. 7A4–7A4.

[21] H. AYAZ, B. WILLEMS, S. BUNCE, P. A. SHEWOKIS, K. IZZETOGLU, S. HAH, A. DESHMUKH, AND B. ONARAL, *Estimation of cognitive workload during simulated air traffic control using optical brain imaging sensors*, in International conference on foundations of augmented cognition, Springer, 2011, pp. 549–558.

[22] F. J. BADESA, R. MORALES, N. GARCIA-ARACIL, J. M. SABATER, C. PEREZ-VIDAL, AND E. FERNANDEZ, *Multimodal interfaces to improve therapeutic outcomes in robot-assisted rehabilitation*, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42 (2012), pp. 1152–1158.

[23] C. L. BALDWIN AND B. PENARANDA, *Adaptive training using an artificial neural network and eeg metrics for within-and cross-task workload classification*, NeuroImage, 59 (2012), pp. 48–56.

[24] A. BASHASHATI, M. FATOURECHI, R. K. WARD, AND G. E. BIRCH, *A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals*, Journal of Neural engineering, 4 (2007), p. R32.

[25] T. BASTICK, *Intuition: How we think and act*, John Wiley & Sons Incorporated, 1982.

[26] I. BATMAZ AND M. OZTURK, *Using pupil diameter changes for measuring mental workload under mental processing*, Journal of Applied Sciences, 8 (2008), pp. 68–76.

[27] J. BEATTY, *Learned regulation of alpha and theta frequency activity in the human electroencephalogram*, New York: Academic Press, 1977.

[28] ——, *Task-evoked pupillary responses, processing load, and the structure of processing resources.*, Psychological bulletin, 91 (1982), p. 276.

[29] J. BEATTY AND J. F. O,ÄôHANLON, *Operant control of posterior theta rhythms and vigilance performance: Repeated treatments and transfer of training*, Progress in operant conditioning of physiological events, (1979), pp. 235–246.

[30] N. BECHTEREVA AND V. GRETCHIN, *Physiological foundations of mental activity*, International review of neurobiology, 11 (1969), pp. 329–352.

[31] A. J. BELL AND T. J. SEJNOWSKI, *An information-maximization approach to blind separation and blind deconvolution*, Neural computation, 7 (1995), pp. 1129–1159.

[32] C. BERKA, D. J. LEVENDOWSKI, C. K. RAMSEY, G. DAVIS, M. N. LUMICAO, K. STANNEY, L. REEVES, S. H. REGLI, P. D. TREMOULET, AND K. STIBLER, *Evaluation of an eeg workload model in an aegis simulation environment*, in Biomonitoring for physiological and cognitive performance during military operations, vol. 5797, SPIE, 2005, pp. 90–99.

[33] K. A. BERNHARDT, D. POLTAVSKI, T. PETROS, F. R. FERRARO, T. JORGENSON, C. CARLSON, P. DRECHSEL, AND C. ISEMINGER, *The effects of dynamic workload and experience on commercially available eeg cognitive state metrics in a high-fidelity air traffic control environment*, Applied ergonomics, 77 (2019), pp. 83–91.

[34] C. L. BETHEL, K. SALOMON, R. R. MURPHY, AND J. L. BURKE, *Survey of psychophysiology measurements applied to human-robot interaction*, in RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2007, pp. 732–737.

[35] A. BEZERIANOS, Y. SUN, Y. CHEN, K. F. WOONG, F. TAYA, P. ARICO, G. BORGHINI, F. BABILONI, AND N. THAKOR, *Cooperation driven coherence: Brains working hard together*, in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 4696–4699.

[36] S. BHATTACHARYYA, A. KONAR, AND D. TIBAREWALA, *Motor imagery and error related potential induced position control of a robotic arm*, IEEE/CAA Journal of Automatica Sinica, 4 (2017), pp. 639–650.

[37] D. BLACKWOOD AND W. J. MUIR, *Cognitive brain potentials and their application*, The British Journal of Psychiatry, 157 (1990), pp. 96–101.

[38] B. BLANKERTZ, L. ACQUALAGNA, S. DÄHNE, S. HAUFE, M. SCHULTZE-KRAFT, I. STURM, M. UŠĆUMLIC, M. A. WENZEL, G. CURIO, AND K.-R. MÜLLER, *The berlin brain-computer interface: progress beyond communication and control*, Frontiers in neuroscience, 10 (2016), p. 530.

[39] G. BORGHINI, P. ARICÒ, L. ASTOLFI, J. TOPPI, F. CINCOTTI, D. MATTIA, P. CHERUBINO, G. VECCHIATO, A. G. MAGLIONE, I. GRAZIANI, ET AL., *Frontal eeg theta changes assess the training improvements of novices in flight simulation tasks*, in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 6619–6622.

[40] G. BORGHINI, P. ARICÒ, G. DI FLUMERI, S. SALINARI, A. COLOSIMO, S. BONELLI, L. NAPOLETANO, A. FERREIRA, AND F. BABILONI, *Avionic technology testing by using a cognitive neurometric index: a study with professional helicopter pilots*, in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 6182–6185.

[41] G. BORGHINI, L. ASTOLFI, G. VECCHIATO, D. MATTIA, AND F. BABILONI, *Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness*, Neuroscience & Biobehavioral Reviews, 44 (2014), pp. 58–75.

[42] G. BORGHINI, R. ISABELLA, G. VECCHIATO, J. TOPPI, L. ASTOLFI, C. CALTAGIRONE, AND F. BABILONI, *Brainshield: Hreeg study of perceived pilot mental workload*, Italian journal of aerospace medicine, 5 (2011), pp. 34–47.

[43] G. BORRAGAN PEDRAZ, *Behavioural bases and functional dynamics of cognitive fatigue*, (2016).

[44] S. BRENNAN, *An experimental report on rating scale descriptor sets for the instantaneous self assessment (isa) recorder*, Portsmouth: DRA Maritime Command and Control Division. DRA Technical Memorandum (CAD5), 92017 (1992).

[45] J. B. BROOKINGS, G. F. WILSON, AND C. R. SWAIN, *Psychophysiological responses to changes in workload during simulated air traffic control*, Biological psychology, 42 (1996), pp. 361–377.

[46] E. A. BYRNE AND R. PARASURAMAN, *Psychophysiology and adaptive automation*, Biological psychology, 42 (1996), pp. 249–268.

[47] R. CABEZA AND L. NYBERG, *Imaging cognition ii: An empirical review of 275 pet and fmri studies*, Journal of cognitive neuroscience, 12 (2000), pp. 1–47.

[48] E. J. CALABRESE, *Dose-response features of neuroprotective agents: an integrative summary*, Critical reviews in toxicology, 38 (2008), pp. 253–348.

[49] ——, *Neuroscience and hormesis: overview and general findings*, Critical Reviews in Toxicology, 38 (2008), pp. 249–252.

[50] X. CAO, P. MACNAUGHTON, L. R. CADET, J. G. CEDENO-LAURENT, S. FLANIGAN, J. VALLARINO, D. DONNELLY-MCLAY, D. C. CHRISTIANI, J. D. SPENGLER, AND J. G. ALLEN, *Heart rate variability and performance of commercial airline pilots during flight simulations*, International journal of environmental research and public health, 16 (2019), p. 237.

[51] A. CAPILLA, J.-M. SCHOFFELEN, G. PATERSON, G. THUT, AND J. GROSS, *Dissociated α-band modulations in the dorsal and ventral visual pathways in visuospatial attention and perception*, Cerebral Cortex, 24 (2014), pp. 550–561.

[52] M. G. CARMICHAEL, S. ALDINI, R. KHONASTY, A. TRAN, C. REEKS, D. LIU, K. J. WALDRON, AND G. DISSANAYAKE, *The anbot: An intelligent robotic co-worker for industrial abrasive blasting*, in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2019, pp. 8026–8033.

[53] M. G. CARMICHAEL AND D. K. LIU, *Human biomechanical model based optimal design of assistive shoulder exoskeleton*, in Field and Service Robotics, Springer, 2015, pp. 245–258.

[54] J. G. CASALI AND W. W. WIERWILLE, *On the measurement of pilot perceptual workload: a comparison of assessment techniques addressing sensitivity and intrusion issues*, Ergonomics, 27 (1984), pp. 1033–1050.

[55] J. F. CAVANAGH AND M. J. FRANK, *Frontal theta as a mechanism for cognitive control*, Trends in cognitive sciences, 18 (2014), pp. 414–421.

[56] Y.-H. CHANG AND C.-H. YEH, *Human performance interfaces in air traffic control*, Applied ergonomics, 41 (2010), pp. 123–129.

[57] R. L. CHARLES AND J. NIXON, *Measuring mental workload using physiological measures: A systematic review*, Applied ergonomics, 74 (2019), pp. 221–232.

[58] G. E. CHATRIAN, E. LETTICH, AND P. L. NELSON, *Ten percent electrode system for topographic studies of spontaneous and evoked eeg activities*, American Journal of EEG technology, 25 (1985), pp. 83–92.

[59] G. CHATTERJI AND B. SRIDHAR, *Measures for air traffic controller workload prediction*, in 1st AIAA, aircraft, technology Integration, and operations Forum, 2001, p. 5242.

[60] U. CHAUDHARY, N. BIRBAUMER, AND A. RAMOS-MURGUIALDAY, *Brain–computer interfaces for communication and rehabilitation*, Nature Reviews Neurology, 12 (2016), pp. 513–525.

[61] R. CHAVARRIAGA, A. BIASIUCCI, K. FÖRSTER, D. ROGGEN, G. TRÖSTER, AND J. D. R. MILLÁN, *Adaptation of hybrid human-computer interaction systems using eeg error-related potentials*, in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 4226–4229.

[62] R. CHAVARRIAGA, A. SOBOLEWSKI, AND J. D. R. MILLÁN, *Errare machinale est: the use of error-related potentials in brain-machine interfaces*, Frontiers in neuroscience, (2014), p. 208.

[63] F. CHEN, J. ZHOU, Y. WANG, K. YU, S. Z. ARSHAD, A. KHAWAJI, AND D. CONWAY, *Robust multimodal cognitive load measurement*, Springer, 2016.

[64] A. CHERUBINI, R. PASSAMA, A. CROSNIER, A. LASNIER, AND P. FRAISSE, *Collaborative manufacturing with physical human–robot interaction*, Robotics and Computer-Integrated Manufacturing, 40 (2016), pp. 1–13.

[65] J. C. CHRISTENSEN AND J. R. ESTEPP, *Coadaptive aiding and automation enhance operator performance*, Human factors, 55 (2013), pp. 965–975.

[66] B. CINAZ, B. ARNRICH, R. LA MARCA, AND G. TRÖSTER, *Monitoring of mental workload levels during an everyday life office-work scenario*, Personal and ubiquitous computing, 17 (2013), pp. 229–239.

[67] B. CINAZ, R. LA MARCA, B. ARNRICH, G. TRÖSTER, ET AL., *Monitoring of mental workload levels*, (2010).

[68] M. X. COHEN, *Analyzing neural time series data: theory and practice*, MIT press, 2014.

[69] E. COLGATE, A. BICCHI, M. A. PESHKIN, AND J. E. COLGATE, *Safety for physical human-robot interaction*, in Springer handbook of robotics, Springer, 2008, pp. 1335–1348.

[70] J. E. COLGATE, J. EDWARD, M. A. PESHKIN, AND W. WANNASUPHOPRASIT, *Cobots: Robots for collaboration with human operators*, (1996).

[71] J. R. COMSTOCK JR AND R. J. ARNEGARD, *The multi-attribute task battery for human operator workload and strategic behavior research*, tech. rep., 1992.

[72] N. COPSEY, L. DRUPSTEEN, J. V. KAMPEN, L. KUIJT-EVERS, E. SCHMITZ-FELTEN, AND M. VERJANS, *A review of accidents and injuries to road transport drivers*, (2010).

[73] P. L. CRAVEN, N. BELOV, P. TREMOULET, M. THOMAS, C. BERKA, D. LEVENDOWSKI, AND G. DAVIS, *Cognitive workload gauge development: comparison of real-time classification methods*, Foundations of Augmented Cognition, (2006), pp. 75–84.

[74] D. N. CRUTCHFIELD, P. M. MOERTL, D. OHRT, ET AL., *Expertise and chess: A pilot study comparing situation awareness methodologies*.

[75] A. CRUZ, G. PIRES, AND U. J. NUNES, *Double errp detection for automatic error correction in an erp-based bci speller*, IEEE transactions on neural systems and rehabilitation engineering, 26 (2017), pp. 26–36.

[76] L. CULLEN, *Validation of a methodology for predicting performance and workload*, Eurocontrol Experimental Center, France, (1999).

[77] M. L. CUMMINGS AND C. TSONIS, *Deconstructing complexity in air traffic control*, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 49, Sage Publications Sage CA: Los Angeles, CA, 2005, pp. 25–29.

[78] Z. DAI, J. C. PRÍNCIPE, A. BEZERIANOS, AND N. V. THAKOR, *Cognitive workload discrimination in flight simulation task using a generalized measure of association*, in International Conference on Neural Information Processing, Springer, 2015, pp. 692–699.

[79] G. L. DANIELS AND K. M. NEWELL, *Attentional focus influences the walk–run transition in human locomotion*, Biological psychology, 63 (2003), pp. 163–178.

[80] A. DE LUCA AND F. FLACCO, *Integrated control for phri: Collision avoidance, detection, reaction and collaboration*, in 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), IEEE, 2012, pp. 288–295.

[81] E. DE VISSER AND R. PARASURAMAN, *Adaptive aiding of human-robot teaming: Effects of imperfect automation on performance, trust, and workload*, Journal of Cognitive Engineering and Decision Making, 5 (2011), pp. 209–231.

[82] D. DE WAARD AND K. BROOKHUIS, *The measurement of drivers' mental workload*, (1996).

[83] D. DE WAARD AND B. LEWIS-EVANS, *Self-report scales alone cannot capture mental workload*, Cognition, technology & work, 16 (2014), pp. 303–305.

[84] A. DELORME AND S. MAKEIG, *Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis*, Journal of neuroscience methods, 134 (2004), pp. 9–21.

[85] A. DELORME, T. MULLEN, C. KOTHE, Z. AKALIN ACAR, N. BIGDELY-SHAMLO, A. VANKOV, AND S. MAKEIG, *Eeglab, sift, nft, bcilab, and erica: new tools for*

advanced eeg processing, Computational intelligence and neuroscience, 2011 (2011).

[86] A. S. DEO AND I. D. WALKER, *Overview of damped least-squares methods for inverse kinematics of robot manipulators*, Journal of Intelligent and Robotic Systems, 14 (1995), pp. 43–68.

[87] G. DI FLUMERI, G. BORGHINI, P. ARICÒ, A. COLOSIMO, S. POZZI, S. BONELLI, A. GOLFETTI, W. KONG, AND F. BABILONI, *On the use of cognitive neurometric indexes in aeronautic and air traffic management environments*, in International Workshop on Symbiotic Interaction, Springer, 2015, pp. 45–56.

[88] G. DI FLUMERI, G. BORGHINI, P. ARICÒ, S. POZZI, S. BONELLI, A. GOLFETTI, A. COLOSIMO, AND F. BABILONI, *Mental workload evaluation of atcos during ecological atm scenarios*, Italian J. Aerosp. Med., 1 (2015), pp. 38–48.

[89] G. DI FLUMERI, G. BORGHINI, P. ARICÒ, N. SCIARAFFA, P. LANZI, S. POZZI, V. VIGNALI, C. LANTIERI, A. BICHICCHI, A. SIMONE, ET AL., *Eeg-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings*, Frontiers in human neuroscience, 12 (2018), p. 509.

[90] F. DOLCOS AND G. MCCARTHY, *Brain systems mediating cognitive interference by emotional distraction*, Journal of Neuroscience, 26 (2006), pp. 2072–2079.

[91] E. DONCHIN, D. KARIS, T. BASHORE, M. COLES, AND G. GRATTON, *Cognitive psychophysiology and human information processing*, (1986).

[92] G. DURANTIN, F. DEHAIS, AND A. DELORME, *Characterization of mind wandering using fnirs*, Frontiers in systems neuroscience, 9 (2015), p. 45.

[93] M. W. DYE, C. S. GREEN, AND D. BAVELIER, *The development of attention skills in action video game players*, Neuropsychologia, 47 (2009), pp. 1780–1789.

[94] ——, *Increasing speed of processing with action video games*, Current directions in psychological science, 18 (2009), pp. 321–326.

[95] A. C. EDMONDSON AND Z. LEI, *Psychological safety: The history, renaissance, and future of an interpersonal construct*, Annual review of organizational psychology and organizational behavior, 1 (2014), pp. 23–43.

[96] F. T. EGGEMEIER, G. F. WILSON, A. F. KRAMER, AND D. L. DAMOS, *Workload assessment in multi-task environments*, in Multiple-task performance, CRC Press, 2020, pp. 207–216.

[97] S. K. EHRLICH AND G. CHENG, *Human-agent co-adaptation using error-related potentials*, Journal of neural engineering, 15 (2018), p. 066014.

[98] ——, *A feasibility study for validating robot actions using eeg-based error-related potentials*, International journal of social robotics, 11 (2019), pp. 271–283.

[99] M. R. ENDSLEY, *Toward a theory of situation awareness in dynamic systems*, in Situational awareness, Routledge, 2017, pp. 9–42.

[100] M. R. ENDSLEY AND D. J. GARLAND, *Pilot situation awareness training in general aviation*, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 44, SAGE Publications Sage CA: Los Angeles, CA, 2000, pp. 357–360.

[101] M. R. ENDSLEY, S. J. SELCON, T. D. HARDIMAN, AND D. G. CROFT, *A comparative analysis of sagat and sart for evaluations of situation awareness*, in Proceedings of the human factors and ergonomics society annual meeting, vol. 42, SAGE Publications Sage CA: Los Angeles, CA, 1998, pp. 82–86.

[102] E. T. ESFAHANI AND V. SUNDARARAJAN, *Using brain–computer interfaces to detect human satisfaction in human–robot interaction*, International Journal of Humanoid Robotics, 8 (2011), pp. 87–101.

[103] S. FAIRCLOUGH, K. EWING, C. BURNS, AND U. KREPLIN, *Neural efficiency and mental workload: locating the red line*, in Neuroergonomics, Elsevier, 2019, pp. 73–77.

[104] S. H. FAIRCLOUGH AND L. VENABLES, *Prediction of subjective states from psychophysiology: A multivariate approach*, Biological psychology, 71 (2006), pp. 100–110.

[105] S. H. FAIRCLOUGH, L. VENABLES, AND A. TATTERSALL, *The influence of task demand and learning on the psychophysiological response*, International Journal of Psychophysiology, 56 (2005), pp. 171–184.

[106] M. FALKENSTEIN, J. HOHNSBEIN, J. HOORMANN, AND L. BLANKE, *Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks*, Electroencephalography and clinical neurophysiology, 78 (1991), pp. 447–455.

[107] M. FALKENSTEIN, J. HOORMANN, S. CHRIST, AND J. HOHNSBEIN, *Erp components on reaction errors and their functional significance: a tutorial*, Biological psychology, 51 (2000), pp. 87–107.

[108] M. FALLAHI, M. MOTAMEDZADE, R. HEIDARIMOGHADAM, A. R. SOLTANIAN, AND S. MIYAKE, *Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study*, Applied ergonomics, 52 (2016), pp. 95–103.

[109] J. FAN, J. I. FLOMBAUM, B. D. MCCANDLISS, K. M. THOMAS, AND M. I. POSNER, *Cognitive and brain consequences of conflict*, Neuroimage, 18 (2003), pp. 42–57.

[110] T. FAWCETT, *Introduction to receiver operator curves*, Pattern Recognit. Lett, 27 (2006), pp. 861–874.

[111] P. FERREZ AND J. MILLÁN, *Eeg-based brain-computer interaction: Improved accuracy by automatic single-trial error detection*, Advances in neural information processing systems, 20 (2007).

[112] P. W. FERREZ AND J. D. R. MILLÁN, *Error-related eeg potentials generated during simulated brain–computer interaction*, IEEE transactions on biomedical engineering, 55 (2008), pp. 923–929.

[113] F. FICUCIELLO, L. VILLANI, AND B. SICILIANO, *Variable impedance control of redundant manipulators for intuitive human–robot physical interaction*, IEEE Transactions on Robotics, 31 (2015), pp. 850–863.

[114] L. R. FOURNIER, G. F. WILSON, AND C. R. SWAIN, *Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training*, International Journal of Psychophysiology, 31 (1999), pp. 129–145.

[115] B. A. FRANCIS AND W. M. WONHAM, *The internal model principle of control theory*, Automatica, 12 (1976), pp. 457–465.

[116] F. G. FREEMAN, P. J. MIKULKA, L. J. PRINZEL, AND M. W. SCERBO, *Evaluation of an adaptive automation system using three eeg indices with a visual tracking task*, Biological psychology, 50 (1999), pp. 61–76.

[117] J. FRIEDENBERG, G. SILVERMAN, AND M. J. SPIVEY, *Cognitive science: an introduction to the study of mind*, Sage Publications, 2021.

[118] T. FRITZ, A. BEGEL, S. C. MÜLLER, S. YIGIT-ELLIOTT, AND M. ZÜGER, *Using psycho-physiological measures to assess task difficulty in software development*, in Proceedings of the 36th international conference on software engineering, 2014, pp. 402–413.

[119] A. GALE, R. DAVIES, AND A. SMALLBONE, *Eeg correlates of signal rate, time in task and individual differences in reaction time during a five-stage sustained attention task*, Ergonomics, 20 (1977), pp. 363–376.

[120] E. GAMBAO, M. HERNANDO, AND D. SURDILOVIC, *A new generation of collaborative robots for material handling*, in ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, vol. 29, IAARC Publications, 2012, p. 1.

[121] T. GASSER, P. BÄCHER, AND J. MÖCKS, *Transformations towards the normal distribution of broad band spectral parameters of the eeg*, Electroencephalography and clinical neurophysiology, 53 (1982), pp. 119–124.

[122] T. GATEAU, H. AYAZ, AND F. DEHAIS, *In silico vs. over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-bci*, Frontiers in human neuroscience, 12 (2018), p. 187.

[123] W. J. GEHRING, B. GOSS, M. G. COLES, D. E. MEYER, AND E. DONCHIN, *A neural system for error detection and compensation*, Psychological science, 4 (1993), pp. 385–390.

[124] A. GEVINS AND M. E. SMITH, *Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style*, Cerebral cortex, 10 (2000), pp. 829–839.

[125] ——, *Neurophysiological measures of cognitive workload during human-computer interaction*, Theoretical issues in ergonomics science, 4 (2003), pp. 113–131.

[126] A. GEVINS, M. E. SMITH, H. LEONG, L. MCEVOY, S. WHITFIELD, R. DU, AND G. RUSH, *Monitoring working memory load during computer-based tasks with eeg pattern recognition methods*, Human factors, 40 (1998), pp. 79–91.

[127] A. S. GEVINS, S. L. BRESSLER, B. A. CUTILLO, J. ILLES, J. C. MILLER, J. STERN, AND H. R. JEX, *Effects of prolonged mental work on functional brain topography*, Electroencephalography and clinical neurophysiology, 76 (1990), pp. 339–350.

[128] D. GOPHER, *13 the skill of attention control: Acquisition*, Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience, 14 (1993), p. 299.

[129] A. GRAMFORT, M. LUESSI, E. LARSON, D. A. ENGEMANN, D. STROHMEIER, C. BRODBECK, R. GOJ, M. JAS, T. BROOKS, L. PARKKONEN, ET AL., *Meg and eeg data analysis with mne-python*, Frontiers in neuroscience, (2013), p. 267.

[130] C. S. GREEN AND D. BAVELIER, *Action-video-game experience alters the spatial resolution of vision*, Psychological science, 18 (2007), pp. 88–94.

[131] S. D. GRONLUND, D. D. OHRT, M. R. DOUGHERTY, J. L. PERRY, AND C. A. MANNING, *Role of memory in air traffic control.*, Journal of Experimental Psychology: Applied, 4 (1998), p. 263.

[132] C. R. GUERRERO, J. C. F. MARINERO, J. P. TURIEL, AND V. MUÑOZ, *Using ‚Äúhuman state aware‚Äù robots to enhance physical human–robot interaction in a cooperative scenario*, Computer methods and programs in biomedicine, 112 (2013), pp. 250–259.

[133] J. GUIOCHET, M. MACHIN, AND H. WAESELYNCK, *Safety-critical advanced robots: A survey*, Robotics and Autonomous Systems, 94 (2017), pp. 43–52.

[134] J. T. GWIN AND D. P. FERRIS, *An eeg-based study of discrete isometric and isotonic human lower limb muscle contractions*, Journal of neuroengineering and rehabilitation, 9 (2012), pp. 1–11.

[135] S. HADDADIN AND E. CROFT, *Physical human–robot interaction*, in Springer handbook of robotics, Springer, 2016, pp. 1835–1874.

[136] G. HAJCAK, J. S. MOSER, N. YEUNG, AND R. F. SIMONS, *On the ern and the significance of errors*, Psychophysiology, 42 (2005), pp. 151–160.

[137] P. A. HANCOCK, *A dynamic model of stress and sustained attention*, Human factors, 31 (1989), pp. 519–537.

[138] P. A. HANCOCK AND G. MATTHEWS, *Workload and performance: Associations, insensitivities, and dissociations*, Human factors, 61 (2019), pp. 374–392.

[139] P. A. HANCOCK AND J. S. WARM, *A dynamic model of stress and sustained attention*, Journal of Human Performance in Extreme Environments, 7 (2003), p. 4.

[140] E. HARMON-JONES AND C. HARMON-JONES, *Cognitive dissonance theory after 50 years of development*, Zeitschrift für Sozialpsychologie, 38 (2007), pp. 7–16.

[141] S. G. HART, *Nasa-task load index (nasa-tlx); 20 years later*, in Proceedings of the human factors and ergonomics society annual meeting, vol. 50, Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.

[142] S. G. HART AND L. E. STAVELAND, *Development of nasa-tlx (task load index): Results of empirical and theoretical research*, in Advances in psychology, vol. 52, Elsevier, 1988, pp. 139–183.

[143] J. A. HARTIGAN AND M. A. WONG, *Algorithm as 136: A k-means clustering algorithm*, Journal of the royal statistical society. series c (applied statistics), 28 (1979), pp. 100–108.

[144] T. HEINE, G. LENIS, P. REICHENSPERGER, T. BERAN, O. DOESSEL, AND B. DEML, *Electrocardiographic features for the measurement of drivers' mental workload*, Applied ergonomics, 61 (2017), pp. 31–43.

[145] C. HERFF, D. HEGER, O. FORTMANN, J. HENNRICH, F. PUTZE, AND T. SCHULTZ, *Mental workload during n-back task‚Äîquantified in the prefrontal cortex using fnirs*, Frontiers in human neuroscience, 7 (2014), p. 935.

[146] B. HILBURN, *Cognitive complexity in air traffic control: A literature review*, EEC note, 4 (2004), pp. 1–80.

[147] B. HILBURN, P. G. JORNA, E. A. BYRNE, AND R. PARASURAMAN, *The effect of adaptive air traffic control (atc) decision aiding on controller mental workload*, Human-automation interaction: Research and practice, (1997), pp. 84–91.

[148] J. M. HISTON, R. J. HANSMAN, B. GOTTLIEB, H. KLEINWAKS, S. YENSON, D. DELAHAYE, AND S. PUECHMOREL, *Structural considerations and cognitive complexity in air traffic control*, in Proceedings. The 21st Digital Avionics Systems Conference, vol. 1, IEEE, 2002, pp. 1C2–1C2.

[149] N. HJORTSKOV, D. RISSÉN, A. K. BLANGSTED, N. FALLENTIN, U. LUNDBERG, AND K. SØGAARD, *The effect of mental stress on heart rate variability and blood pressure during computer work*, European journal of applied physiology, 92 (2004), pp. 84–89.

[150] G. R. J. HOCKEY, P. NICKEL, A. C. ROBERTS, AND M. H. ROBERTS, *Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control*, Applied ergonomics, 40 (2009), pp. 1011–1018.

[151] D. A. HOFMANN, M. J. BURKE, AND D. ZOHAR, *100 years of occupational safety research: From basic protections and work analysis to a multilevel view of workplace safety and risk.*, Journal of applied psychology, 102 (2017), p. 375.

[152] M. A. HOGERVORST, A.-M. BROUWER, AND J. B. VAN ERP, *Combining and comparing eeg, peripheral physiology and eye-related measures for the assessment of mental workload*, Frontiers in neuroscience, 8 (2014), p. 322.

[153] D. HOLGADO, M. ZABALA, AND D. SANABRIA, *No evidence of the effect of cognitive load on self-paced cycling performance*, PloS one, 14 (2019), p. e0217825.

[154] A. HOLM, K. LUKANDER, J. KORPELA, M. SALLINEN, AND K. M. MÜLLER, *Estimating brain load from the eeg*, TheScientificWorldJOURNAL, 9 (2009), pp. 639–651.

[155] C. B. HOLROYD AND M. G. COLES, *The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity.*, Psychological review, 109 (2002), p. 679.

[156] R. J. INNES, N. J. EVANS, Z. L. HOWARD, A. EIDELS, AND S. D. BROWN, *A broader application of the detection response task to cognitive tasks and online environments*, Human factors, 63 (2021), pp. 896–909.

[157] I. ITURRATE, R. CHAVARRIAGA, L. MONTESANO, J. MINGUEZ, AND J. D. R. MILLÁN, *Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control*, Scientific reports, 5 (2015), pp. 1–10.

[158] I. ITURRATE, J. GRIZOU, J. OMEDES, P.-Y. OUDEYER, M. LOPES, AND L. MONTESANO, *Exploiting task constraints for self-calibrated brain-machine interface control using error-related potentials*, PloS one, 10 (2015), p. e0131491.

[159] I. ITURRATE, L. MONTESANO, AND J. MINGUEZ, *Robot reinforcement learning using eeg-based reward signals*, in 2010 IEEE international conference on robotics and automation, IEEE, 2010, pp. 4822–4829.

[160] J. JACOBS, G. HWANG, T. CURRAN, AND M. J. KAHANA, *Eeg oscillations and recognition memory: theta correlates of memory retrieval and decision making*, Neuroimage, 32 (2006), pp. 978–987.

[161] H. F. JELINEK, K. G. AUGUST, M. H. IMAM, A. H. KHANDOKER, A. KOENIG, AND R. RIENER, *Cortical response to psycho-physiological changes in auto-adaptive robot assisted gait training*, in 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2011, pp. 7409–7412.

[162] R. JENKE, A. PEER, AND M. BUSS, *Feature extraction and selection for emotion recognition from eeg*, IEEE Transactions on Affective computing, 5 (2014), pp. 327–339.

[163] O. JENSEN AND C. D. TESCHE, *Frontal theta activity in humans increases with memory load in a working memory task*, European journal of Neuroscience, 15 (2002), pp. 1395–1399.

[164] J.-H. JEONG, K.-T. KIM, Y.-D. YUN, AND S.-W. LEE, *Design of a brain-controlled robot arm system based on upper-limb movement imagery*, in 2018 6th International Conference on Brain-Computer Interface (BCI), IEEE, 2018, pp. 1–3.

[165] A. JIMENEZ-MOLINA, C. RETAMAL, AND H. LIRA, *Using psychophysiological sensors to assess mental workload during web browsing*, Sensors, 18 (2018), p. 458.

[166] A. R. JOHN, A. K. SINGH, T.-T. N. DO, A. EIDELS, E. NALIVAIKO, A. M. GAVGANI, S. BROWN, M. BENNETT, S. LAL, A. M. SIMPSON, ET AL., *Unraveling the physiological correlates of mental workload variations in tracking and collision*

prediction tasks, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30 (2022), pp. 770–781.

[167] C. JORDAN, *Experimental study of the effect of an instantaneous self assessment workload recorder on task performance*, Report No. DRA/TM (CAD5)/92011. Farnborough: Defence Evaluation & Research Agency, (1992).

[168] D. B. KABER AND M. R. ENDSLEY, *The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task*, Theoretical issues in ergonomics science, 5 (2004), pp. 113–153.

[169] D. KAHNEMAN, *Attention and effort*, vol. 1063, Citeseer, 1973.

[170] F. P. KALAGANIS, E. CHATZILARI, S. NIKOLOPOULOS, I. KOMPATSIARIS, AND N. A. LASKARIS, *An error-aware gaze-based keyboard by means of a hybrid bci system*, Scientific reports, 8 (2018), pp. 1–11.

[171] M. V. KAMATH AND E. L. FALLEN, *Power spectral analysis of heart rate variability: a noninvasive signature of cardiac autonomic function.*, Critical reviews in biomedical engineering, 21 (1993), pp. 245–311.

[172] M. KASSNER, W. PATERA, AND A. BULLING, *Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction*, in Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication, 2014, pp. 1151–1160.

[173] Y. KE, H. QI, F. HE, S. LIU, X. ZHAO, P. ZHOU, L. ZHANG, AND D. MING, *An eeg-based mental workload estimator trained on working memory task can work well under simulated multi-attribute task*, Frontiers in human neuroscience, 8 (2014), p. 703.

[174] P. KEARNEY, W.-C. LI, C.-S. YU, AND G. BRAITHWAITE, *The impact of alerting designs on air traffic controller‚Äôs eye movement patterns and situation awareness*, Ergonomics, 62 (2019), pp. 305–318.

[175] S. K. KIM AND E. A. KIRCHNER, *Classifier transferability in the detection of error related potentials from observation to interaction*, in 2013 IEEE international conference on systems, man, and cybernetics, IEEE, 2013, pp. 3360–3365.

[176] ——, *Handling few training data: classifier transfer between different types of error-related potentials*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24 (2015), pp. 320–332.

[177] S. K. KIM, E. A. KIRCHNER, A. STEFES, AND F. KIRCHNER, *Intrinsic interactive reinforcement learning–using error-related potentials for real world human-robot interaction*, Scientific reports, 7 (2017), pp. 1–16.

[178] E. A. KIRCHNER, S. K. KIM, M. TABIE, H. WÖHRLE, M. MAURUS, AND F. KIRCHNER, *An intelligent man-machine interface‚Äîmulti-robot control adapted for task engagement based on single-trial detectability of p300*, Frontiers in human neuroscience, 10 (2016), p. 291.

[179] D. KIRSH, *A few thoughts on cognitive overload*, Intellectica, 1 (2000).

[180] B. KIRWAN, A. EVANS, L. DONOHOE, A. KILNER, T. LAMOUREUX, T. ATKINSON, AND H. MACKENDRICK, *Human factors in the atm system design life cycle*, in FAA/Eurocontrol ATM R&D Seminar, 1997, pp. 16–20.

[181] B. KIRWAN, R. SCAIFE, AND R. KENNEDY, *Investigating complexity factors in uk air traffic management*, Human Factors and Aerospace Safety, 1 (2001).

[182] W. KLIMESCH, *Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis*, Brain research reviews, 29 (1999), pp. 169–195.

[183] W. KLIMESCH, M. DOPPELMAYR, H. RUSSEGGER, T. PACHINGER, AND J. SCHWAIGER, *Induced alpha band power changes in the human eeg and attention*, Neuroscience letters, 244 (1998), pp. 73–76.

[184] A. KOENIG, D. NOVAK, X. OMLIN, M. PULFER, E. PERREAULT, L. ZIMMERLI, M. MIHELJ, AND R. RIENER, *Real-time closed-loop control of cognitive load in neurological patients during robot-assisted gait training*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 19 (2011), pp. 453–464.

[185] J. KOHLMORGEN, G. DORNHEGE, M. BRAUN, B. BLANKERTZ, K.-R. MÜLLER, G. CURIO, K. HAGEMANN, A. BRUNS, M. SCHRAUF, W. KINCSES, ET AL., *Improving human performance in a real operating environment through real-time mental workload detection*, Toward brain-computer interfacing, 409422 (2007), pp. 409–422.

[186] A. KOK, *On the utility of p3 amplitude as a measure of processing capacity*, Psychophysiology, 38 (2001), pp. 557–577.

[187] M. A. KOMPIER AND T. S. KRISTENSEN, *Organizational work stress interventions in a theoretical, methodological and practical context.*, (2001).

[188] B. KOPP, F. RIST, AND U. MATTLER, *N200 in the flanker task as a neurobehavioral tool for investigating executive control*, Psychophysiology, 33 (1996), pp. 282–294.

[189] T. KOSCH, M. HASSIB, D. BUSCHEK, AND A. SCHMIDT, *Look into my eyes: using pupil dilation to estimate mental workload for task complexity adaptation*, in Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, 2018, pp. 1–6.

[190] N. KOSTYK AND S. LEBID, *Physical workloads vs cognitive tasks performance-mobile monitoring capabilities*, in 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), IEEE, 2020, pp. 335–340.

[191] C. KOTHE ET AL., *Lab streaming layer (lsl)*, 2014.

[192] A. KRAMER AND J. SPINKS, *Capacity views of human information processing.*, (1991).

[193] W. N. KUHLMAN, *Eeg feedback training: enhancement of somatosensory cortical activity*, Electroencephalography and clinical neurophysiology, 45 (1978), pp. 290–294.

[194] D. KULIC AND E. A. CROFT, *Affective state estimation for human–robot interaction*, IEEE transactions on robotics, 23 (2007), pp. 991–1000.

[195] J. L. LANCASTER, M. G. WOLDORFF, L. M. PARSONS, M. LIOTTI, C. S. FREITAS, L. RAINEY, P. V. KOCHUNOV, D. NICKERSON, S. A. MIKITEN, AND P. T. FOX, *Automated talairach atlas labels for functional brain mapping*, Human brain mapping, 10 (2000), pp. 120–131.

[196] P. A. LANDSBERGIS, *The changing organization of work and the safety and health of working people: a commentary*, Journal of occupational and environmental medicine, (2003), pp. 61–72.

[197] P. A. LASOTA, T. FONG, J. A. SHAH, ET AL., *A survey of methods for safe human-robot interaction*, Foundations and Trends® in Robotics, 5 (2017), pp. 261–349.

[198] V. J. LAWHERN, A. J. SOLON, N. R. WAYTOWICH, S. M. GORDON, C. P. HUNG, AND B. J. LANCE, *Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces*, Journal of neural engineering, 15 (2018), p. 056013.

[199] I. LAZAROU, S. NIKOLOPOULOS, P. C. PETRANTONAKIS, I. KOMPATSIARIS, AND M. TSOLAKI, *Eeg-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: a novel approach of the 21 st century*, Frontiers in human neuroscience, 12 (2018), p. 14.

[200] C.-T. LIN, S.-A. CHEN, T.-T. CHIU, H.-Z. LIN, AND L.-W. KO, *Spatial and temporal eeg dynamics of dual-task driving performance*, Journal of neuroengineering and rehabilitation, 8 (2011), pp. 1–13.

[201] C.-T. LIN, C.-Y. CHIU, A. K. SINGH, J.-T. KING, L.-W. KO, Y.-C. LU, AND Y.-K. WANG, *A wireless multifunctional ssvep-based brain–computer interface assistive system*, IEEE Transactions on Cognitive and Developmental Systems, 11 (2018), pp. 375–383.

[202] J. L. LOBO, J. D. SER, F. DE SIMONE, R. PRESTA, S. COLLINA, AND Z. MORAVEK, *Cognitive workload classification using eye-tracking and eeg data*, in Proceedings of the International Conference on Human-Computer Interaction in Aerospace, 2016, pp. 1–8.

[203] S. LOFT, P. SANDERSON, A. NEAL, AND M. MOOIJ, *Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications*, Human factors, 49 (2007), pp. 376–399.

[204] M. A. LOPEZ-GORDO, D. SANCHEZ-MORILLO, AND F. P. VALLE, *Dry eeg electrodes*, Sensors, 14 (2014), pp. 12847–12870.

[205] E. LOPEZ-LARRAZ, I. ITURRATE, L. MONTESANO, AND J. MINGUEZ, *Real-time recognition of feedback error-related potentials during a time-estimation task*, in 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 2670–2673.

[206] E. LÓPEZ-LARRAZ, A. SARASOLA-SANZ, N. IRASTORZA-LANDA, N. BIRBAUMER, AND A. RAMOS-MURGUIALDAY, *Brain-machine interfaces for rehabilitation in stroke: a review*, NeuroRehabilitation, 43 (2018), pp. 77–97.

[207] S. J. LUCK, *Event-related potentials.*, (2012).

[208] S. MAGNUSSON NÄHLINDER, *Flight simulator training: Assessing the potential*, PhD thesis, Linköping University Electronic Press, 2009.

[209] A. MAJUMDAR, W. Y. OCHIENG, G. MCAULEY, J. M. LENZI, AND C. LEPADATU, *The factors affecting airspace capacity in europe: A cross-sectional time-series analysis using simulated controller workload data*, The Journal of Navigation, 57 (2004), pp. 385–405.

[210] S. MAKEIG, *Auditory event-related dynamics of the eeg spectrum and effects of exposure to tones*, Electroencephalography and clinical neurophysiology, 86 (1993), pp. 283–293.

[211] ——, *Linking brain, mind, and behavior*, International Journal of Psychophysiology, 3 (2008), p. 137.

[212] B. R. MALCOLM, J. J. FOXE, J. S. BUTLER, S. MOLHOLM, AND P. DE SANCTIS, *Cognitive load reduces the effects of optic flow on gait and electrocortical dynamics during treadmill walking*, Journal of neurophysiology, 120 (2018), pp. 2246–2259.

[213] K. MANDRICK, V. PEYSAKHOVICH, F. RÉMY, E. LEPRON, AND M. CAUSSE, *Neural and psychophysiological correlates of human performance under stress and high mental workload*, Biological psychology, 121 (2016), pp. 62–73.

[214] X. MAO, M. LI, W. LI, L. NIU, B. XIAN, M. ZENG, AND G. CHEN, *Progress in eeg-based brain robot interaction systems*, Computational intelligence and neuroscience, 2017 (2017).

[215] A. C. MARINESCU, S. SHARPLES, A. C. RITCHIE, T. SANCHEZ LOPEZ, M. MCDOWELL, AND H. P. MORVAN, *Physiological parameter response to variation of mental workload*, Human factors, 60 (2018), pp. 31–56.

[216] G. MARQUART, C. CABRALL, AND J. DE WINTER, *Review of eye-related measures of drivers‚Äô mental workload*, Procedia Manufacturing, 3 (2015), pp. 2854–2861.

[217] G. MARQUART AND J. DE WINTER, *Workload assessment for mental arithmetic tasks using the task-evoked pupillary response*, PeerJ Computer Science, 1 (2015), p. e16.

[218] L. MASIA, H. KREBS, P. CAPPA, AND N. HOGAN, *Whole-arm rehabilitation following stroke: Hand module*, in The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, 2006. BioRob 2006., IEEE, 2006, pp. 1085–1089.

[219] M. MATOUSEK AND I. PETERSÉN, *A method for assessing alertness fluctuations from eeg spectra*, Electroencephalography and clinical neurophysiology, 55 (1983), pp. 108–113.

[220] G. MATTHEWS, S. E. CAMPBELL, S. FALCONER, L. A. JOYNER, J. HUGGINS, K. GILLILAND, R. GRIER, AND J. S. WARM, *Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry.*, Emotion, 2 (2002), p. 315.

[221] G. MATTHEWS, L. E. REINERMAN-JONES, D. J. BARBER, AND J. ABICH IV, *The psychometrics of mental workload: Multiple measures are sensitive but divergent*, Human factors, 57 (2015), pp. 125–143.

[222] G. MATTHEWS, J. S. WARM, L. E. REINERMAN, L. K. LANGHEIM, AND D. J. SAXBY, *Task engagement, attention, and executive control*, in Handbook of individual differences in cognition, Springer, 2010, pp. 205–230.

[223] L. K. MCEVOY, E. PELLOUCHOUD, M. E. SMITH, AND A. GEVINS, *Neurophysiological signals of working memory in normal aging*, Cognitive Brain Research, 11 (2001), pp. 363–376.

[224] R. MCKENDRICK, T. SHAW, E. DE VISSER, H. SAQER, B. KIDWELL, AND R. PARASURAMAN, *Team performance in networked supervisory control of unmanned air vehicles: Effects of automation, working memory, and communication content*, Human factors, 56 (2014), pp. 463–475.

[225] T. MCMORRIS, A. TURNER, B. J. HALE, AND J. SPROULE, *Beyond the catecholamines hypothesis for an acute exercise–cognition interaction: A neurochemical perspective.*, (2016).

[226] B. MEHLER, B. REIMER, AND Y. WANG, *A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload*, in Driving Assesment Conference, vol. 6, University of Iowa, 2011.

[227] E. MEISNER, V. ISLER, AND J. TRINKLE, *Controller design for human-robot interaction*, Autonomous Robots, 24 (2008), pp. 123–134.

[228] A. H. MEMAR AND E. T. ESFAHANI, *Eeg correlates of motor control difficulty in physical human-robot interaction: A frequency domain analysis*, in 2018 IEEE Haptics Symposium (HAPTICS), IEEE, 2018, pp. 229–234.

[229] ——, *Objective assessment of human workload in physical human-robot cooperation using brain monitoring*, ACM Transactions on Human-Robot Interaction (THRI), 9 (2019), pp. 1–21.

[230] T. MILEKOVIC, T. BALL, A. SCHULZE-BONHAGE, A. AERTSEN, AND C. MEHRING, *Detection of error related neuronal responses recorded by electrocorticography in humans during continuous movements*, PloS one, 8 (2013), p. e55235.

[231] J. MISHRA, J. A. ANGUERA, D. A. ZIEGLER, AND A. GAZZALEY, *A cognitive framework for understanding and improving interference resolution in the brain*, Progress in brain research, 207 (2013), pp. 351–377.

[232] S. MIXTER, S. E. MATHIASSEN, AND D. HALLMAN, *Alternations between physical and cognitive tasks in repetitive work–effect of cognitive task difficulty on fatigue development in women*, Ergonomics, 62 (2019), pp. 1008–1022.

[233] H. N. MODI, H. SINGH, G.-Z. YANG, A. DARZI, AND D. R. LEFF, *A decade of imaging surgeons' brain function (part i): Terminology, techniques, and clinical translation*, Surgery, 162 (2017), pp. 1121–1130.

[234] R. H. MOGFORD, J. GUTTMAN, S. MORROW, AND P. KOPARDEKAR, *The complexity construct in air traffic control: A review and synthesis of the literature.*, (1995).

[235] D. D. MONTGOMERY, J. ROBB, K. V. DWYER, AND S. T. GONTKOVSKY, *Single channel qeeg amplitudes in a bright, normal young adult sample*, Journal of Neurotherapy, 2 (1998), pp. 1–7.

[236] K. MOORE AND L. GUGERTY, *Development of a novel measure of situation awareness: The case for eye movement analysis*, in Proceedings of the human factors and ergonomics society annual meeting, vol. 54, SAGE Publications Sage CA: Los Angeles, CA, 2010, pp. 1650–1654.

[237] M. E. MORAN, *Evolution of robotic arms*, Journal of robotic surgery, 1 (2007), pp. 103–111.

[238] E. MOWER, D. J. FEIL-SEIFER, M. J. MATARIC, AND S. NARAYANAN, *Investigating implicit cues for user state estimation in human-robot interaction using physiological measurements*, in RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2007, pp. 1125–1130.

[239] K. MURAI, Y. HAYASHI, AND S. INOKUCHI, *A basic study on teammates' mental workload among ship's bridge team*, IEICE TRANSACTIONS on Information and Systems, 87 (2004), pp. 1477–1483.

[240] J. M. MURKIN AND M. ARANGO, *Near-infrared spectroscopy as an index of brain and tissue oxygenation*, British journal of anaesthesia, 103 (2009), pp. i3–i13.

[241] A. MURPHY, D. KAEGI, R. GOBBLE, A. DUBIN, S. HOWARD, D. GABA, Y. SOWB, AND L. HALAMEK, *245 validation of simulation-based training in neonatal resuscitation: Use of heart rate variability as a marker for mental workload.*, 2004.

[242] S. NÄHLINDER, *ACES-Air combat evaluation system*, Command and Control Systems, Swedish Defence Research Agency, 2004.

[243] B. NAVARRO, A. FONTE, P. FRAISSE, G. POISSON, AND A. CHERUBINI, *Openphri: an open-source library for safe physical human-robot interaction*, IEEE Robotics and Automation Magazine, (2017).

[244] L. H. NEGRI, *Peakutils*, URL: https://pypi. python. org/pypi/PeakUtils (visited on 07/20/2016), (2018).

[245] P. NICKEL AND F. NACHREINER, *Sensitivity and diagnosticity of the 0.1-hz component of heart rate variability as an indicator of mental workload*, Human factors, 45 (2003), pp. 575–590.

[246] E. NIEDERMEYER AND F. L. DA SILVA, *Electroencephalography: basic principles, clinical applications, and related fields*, Lippincott Williams & Wilkins, 2005.

[247] M. NIJBOER, J. BORST, H. VAN RIJN, AND N. TAATGEN, *Single-task fmri overlap predicts concurrent multitasking interference*, NeuroImage, 100 (2014), pp. 60–74.

[248] N. NIOSH, *The changing organization of work and the safety and health of working people*, 2002.

[249] D. A. NORMAN AND D. G. BOBROW, *On data-limited and resource-limited processes*, Cognitive psychology, 7 (1975), pp. 44–64.

[250] D. NOVAK, B. BEYELER, X. OMLIN, AND R. RIENER, *Workload estimation in physical human–robot interaction using physiological measurements*, Interacting with computers, 27 (2015), pp. 616–629.

[251] D. NOVAK, M. MIHELJ, J. ZIHERL, A. OLENSEK, AND M. MUNIH, *Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 19 (2011), pp. 400–410.

[252] P. L. NUNEZ, R. SRINIVASAN, ET AL., *Electric fields of the brain: the neurophysics of EEG*, Oxford University Press, USA, 2006.

[253] D. NURSEITOV, A. SEREKOV, A. SHINTEMIROV, AND B. ABIBULLAEV, *Design and evaluation of a p300-erp based bci system for real-time control of a mobile robot*, in 2017 5th International Winter Conference on Brain-Computer Interface (BCI), IEEE, 2017, pp. 115–120.

[254] R. D. O'DONNELL, *Workload assessment methodology*, Cognitive processes and performance, (1986).

[255] D. OETOMO AND M. H. ANG JR, *Singularity robust algorithm in serial manipulators*, Robotics and Computer-Integrated Manufacturing, 25 (2009), pp. 122–134.

[256] D. O'HARE, M. WIGGINS, R. BATT, AND D. MORRISON, *Cognitive failure analysis for aircraft accident investigation*, Ergonomics, 37 (1994), pp. 1855–1869.

[257] D. M. OLVET AND G. HAJCAK, *The stability of error-related brain activity with increasing trials*, Psychophysiology, 46 (2009), pp. 957–961.

[258] R. OOSTENVELD, P. FRIES, E. MARIS, AND J.-M. SCHOFFELEN, *Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data*, Computational intelligence and neuroscience, 2011 (2011).

[259] R. OOSTENVELD AND T. F. OOSTENDORP, *Validating the boundary element method for forward and inverse eeg computations in the presence of a hole in the skull*, Human brain mapping, 17 (2002), pp. 179–192.

[260] T. OTA, R. TOYOSHIMA, AND T. YAMAUCHI, *Measurements by biphasic changes of the alpha band amplitude as indicators of arousal level*, International journal of psychophysiology, 24 (1996), pp. 25–37.

[261] H. PA VAN DONGEN, G. BELENKY, AND J. M KRUEGER, *A local, bottom-up perspective on sleep deprivation and neurobehavioral performance*, Current topics in medicinal chemistry, 11 (2011), pp. 2414–2422.

[262] D. C. PACHECO, A. I. A. MONIZ, AND S. N. CALDEIRA, *Silence in organizations and psychological safety: a literature review*, European Scientific Journal, (2015), pp. 293–308.

[263] R. PARASURAMAN, *Adaptive automation: from theory to practice*, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 45, SAGE Publications Sage CA: Los Angeles, CA, 2001, pp. 291–292.

[264] ——, *Neuroergonomic perspectives on human systems integration: Mental workload, vigilance, adaptive automation, and training.*, (2015).

[265] R. PARASURAMAN, T. BAHRI, J. E. DEATON, J. G. MORRISON, AND M. BARNES, *Theory and design of adaptive automation in aviation systems*, tech. rep., Catholic Univ of America Washington DC cognitive science lab, 1992.

[266] R. PARASURAMAN AND M. RIZZO, *Neuroergonomics: The brain at work*, vol. 3, Oxford University Press, 2006.

[267] L. C. PARRA, C. D. SPENCE, A. D. GERSON, AND P. SAJDA, *Response error correction-a demonstration of improved human-machine performance using real-time eeg monitoring*, IEEE transactions on neural systems and rehabilitation engineering, 11 (2003), pp. 173–177.

[268] P. PARSINEJAD AND R. SIPAHI, *A touchscreen game to induce mental workload on human subjects*, in 2014 40th Annual Northeast Bioengineering Conference (NEBEC), IEEE, 2014, pp. 1–2.

[269] H. PASHLER, *Dual-task interference in simple tasks: data and theory.*, Psychological bulletin, 116 (1994), p. 220.

[270] V. PATWARDHAN, M. A. RIBEIRO, V. PAYINI, K. M. WOOSNAM, J. MALLYA, AND P. GOPALAKRISHNAN, *Visitors‚Äô place attachment and destination loyalty: Examining the roles of emotional solidarity and perceived safety*, Journal of Travel Research, 59 (2020), pp. 3–21.

[271] C. I. PENALOZA, Y. MAE, M. KOJIMA, AND T. ARAI, *Brain signal-based safety measure activation for robotic systems*, Advanced Robotics, 29 (2015), pp. 1234–1242.

[272] B. PENARANDA AND C. L. BALDWIN, *Temporal factors of eeg and artificial neural network classifiers of mental workload*, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, SAGE Publications Sage CA: Los Angeles, CA, 2012, pp. 188–192.

[273] A. PERVEZ AND J. RYU, *Safe physical human robot interaction-past, present and future*, Journal of Mechanical Science and Technology, 22 (2008), pp. 469–483.

[274] G. PFURTSCHELLER, *Eeg event-related desynchronization (erd) and synchronization (ers)*, Electroencephalography and Clinical Neurophysiology, 1 (1997), p. 26.

[275] G. PFURTSCHELLER, B. Z. ALLISON, G. BAUERNFEIND, C. BRUNNER, T. SOLIS ESCALANTE, R. SCHERER, T. O. ZANDER, G. MUELLER-PUTZ, C. NEUPER, AND N. BIRBAUMER, *The hybrid bci*, Frontiers in neuroscience, 4 (2010), p. 3.

[276] L. PION-TONACHINI, K. KREUTZ-DELGADO, AND S. MAKEIG, *Iclabel: An automated electroencephalographic independent component classifier, dataset, and website*, NeuroImage, 198 (2019), pp. 181–197.

[277] D. POPOVIC, M. STIKIC, T. ROSENTHAL, D. KLYDE, AND T. SCHNELL, *Sensitive, diagnostic and multifaceted mental workload classifier (physioprint)*, in International Conference on Augmented Cognition, Springer, 2015, pp. 101–111.

[278] M. I. POSNER AND S. E. PETERSEN, *The attention system of the human brain*, Annual review of neuroscience, 13 (1990), pp. 25–42.

[279] L. J. PRINZEL, F. G. FREEMAN, M. W. SCERBO, P. J. MIKULKA, AND A. T. POPE, *A closed-loop system for examining psychophysiological measures for adaptive task allocation*, The International journal of aviation psychology, 10 (2000), pp. 393–410.

[280] C. M. PRIVITERA, L. W. RENNINGER, T. CARNEY, S. KLEIN, AND M. AGUILAR, *Pupil dilation during visual target detection*, Journal of Vision, 10 (2010), pp. 3–3.

[281] A. PUCE AND M. S. HÄMÄLÄINEN, *A review of issues related to data acquisition and analysis in eeg/meg studies*, Brain sciences, 7 (2017), p. 58.

[282] S. PUMA, N. MATTON, P.-V. PAUBEL, É. RAUFASTE, AND R. EL-YAGOUBI, *Using theta and alpha band power to assess cognitive workload in multitasking environments*, International Journal of Psychophysiology, 123 (2018), pp. 111–120.

[283] E. D. PUTRI, H. NURMAIDA, T. WARSITO, A. SODIKIN, AND S. GULTOM, *The effect fatigue levels of air traffic control (atc) on work effectiveness in soekarno-hatta international airport*, Advances in Transportation and Logistics Research, 2 (2019), pp. 46–50.

[284] Z. W. PYLYSHYN AND E. LEPORE, *What is cognitive science?*, Blackwell, 1999.

[285] Z. W. PYLYSHYN AND R. W. STORM, *Tracking multiple independent targets: Evidence for a parallel tracking mechanism*, Spatial vision, 3 (1988), pp. 179–197.

[286] F. S. RACZ, P. MUKLI, Z. NAGY, AND A. EKE, *Increased prefrontal cortex connectivity during cognitive challenge assessed by fnirs imaging*, Biomedical optics express, 8 (2017), pp. 3842–3855.

[287] T. RADÜNTZ, *Dual frequency head maps: A new method for indexing mental workload continuously during execution of cognitive tasks*, Frontiers in physiology, 8 (2017), p. 1019.

[288] ——, *Signal quality evaluation of emerging eeg devices*, Frontiers in physiology, 9 (2018), p. 98.

[289] T. RADÜNTZ, N. FÜRSTENAU, T. MÜHLHAUSEN, AND B. MEFFERT, *Indexing mental workload during simulated air traffic control tasks by means of dual frequency head maps*, Frontiers in physiology, 11 (2020), p. 300.

[290] T. RADÜNTZ, B. MEFFERT, ET AL., *User experience of 7 mobile electroencephalography devices: comparative study*, JMIR mHealth and uHealth, 7 (2019), p. e14474.

[291] N. RAMNANI AND A. M. OWEN, *Anterior prefrontal cortex: insights into function from anatomy and neuroimaging*, Nature reviews neuroscience, 5 (2004), pp. 184–194.

[292] N. F. RAMSEY, J. M. JANSMA, G. JAGER, T. VAN RAALTEN, AND R. KAHN, *Neurophysiological factors in human information processing capacity*, Brain, 127 (2004), pp. 517–525.

[293] P. RANI, J. SIMS, R. BRACKIN, AND N. SARKAR, *Online stress detection using psychophysiological signals for implicit human-robot cooperation*, Robotica, 20 (2002), pp. 673–685.

[294] J. REASON, *Human error: models and management*, Bmj, 320 (2000), pp. 768–770.

[295] M. Á. RECARTE, E. PÉREZ, Á. CONCHILLO, AND L. M. NUNES, *Mental workload and visual impairment: Differences between pupil, blink, and subjective rating*, The Spanish journal of psychology, 11 (2008), pp. 374–385.

[296] G. B. REID AND T. E. NYGREN, *The subjective workload assessment technique: A scaling procedure for measuring mental workload*, in Advances in psychology, vol. 52, Elsevier, 1988, pp. 185–218.

[297] F. REZAZADEGAN, J. GENG, M. GHIRARDI, G. MENGA, S. MURE, G. CAMUNCOLI, AND M. DEMICHELA, *Risked-based design for the physical human-robot interaction (phri): An overview*, Chemical Engineering Transactions, 43 (2015), pp. 1249–1254.

[298] K. R. RIDDERINKHOF, W. P. VAN DEN WILDENBERG, S. J. SEGALOWITZ, AND C. S. CARTER, *Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning*, Brain and cognition, 56 (2004), pp. 129–140.

[299] A. L. ROACH AND V. G. DUFFY, *Emerging applications of cognitive ergonomics: a bibliometric and content analysis*, in HCI International 2021-Late Breaking Papers: Cognition, Inclusion, Learning, and Culture: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings 23, Springer, 2021, pp. 77–89.

[300] M. D. RODGERS, G. K. DRECHSLER, ET AL., *Conversion of the cta, inc., en route operations concepts database into a formal sentence outline job task taxonomy*, tech. rep., Civil Aerospace Medical Institute, 1993.

[301] S. A. ROSET, K. GANT, A. PRASAD, AND J. C. SANCHEZ, *An adaptive brain actuated system for augmenting rehabilitation*, Frontiers in neuroscience, 8 (2014), p. 415.

[302] M. RUBAGOTTI, I. TUSSEYEVA, S. BALTABAYEVA, D. SUMMERS, AND A. SANDYGULOVA, *Perceived safety in physical human–robot interaction‚Äîa survey*, Robotics and Autonomous Systems, 151 (2022), p. 104047.

[303] M. S. RYOO, *Human activity prediction: Early recognition of ongoing activities from streaming videos*, in 2011 International Conference on Computer Vision, IEEE, 2011, pp. 1036–1043.

[304] K. RYU AND R. MYUNG, *Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic*, International Journal of Industrial Ergonomics, 35 (2005), pp. 991–1009.

[305] K. SAKATA, T. TAKUBO, K. INOUE, S. NONAKA, Y. MAE, AND T. ARAI, *Psychological evaluation on shape and motions of real humanoid robot*, in RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759), IEEE, 2004, pp. 29–34.

[306] A. F. SALAZAR-GOMEZ, J. DELPRETO, S. GIL, F. H. GUENTHER, AND D. RUS, *Correcting robot mistakes in real time using eeg signals*, in 2017 IEEE international conference on robotics and automation (ICRA), IEEE, 2017, pp. 6570–6577.

[307] J. SANDSON AND M. L. ALBERT, *Varieties of perseveration*, Neuropsychologia, 22 (1984), pp. 715–732.

[308] P. SAUSENG, B. GRIESMAYR, R. FREUNBERGER, AND W. KLIMESCH, *Control mechanisms in working memory: a possible function of eeg theta oscillations*, Neuroscience & Biobehavioral Reviews, 34 (2010), pp. 1015–1022.

[309] M. W. SCERBO, *The efficacy of psychophysiological measures for implementing adaptive technology*, (2001).

[310] ——, *Theoretical perspectives on adaptive automation*, in Automation and human performance: Theory and applications, CRC Press, 2018, pp. 37–63.

[311] M. W. SCERBO, F. G. FREEMAN, AND P. J. MIKULKA, *A brain-based system for adaptive automation*, Theoretical Issues in Ergonomics Science, 4 (2003), pp. 200–219.

[312] M. SCHAPSCHRÖER, S. LEMEZ, J. BAKER, AND J. SCHORER, *Physical load affects perceptual-cognitive performance of skilled athletes: A systematic review*, Sports medicine-open, 2 (2016), pp. 1–16.

[313] J. SCHEPERS, A. DE JONG, M. WETZELS, AND K. DE RUYTER, *Psychological safety and social support in groupware adoption: A multi-level assessment in education*, Computers & Education, 51 (2008), pp. 757–775.

[314] R. T. SCHIRRMEISTER, J. T. SPRINGENBERG, L. D. J. FIEDERER, M. GLASSTETTER, K. EGGENSPERGER, M. TANGERMANN, F. HUTTER, W. BURGARD, AND T. BALL, *Deep learning with convolutional neural networks for eeg decoding and visualization*, Human brain mapping, 38 (2017), pp. 5391–5420.

[315] N. M. SCHMIDT, B. BLANKERTZ, AND M. S. TREDER, *Online detection of error-related potentials boosts the performance of mental typewriters*, BMC neuroscience, 13 (2012), pp. 1–13.

[316] D. SCHMORROW, K. M. STANNEY, G. WILSON, AND P. YOUNG, *Augmented cognition in human–system interaction*, Handbook of human factors and ergonomics, (2006), pp. 1364–1383.

[317] M. SCHOPPENHORST, F. BRAUER, G. FREUND, AND S. KUBICKI, *The significance of coherence estimates in determining central alpha and mu activities*, Electroencephalography and Clinical Neurophysiology, 48 (1980), pp. 25–33.

[318] S. SHAPPELL, D. WIEGMANN, J. FRASER, G. GREGORY, P. KINSEY, AND H. SQUIER, *Beyond mishap rates: A human factors analysis of us navy/-marine corps tacair and rotary wing mishaps using hfacs*, Aviation, Space, and Environmental Medicine, 70 (1999), pp. 416–417.

[319] G. SHOU, L. DING, AND D. DASARI, *Probing neural activations from continuous eeg in a real-world task: time-frequency independent component analysis*, Journal of neuroscience methods, 209 (2012), pp. 22–34.

[320] A. K. SINGH, H.-T. CHEN, Y.-F. CHENG, J.-T. KING, L.-W. KO, K. GRAMANN, AND C.-T. LIN, *Visual appearance modulates prediction error in virtual reality*, IEEE Access, 6 (2018), pp. 24617–24624.

[321] S. SLOBOUNOV, K. FUKADA, R. SIMON, M. REARICK, AND W. RAY, *Neurophysiological and behavioral indices of time pressure effects on visuomotor task performance*, Cognitive Brain Research, 9 (2000), pp. 287–298.

[322] J. SMALLWOOD, E. BEACH, J. W. SCHOOLER, AND T. C. HANDY, *Going awol in the brain: Mind wandering reduces cortical analysis of external events*, Journal of cognitive neuroscience, 20 (2008), pp. 458–469.

[323] J. SMALLWOOD AND J. W. SCHOOLER, *The science of mind wandering: empirically navigating the stream of consciousness*, Annual review of psychology, (2015), pp. 487–518.

[324] M. E. SMITH, A. GEVINS, H. BROWN, A. KARNIK, AND R. DU, *Monitoring task loading with multivariate eeg measures during complex forms of human-computer interaction*, Human Factors, 43 (2001), pp. 366–380.

[325] M. SØRENSEN AND M. MOSSLEMI, *Subjective and objective safety*, The effect of road, (2009).

[326] J. SPERANDIO, *Variation of operator's strategies and regulating effects on workload*, Ergonomics, 14 (1971), pp. 571–577.

[327] M. SPÜLER AND C. NIETHAMMER, *Error-related potentials during continuous feedback: using eeg to detect errors of different type and severity*, Frontiers in human neuroscience, 9 (2015), p. 155.

[328] C. STAGG, P. HINDLEY, A. TALES, AND S. BUTLER, *Visual mismatch negativity: the detection of stimulus change*, Neuroreport, 15 (2004), pp. 659–663.

[329] B. STANDARD, *Safety of machinery Ð guards Ð general requirements for the design and construction of fixed and movable guards*, (2009).

[330] M. STERMAN AND C. MANN, *Concepts and applications of eeg analysis in aviation performance evaluation*, Biological psychology, 40 (1995), pp. 115–130.

[331] J. SWELLER, *Cognitive load theory: Recent theoretical advances.*, (2010).

[332] D. SZAFIR AND B. MUTLU, *Pay attention! designing adaptive agents that monitor and improve user engagement*, in Proceedings of the SIGCHI conference on human factors in computing systems, 2012, pp. 11–20.

[333] A. J. SZAMEITAT, T. SCHUBERT, AND H. J. MÜLLER, *How to test for dual-task-specific effects in brain imaging studies‚Äîan evaluation of potential analysis methods*, Neuroimage, 54 (2011), pp. 1765–1773.

[334] A. J. TATTERSALL AND P. S. FOORD, *An experimental evaluation of instantaneous self-assessment as a measure of workload*, Ergonomics, 39 (1996), pp. 740–748.

[335] A. TJOLLENG, K. JUNG, W. HONG, W. LEE, B. LEE, H. YOU, J. SON, AND S. PARK, *Classification of a driver's cognitive workload levels using artificial neural network on ecg signals*, Applied ergonomics, 59 (2017), pp. 326–332.

[336] G. TOBARUELA, W. SCHUSTER, A. MAJUMDAR, W. Y. OCHIENG, L. MARTINEZ, AND P. HENDRICKX, *A method to estimate air traffic controller mental workload based on traffic clearances*, Journal of Air Transport Management, 39 (2014), pp. 59–71.

[337] V. J. TRAVER, A. P. DEL POBIL, AND M. PÉREZ-FRANCISCO, *Making service robots human-safe*, in Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)(Cat. No. 00CH37113), vol. 1, IEEE, 2000, pp. 696–701.

[338] M. TRUSCHZINSKI, A. BETELLA, G. BRUNNETT, AND P. F. VERSCHURE, *Emotional and cognitive influences in air traffic controller tasks: An investigation using a virtual environment?*, Applied ergonomics, 69 (2018), pp. 1–9.

[339] M. ULLSPERGER, A. G. FISCHER, R. NIGBUR, AND T. ENDRASS, *Neural mechanisms and temporal dynamics of performance monitoring*, Trends in cognitive sciences, 18 (2014), pp. 259–267.

[340] D. A. VALENTINO, J. ARRUDA, AND S. GOLD, *Comparison of qeeg and response accuracy in good vs poorer performers during a vigilance task*, International Journal of Psychophysiology, 15 (1993), pp. 123–133.

[341] B. B. VAN ACKER, D. D. PARMENTIER, P. VLERICK, AND J. SALDIEN, *Understanding mental workload: from a clarifying concept analysis toward an implementable framework*, Cognition, technology & work, 20 (2018), pp. 351–365.

[342] J. B. VAN ERP, H. J. VELTMAN, AND M. GROOTJEN, *Brain-based indices for user system symbiosis*, in Brain-Computer Interfaces, Springer, 2010, pp. 201–219.

[343] H. T. VAN SCHIE, R. B. MARS, M. G. COLES, AND H. BEKKERING, *Modulation of activity in medial frontal and motor cortices during error observation*, Nature neuroscience, 7 (2004), pp. 549–554.

[344] J. VELTMAN AND A. GAILLARD, *Physiological indices of workload in a simulated flight task*, Biological psychology, 42 (1996), pp. 323–342.

[345] J. VERA, J. C. PERALES, R. JIMÉNEZ, AND D. CÁRDENAS, *A test-retest assessment of the effects of mental load on ratings of affect, arousal and perceived exertion during submaximal cycling*, Journal of Sports Sciences, 36 (2018), pp. 2521–2530.

[346] V. VILLANI, F. PINI, F. LEALI, AND C. SECCHI, *Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications*, Mechatronics, 55 (2018), pp. 248–266.

[347] V. VILLANI, L. SABATTINI, C. SECCHI, AND C. FANTUZZI, *Natural interaction based on affective robotics for multi-robot systems*, in 2017 International Symposium on Multi-Robot and Multi-Agent Systems (MRS), IEEE, 2017, pp. 56–62.

[348] R. VOCAT, G. POURTOIS, AND P. VUILLEUMIER, *Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task*, Neuropsychologia, 46 (2008), pp. 2545–2555.

[349] Z. WANG, R. M. HOPE, Z. WANG, Q. JI, AND W. D. GRAY, *Cross-subject workload classification with a hierarchical bayes model*, NeuroImage, 59 (2012), pp. 64–69.

[350] X. WANYAN, D. ZHUANG, Y. LIN, X. XIAO, AND J.-W. SONG, *Influence of mental workload on detecting information varieties revealed by mismatch negativity during flight simulation*, International Journal of Industrial Ergonomics, 64 (2018), pp. 1–7.

[351] X. WANYAN, D. ZHUANG, AND H. ZHANG, *Improving pilot mental workload evaluation with combined measures*, Bio-medical materials and engineering, 24 (2014), pp. 2283–2290.

[352] P. WELCH, *The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms*, IEEE Transactions on audio and electroacoustics, 15 (1967), pp. 70–73.

[353] J. R. WESSEL, C. DANIELMEIER, AND M. ULLSPERGER, *Error awareness revisited: accumulation of multimodal evidence from central and autonomic nervous systems*, Journal of cognitive neuroscience, 23 (2011), pp. 3021–3036.

[354] C. WICKENS, *Processing resources and attention, varieties of attention*, R. Parasuraman and D. Davis, Eds. Academic Press, (1984).

[355] C. WICKENS AND P. S. TSANG, *Workload.*, (2015).

[356] C. D. WICKENS, *Multiple resources and performance prediction*, Theoretical issues in ergonomics science, 3 (2002), pp. 159–177.

[357] ——, *Multiple resources and mental workload*, Human factors, 50 (2008), pp. 449–455.

[358] ——, *Processing resources and attention*, in Multiple-task performance, CRC Press, 2020, pp. 3–34.

[359] C. D. WICKENS, W. S. HELTON, J. G. HOLLANDS, AND S. BANBURY, *Engineering psychology and human performance*, Routledge, 2021.

[360] C. D. WICKENS AND Y. LIU, *Codes and modalities in multiple resources: A success and a qualification*, Human factors, 30 (1988), pp. 599–616.

[361] A. WIDYANTI, N. F. SOFIANI, H. R. SOETISNA, AND K. MUSLIM, *Eye blink rate as a measure of mental workload in a driving task: convergent or divergent with other measures?*, Eye, 8 (2017).

[362] W. W. WIERWILLE AND F. T. EGGEMEIER, *Recommendations for mental workload measurement in a test and evaluation environment*, Human factors, 35 (1993), pp. 263–281.

[363] G. F. WILSON, *An analysis of mental workload in pilots during flight using multiple psychophysiological measures*, The International Journal of Aviation Psychology, 12 (2002), pp. 3–18.

[364] G. F. WILSON AND C. A. RUSSELL, *Operator functional state classification using multiple psychophysiological features in an air traffic control task*, Human Factors, 45 (2003), pp. 381–389.

[365] G. F. WILSON AND C. A. RUSSELL, *Real-time assessment of mental workload using psychophysiological measures and artificial neural networks*, Human factors, 45 (2003), pp. 635–644.

[366] J. R. WOLPAW AND E. W. WOLPAW, *Brain-computer interfaces: something new under the sun*, Brain-computer interfaces: principles and practice, 14 (2012).

[367] J. N. WOOD AND J. GRAFMAN, *Human prefrontal cortex: processing and representational perspectives*, Nature reviews neuroscience, 4 (2003), pp. 139–147.

[368] B. XIE AND G. SALVENDY, *Prediction of mental workload in single and multiple tasks environments*, International journal of cognitive ergonomics, 4 (2000), pp. 213–242.

[369] B. YAZMIR AND M. REINER, *Monitoring brain potentials to guide neurorehabilitation of tracking impairments*, in 2017 International Conference on Rehabilitation Robotics (ICORR), IEEE, 2017, pp. 983–988.

[370] Y.-Y. YEH AND C. D. WICKENS, *Dissociation of performance and subjective measures of workload*, Human factors, 30 (1988), pp. 111–120.

[371] R. M. YERKES, J. D. DODSON, ET AL., *The relation of strength of stimulus to rapidity of habit-formation*, (1908).

[372] M. S. YOUNG, K. A. BROOKHUIS, C. D. WICKENS, AND P. A. HANCOCK, *State of science: mental workload in ergonomics*, Ergonomics, 58 (2015), pp. 1–17.

[373] M. S. YOUNG AND N. A. STANTON, *Malleable attentional resources theory: a new explanation for the effects of mental underload on performance*, Human factors, 44 (2002), pp. 365–375.

[374] H. YU, S. HUANG, G. CHEN, Y. PAN, AND Z. GUO, *Human–robot interaction control of rehabilitation robots with series elastic actuators*, IEEE Transactions on Robotics, 31 (2015), pp. 1089–1100.

[375] T. O. ZANDER AND S. JATZEV, *Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment*, Journal of Neural Engineering, 9 (2011), p. 016003.

[376] T. O. ZANDER, C. KOTHE, S. JATZEV, AND M. GAERTNER, *Enhancing human-computer interaction with input from active and passive brain-computer interfaces*, in Brain-computer interfaces, Springer, 2010, pp. 181–199.

[377] T. O. ZANDER, K. SHETTY, R. LORENZ, D. R. LEFF, L. R. KROL, A. W. DARZI, K. GRAMANN, AND G.-Z. YANG, *Automated task load detection with electroencephalography: towards passive brain–computer interfacing in robotic surgery*, Journal of Medical Robotics Research, 2 (2017), p. 1750003.

[378] T. ZEYL, E. YIN, M. KEIGHTLEY, AND T. CHAU, *Adding real-time bayesian ranks to error-related potential scores improves error detection and auto-correction in a p300 speller*, IEEE transactions on neural systems and rehabilitation engineering, 24 (2015), pp. 46–56.

[379] D. ZHANG, B. WEI, AND M. ROSEN, *Overview of an engineering teaching module on robotics safety*, in Mechatronics and Robotics Engineering for Advanced and Intelligent Manufacturing, Springer, 2017, pp. 29–43.

[380] G. ZHAO, Y.-J. LIU, AND Y. SHI, *Real-time assessment of the cross-task mental workload using physiological measures during anomaly detection*, IEEE Transactions on Human-Machine Systems, 48 (2018), pp. 149–160.