Driver Mental States Monitoring Based on Brain Signals

vorgelegt von
Master of Engineering
Shengguang Lei
aus Hunan, China

Von der Fakultät V - Verkehrs- und Maschinensysteme
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften
Dr. -Ing

genehmigte Dissertation

Promotionsausschuss:
Vorsitzender: Prof. Dr. phil. Manfred Thüring
Berichter: Prof. Dr. -Ing. Matthias Rötting
Berichter: Prof. Dr. -Ing. Takashi Toriizuka

Tag der wissenschaftlichen Aussprache: 19.7.2011

Berlin 2011

D 83

Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisor Prof. Dr.-Ing Matthias Rötting for his encouragement, guidance and continuous support of my Ph.D study. His patience, enthusiasm, and immense knowledge helped me in all the time of my research and writing of this dissertation. I would also like to thank Prof. Dr. Takashi Toriizuka and the rest of my thesis committee for their encouragement, insightful comments and questions.

My sincere thanks also go to all of my colleagues in the Chair of Human-Machine-Systems for the stimulating discussions and the enjoyable time in the last four years. In particular, I would like to thank Mario Lasch and Stefan Damke for their untired help and support during the experiments. Also I would like to thank my colleagues Sebastian Welke and Marco Pedrotti for the collaboration in this project, and Micheal Beckman for helping me with the German abstract translation.

Last but not the least, I am thankful to my wife, Mrs. Peng Cheng, supporting me spiritually throughout my life. I would also like to thank my family: my parents Yanghuai Lei and Baimei Liu, for giving birth to me, and my parents-in-law Zhaoyi Cheng and Shufen Xing for their thoughtful care of my life.

Contents

Summary I
Zusammenfassung V
Chapter 1. Introduction 1
Chapter 2. Theoretical Background 9
2.1 Adaptive task allocation 9
2.1.1 The concept of adaptive task allocation 9
2.1.2 Mental workload 12
2.1.3 Task demand, workload, and performance 18
2.1.4 The demand-workload-matched model for adaptive task allocation (DWM-ATA) 23
2.2 The measurement of mental workload 27
2.2.1 Subjective rating 30
2.2.2 Performance measures 32
2.2.3 Physiological measures 35
2.3 Electrocardiogram (ECG) 42
2.3.1 ECG and ECG measures 42
2.3.2 ECG as index of workload 44
2.4 Electroencephalogram (EEG) 46
2.4.1 Mechanism of EEG generation: the brain as a bioelectric generator 47
2.4.2 EEG measurement and parameters 48
2.4.3 EEG as an index of mental workload 55
2.5 Psychophysiology-driven adaptive aiding design 62
2.6 Driving task and driver task load 65
2.6.1 Driving task and driver mental workload 66
2.6.2 Neural correlates of driving 70
2.6.3 State-of-the-art driver workload assessment using psychophysiological signals 75
2.7 Limitations of the current EEG-workload research 78
2.8 Summary of the theoretical background 80
Chapter 3. Representation of driver of workload in EEG: ERP or Band Powers? 83
3.1 Motivation 83
3.2 Introduction of the tasks 84
3.2.1 Lane Change Task 84
3.2.2 Paced Auditory Serial Addition Task (PASAT) 86
3.3 Pre-study: Manipulating workload in Lane Change Task 86

3.4 Assessment of driver's mental workload with EEG 89
3.4.1 Participants 89
3.4.2 Experiment apparatus 89
3.4.3 Experiment procedure 91
3.4.4 Data analysis 91
3.5 Results 96
3.5.1 Task performance 96
3.5.2 ERP in LCT 97
3.5.3 ERP and workload 99
3.5.4 Band Powers and workload 103
3.5.5 Classification accuracy 109
3.6 Discussion 110
3.6.1 What are these components in ERP: A Task Analysis 110
3.6.2 Effect of task load on the amplitude and latency of P300 111
3.6.3 Effect of task load on the EEG spectrum parameters 113
3.6.4 Which is robust for workload representation: ERPs or band powers? 116
3.7 Summary 119
Chapter 4. EEG spectrum modulation with task combination 121
4.1 Motivation 121
4.2 Methods 122
4.2.1 Participants 122
4.2.2 Experiment apparatus 123
4.2.3 Tasks 124
4.2.4 Procedures 125
4.2.5 Data analysis 126
4.3 Results 128
4.3.1 Subjective load (NASA-TLX) 128
4.3.2 Task performance 129
4.3.3 Heart rate and heart rate variability 132
4.3.4 General modulation of the EEG parameters 133
4.3.5 Short-term modulation of the EEG parameters 136
4.3.6 Correlation of EEG parameters to other variables 141
4.4 Discussion 142
4.4.1 Modulation of theta and alpha power with workload 142
4.4.2 Other variables and their correlations to EEG parameters 145
4.5 Summary 146
Chapter 5. A computational model for online workload quantification 148

5.1 Motivation 148
5.2 P-quantile as a generalization method 149
5.3 Logistic regression model for workload quantification 156
5.3.1 The logistic function model (LFM) for workload quantification 157
5.3.2 Regression of the logistic function coefficients with p-quantiles 159
5.3.3 Results with the logistic regression model 161
5.4 Discussion 166
5.4.1 P-quantiles and z-scores 167
5.4.2 Merits and demerits of the logistic function model 168
5.5 Summary 170
Chapter 6. Driver adaptive task allocation in driving simulator 172
6.1 Motivation 172
6.2. Methods 173
6.2.1 Participants 173
6.2.2 Tasks and task load manipulation 174
6.2.3 Experiment setup 174
6.2.4 Quantification of workload with EEG signal 177
6.2.5 Task demand adjustment using DWM-ATA 178
6.2.6 Experiment procedure 180
6.2.7 Data analysis 181
6.3 Results 182
6.3.1 Task demand 182
6.3.2 Subjective load (NASA-TLX) 184
6.3.3 Task Performance 185
6.3.4 HR and HRV 188
6.3.5 EEG Parameters 190
6.3.6 DWM-ATA model and performance 192
6.4 Discussion 194
6.4.1 Psychophysiology based adaptive task allocation 194
6.4.2 The needs of DWM-ATA 196
6.4.3 DWM-ATA model in driving context 198
6.5 Summary 200
Chapter 7. Overall discussion, conclusion, and outlook 202
7.1 Overall discussion and conclusion 202
7.2 Originality, innovations, and new findings 206
7.3 Outlook for future research 209
References 212

List of Abbreviations 246
List of Tables 248
List of Figures 249
Appendix 1. The labview interface developed for experiment 3 253
Appendix 2. The Matlab toolbox for workload detection 254

Summary

Traffic safety has been a serious problem for more than a century. Recently, particular attention has been paid to the human factor issues (e.g. mental workload) associated with in-vehicle high-technologies. In light of *human-centred design*, a newly proposed concept, *adaptive task allocation (ATA)*, is supposed to be an effective solution for these issues. The ATA suggests an adaptive regulation of the task demands upon human operators according to their mental states, which can be assessed using various psychophysiological signals. This dissertation focuses on driver mental workload detection and driver adaptive task allocation, based on the Electroencephalogram (EEG) signals. With such a focus, a serial of studies have been conducted.

In Chapter 1, the research questions and the scopes of this dissertation are introduced. The motivation for this dissertation is the need to address human factor issues associated with in-vehicle high technologies. It has been suggested that, on the one hand, the implementation of the in-vehicle devices (e.g. cellphone, navigator, entertainment systems) forces the driver to engage often in multiple driving-unrelated tasks, which may lead to driver's distraction and mental overload. On the other hand, the highly developed vehicle automation technology (e.g. ACC[1]) seems able to simplify or even monotonize the driving task. Such a case is often thought to be associated with degraded driver mental states (e.g. low vigilance, fatigue, etc). One of the most effective solutions for these issues is the ATA, which *dynamically regulates the information flow to the driver from in-vehicle devices and adaptively modulates the level of automation according to the driver's mental workload*. However, for such a concept, a reliable and accurate

[1] Adaptive cruise control

computational model for the workload assessment using psychophysiological signals as well as the evidential validation of its feasibility in the driving context are definitely essential.

In Chapter 2, the theoretical aspects of the adaptive task allocation, mental workload, and relations of mental workload with the task demand and task performance are introduced. Various context-specific definitions of mental workload as well as theoretical models concerning the relations of mental workload, task demand and task performance are reviewed. An important conclusion from this review is that *the moderate workload level is the optimal stage for the operator's task performance*. Following this, a new model, the *demand-workload-matched adaptive task allocation (DWM-ATA)*, is proposed in this research for addressing the shortcoming of the existing models. Later, an overview of the general characteristics of different workload measures, including subjective reporting, performance measures, and physiological parameters, is given concentrating particularly on the EEG and Electrocardiogram (ECG). Additionally, in this chapter, the driving task analysis, driver behavior modeling, and the neural correlates of driving are provided. Finally, the state-of-the-art studies of psychophysiology-driven workload assessment are systematically reviewed before several limitations of current EEG-workload research are addressed.

In Chapter 3, the details of the first experiment are set forth. This experiment was designed to investigate the changes in EEG parameters including both ERP components and EEG frequency bands (theta, alpha, and beta) with the task load in a simulated driving task, namely, the Lane Change Task (LCT). In the dual task paradigm, another secondary task, the Paced Auditory Serial Addition Task (PASAT) was also used. The comparison between the robustness of ERPs and EEG band powers for driver mental workload assessment was a particular concern

in both single and dual task paradigms. Results indicated that the amplitude of the P300 significantly attenuated with the task load in both single and dual task paradigms, while the theta and alpha power also demonstrate significant changes with the task load. However, the classification of workload using these two groups of parameters (ERPs and EEG band powers) indicated that *the band powers would be more efficient for instantaneous workload detection*, even though both methods have unique merits.

In Chapter 4, the second experiment is presented. In this experiment, the changes of EEG band powers with the task load were continuously investigated in which the driver mental workload was simultaneously attributed to multiple factors. The LCT and another working memory task (n-back task) were used while the task load levels were manipulated in two dimensions (i.e. the driving task load and working memory load), with each containing three task load conditions. Generally, the results consistently indicated in previous studies were reproduced in the driving context: frontal theta activity increased while parietal alpha activity decreased with the task load. However, task-related differences such as driving task load contributed more to the changes in alpha power, whereas the working memory load contributed more to changes in theta power. Additionally, a new finding is also presented which showed that *the variation of the short-term alpha power was decreased with increased task load*.

In Chapter 5, a new computational model for driver mental workload detection is proposed. The analysis of the variation in short-term EEG parameters with the task load suggests that *p-quantiles* could be used to customize the individual differences. A *logistic function model (LFM)* was proposed to quantify the mental workload through a combination of the theta and alpha power into scales that ranged from 0 to 1. The results indicated that the workload scores using the LFM increased with the task load and showed improved correlations with other

workload parameters, more than the single theta or alpha power. The biggest merit of the proposed model is that it can reliably reflect the workload states and enables an easy definition of the workload thresholds, which are also adjustable for various application cases.

In Chapter 6, the third experiment is presented. This experiment took the theoretical DWM-ATA model and the proposed LFM to task by aiming to investigate the feasibility of the EEG-driven adaptive task allocation in a simulated driving environment. Again, the LCT and the n-back task were used in this experiment. The EEG estimated workload states were immediately used to regulate the driving task load induced by the driving speed. The results from this study demonstrated that the LFM can effectively regulate the task load into a moderate level, and that the psychophysiologically driven adaptive task allocation can *improve the operator's performance* and *reduce the operator workload* in a high task load condition. Additionally, *a paradigm integrating the DWM-ATA model into the driver-vehicle-environment loop* is also proposed in the discussion section.

In Chapter 7, the results presented in the previous chapters are discussed in a general manner. From these results, it can be concluded that the *EEG signal provides useful information for inferring driver's workload states*; *the proposed LFM and the theoretical DWM-ATA models are valuable for regulating the task flow to the driver.* Additionally, the originality, innovations, and new findings of this dissertation are addressed in this chapter as well as an outlook into future research on this topic.

Zusammenfassung

Verkehrssicherheit stellt seit dem letzten Jahrhundert ein ernstes Problem dar. In der aktuellen Forschung werden die Themen um die menschlichen Faktoren (z.B. psychische Arbeitsbelastung) im Bereich der fahrzeuginternen Technologien besondere Aufmerksamkeit geschenkt. Adaptive Task Allokation(ATA) ist ein neu Entwickeltes Konzept aus dem Bereich des menschlichen-zentrierten Designs und stellt eine Lösung zur Beachtung von menschlichen Faktoren dar. Die ATA schlägt eine adaptive Regulierung der Aufgabenanforderungen basierend auf dem geistigen Zustand des Menschen vor. Diese können mit verschiedenen psychophysiologischen Signalen bewertet werden. Diese Dissertation konzentriert sich auf die Erkennung von psychischer Arbeitsbelastung und die adaptive Aufteilung von Aufgaben bei Kraftfahrern mit Hilfe eines Elektroenzephalogramms (EEG). Mehrere Studien wurden zur Untersuchung dieser Sachverhalte durchgeführt.

In Kapitel 1 werden die Forschungsfragen eingeführt und die Aufgabenbereiche der Dissertation dargestellt. Die Motivation dieser Dissertation ist die Berücksichtigung der menschlichen Faktoren im Zusammenhang mit fahrzeuginternen Technologien. Die aktuelle Forschung zeigt, dass die Durchführung des in-vehicle Geräte (z. B. Handy, Navigator, Unterhaltung Systeme) einerseits machen die Fahrer oft sich in mehrere driving-unabhängige Aufgaben machen, die zu Fahrer Ablenkung und psychischen Überlastung führen. Zum anderen neigt die zunehmende Automatisierung von Fahrzeugen (z.B. die ACC-Regelung) zu vereinfachten oder sogar monotonen Fahraufgaben. Ein solcher Fall wird oft mit einem degradierten geistigen Zustand des Fahrers in Verbindung gebracht. Beispiele sind niedrige Wachsamkeit und Müdigkeit. ATA

stellt eine effektive Lösung für diese Fragen dar. Diese Methode schlägt eine dynamische Kontrolle des Informationsflusses zwischen dem Kraftfahrer und fahrzeuginternen Geräten vor um adaptiv den Grad der Automatisierung an die mentale Arbeitsbelastung anzupassen. Um solch ein Konzept umsetzen zu können ist ein zuverlässiges und genaues Berechnungsmodell für die Arbeitsbelastung unter Verwendung von psychophysiologischen Signalen, sowie die Validierung der Durchführbarkeit im Fahrzeugkontext essentiell.

In Kapitel 2 werden die theoretischen Aspekte der ATA, psychische Arbeitsbelastung, und die Beziehungen zwischen der geistigen Arbeitsbelastung, der Aufgabenanforderung und der Aufgabeleistung eingeführt. Verschiedene kontext-spezifische Definitionen von psychischer Arbeitsbelastung sowie theoretische Modelle über Beziehungen zwischen psychischer Arbeitsbelastung, Aufgabenanforderungen und Aufgabenleistung werden überprüft. Eine wichtige Schlussfolgerung aus dieser Überprüfung ist, dass eine mäßige Arbeitsbelastung die Aufgabenleistung maximiert. Basierend auf dieser Feststellung wird ein neues Modell, Demand-Workload-Matched Adpative Task Allocation, (DWM-ATA), vorgeschlagen um die Mängel der existierenden Modelle zu bewältigen. Im weiteren Verlauf dieses Kapitels wird eine Übersicht über die allgemeinen Merkmale der verschiedenen Methoden zur Messung der Arbeitsbelastung gegeben. Diese umfassen die subjektive Berichterstattung, Leistungsmessung und physiologische Parameter mit besonderer Konzentration auf das EEG und Elektrokardiogramm (EKG). Zusätzlich werden die Analyse der Fahraufgabe, die Modellierung des Fahrerverhaltens und die neuronalen Zusammenhänge beim Autofahren beschrieben. Schließlich werden aktuelle Studien über psychophysiologische Messungen von Arbeitsbelastung systematisch überprüft und einige Einschränkungen bei der Messung von Arbeitsbelastung mittels EEG betrachtet.

In Kapitel 3 werden die Einzelheiten des ersten Versuchs eingeführt. Dieses Experiment basiert auf einer simulierten Fahraufgabe, dem Lane Change Task (LTC), und dient der Untersuchung der Änderungen der EEG Parameter einschließlich der beiden Event-Related-Potentials (ERP) Komponenten und EEG-Frequenzbänder (Theta, Alpha und Beta) bei sich verändernden Aufgabenanforderungen. Im Zweiaufgabenparadigma wurde die Paced Auditory Serial Addition Task (PASAT) als sekundäre Aufgabe verwendet. Der Vergleich zwischen der Robustheit der ERPs und den EEG Frequenzbänder für die Erkennung der psychischen Arbeitsbelastung des Fahrers wurde beim Einzel- und Zweiaufgabenparadigma betrachtet. Die Ergebnisse zeigen, dass die Amplitude der P300-Komponente erheblich mit steigender Aufgabeanforderung unabhängig von den verwendeten Paradigmen gedämpft wird. Weiterhin gab es auch signifikante Änderungen im Theta- und Alpha-Band bei ändernder Aufgabenanforderung. Die Klassifizierung der Arbeitsbelastung mit diesen zwei Gruppen von Parametern (ERPs und EEG-Frequenzbänder) ergaben, dass die EEG-Frequenzbänder effizienter für die Erkennung der spontanen Arbeitsbelastung sind, obwohl beide Methoden einzigartige Vorteile haben.

In Kapitel 4 wird das zweite Experiment vorgestellt. In diesem Experiment wurden die Änderungen in den EEG-Frequenzbänder mit den Aufgabenanforderungen weiter untersucht bei dem die psychische Arbeitsbelastung des Fahrers gleichzeitig durch mehrere Faktoren erzeugt wurde. Die LCT und eine andere Gedächtnis-Aufgabe, die n-back Aufgabe, wurden verwendet. Das Niveau der Aufgabeanforderung wurde in zwei Dimensionen manipuliert, die Fahranforderungen und Arbeitsgedächtnisauslastung mit jeweils drei Stufen. Im Allgemeinen wurden die Ergebnisse aus früheren Studien auch im Kontext des Autofahrens bestätigt: Erhöhung der frontalen Theta-Aktivität und Verminderung der parietal Alpha-Aktivität bei ansteigender Aufgabenanforderung. Weiterhin wurden noch von der Aufgabe abhängige Unterschiede gefunden.

Fahranforderungen hatten eine größere Auswirkung auf Aktivität im Alpha-Band und die Arbeitsauslastung hat stärker das Theta-Band beeinflusst. Zusätzlich wurde noch festgestellt, dass die Variation der kurzzeitigen Alpha-Aktivität bei verstärkter Aufgabenanforderung abnimmt.

In Kapitel 5 wird ein neues Berechnungsmodell für die Erkennung der psychischen Arbeitsbelastung des Fahrers vorgeschlagen. Die Analyse der Variation der kurzzeitigen EEG-Parameter mit der Aufgabenanforderung suggeriert, dass p-quantiles zur Anpassung der individuellen Unterschiede verwendet werden können. Eine logistische Funktionmodell (LFM) wurde vorgeschlagen um die psychische Arbeitsbelastung durch Kombination der Theta-Aktivität und Alpha-Aktivität in den Bereich von 0 bis 1 zu quantifizieren. Die Ergebnisse zeigen, dass bei steigender Aufgabenanforderung sich auch der Wert der LFM erhöhte. Weiterhin konnte eine erhöhte Korrelationen mit anderen Arbeitsbelastungs-Parametern festgestellt werden als nur mit der Theta-Aktivität oder Alpha-Aktivität. Der größte Vorteil des vorgeschlagenen Modells ist, dass es zuverlässig den Zustand der Arbeitsbelastung widerspiegeln kann und eine einfache Definition von Schwellenwerten der Arbeitsbelastung ermöglicht, die sich an verschiedene Anwendungsfälle anpassen lässt.

In Kapitel 6 wird das dritte Experiment vorgestellt. Dieses Experiment nahm die theoretische DWM-ATA Modell und die vorgeschlagenen LFM in die Praxis zur Untersuchung der Durchführbarkeit der EEG-gestützten ATA in einer simulierten Fahrumwelt. Wieder wurden die LCT und n-back Aufgabe in diesem Experiment verwendet. Die EEG geschätzte Arbeitsbelastung wurde sofort zur Regulierung der durch Fahrgeschwindigkeit induzierten Fahranforderung verwendet. Die Ergebnisse der Studie zeigte, dass die LFM effektiv die Aufgabenanforderung in ein moderates Niveau regeln kann, und somit die psychophysiologie-getriebene ATA bei hohen Arbeitsanforderungen die Leistung des Fahrers verbessern und die

Fahrerarbeitsbelastung reduzieren kann. Zusätzlich wird ein Modell für die Integration der DWM-ATA Modell in die Fahr-Fahrzeug-Umwelt Schleife in dem Diskussionsabschnitt vorgeschlagen.

In Kapitel 7 werden die Ergebnisse in den vorhergehenden Kapiteln in allgemeiner Weise erörtert. Aus diesen Ergebnissen kann der Schluss gezogen werden, dass EEG-Signale nützliche Informationen für die Erkennung der Arbeitsbelastung von Kraftfahrern liefern können. Das vorgeschlagene LFM und das theoretische DWM-ATA Modell sind wertvoll zur Regulierung des Aufgabeabflusses für den Fahrer. Zusätzlich werden in diesem Kapitel die ursprünglichen Modelle, Innovationen und neue Erkenntnisse der Dissertation adressiert und es wird auf Perspektiven der künftigen Forschung zu diesem Thema eingegangen.

Chapter 1. Introduction

Traffic accident has been a serious problem in the modern society. According to Federal Statistical Office of Germany[2], in 2010, there were in total 2,398,414 traffic accidents, in which 3,651 people were dead and 37,114 got injuired. Although these numbers were steadily decreased compared to previous years, traffic accident continues to be one of the leading causes of death and injuries for individuals. Therefore, the research in driving safety still needs to be intensively addressed.

Driving is a common yet complex skill that requires continuous attention and integration of different simultaneous streams of information by a driver (Ho & Spence, 2008). In the past decades, increasing attention has been directed towards human factor issues in the driving task with the aim of improving traffic safety (De Waard, 1996; Baldwin & Coyne, 2005), particularly in situations where driving is increasingly subject to intense change with the continuous advancement in technology (Sayer et al., 2007). The increasing use of new technologies tends to 'drive' the driving task into two paradoxes: complexity and simplicity.

On the one hand, with the increasing traffic density, road complexity and implementation of in-vehicle technologies (IVTs), such as navigation system, cellular phone, in-vehicle entertainment systems, etc., the driving task has become a more complex and interactive activity rather than a simple act of mobility. In such situations, the driver's attention is increasingly diverted from the primary task of driving and hazard detection. Lenneman and Backs (2009) stated that when using the in-vehicle displays and aftermarket devices, the driver often engages in multiple attentional tasks which might not relate to vehicle control and

[2]http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Verkehr/Verkehrsunfaelle/Tabellen/Content75/PolizeilichErfassteUnfaelle,templateId=renderPrint.psml

navigation. Since drivers' functional capacity is limited and can often be exceeded when exposed to dual or multiple tasks, especially in today's complex driving environment (such as increased traffic density or poor weather), the likelihood of driver error may increase (De Waard, 1996). Sayer et al. (2007) reported a field-driving study investigating the driving performance under different secondary tasks including cellphone use. They concluded that all categories of secondary behavior were associated with significantly higher variability in the steering angle. In the last decades, a large number of studies show that the in-vehicle secondary tasks do impair driver's performance (Strayer et al., 2003; Horrey & Wickens, 2004; Caird et al., 2005; Collet et al., 2010).

On the other hand, high technology, which enables cars to be fully automated (e.g. the adaptive cruise control or ACC system enables automatic car driving without driver's input in the highway), can now drive the car to its destination with simplicity and monotony. Indeed, such highly automated system could relieve the human operator from the high demanding task. However, the human factor issues associated with highly automated systems have also been addressed by researchers (Sarter & Woods, 1992; Woods, 1994; Sheridan & Parasuraman, 2006; Lee & Seppelt, 2009; Rauch et al., 2009). In this highly automated system, there are fewer activities for the operator to perform while the role of the operator shifts from an active participant to a passive monitor (Scerbo, 1996). Unfortunately, the new role of passive monitoring is prone to bring the operator into a 'low workload' state, wherein the operator's performance is supposed to decline (Parasuraman et al., 1993, 1994) with operator's fatigue or drowsiness as a possible result (May & Baldwin, 2009). The low workload effect associated with automation is a common concern in aviation psychology research. Recently, it has also addressed the driving context (Brookhuis & De Waard, 1993; De Waard, 1996; Brookhuis et al., 2001; Caleefato et al., 2007). Brookhuis et al. (2001) pointed out that driver alertness in general and paying attention to the driving task

per se is reduced in the case of driving task automation. Therefore, it is of utmost importance to ensure that the driver is in the loop on higher levels of automation when required so that s/he is able to react properly to a potential critical situation (Rauch et al., 2009).

The problems associated with both the complexity and simplicity of the driving task are shown in Figure 1.1. An emerging concept addressing the complexity of driving induced by IVTs is to adaptively modulate the information flow from IVTs (e.g. change the availability of in-vehicle devices) and reallocate the tasks between driver and vehicle according to the driver's functional states (e.g. workload). For instance, the workload manager was proposed to assess the difficulty of driving and regulate the flow of information to drivers that could interfere with driving (Green, 2004). To address the shortcomings of automation, it is also proposed that the level of automation (LOA; Scerbo, 2007) in the vehicle ought to be adjusted according to the driver's function states, what is termed as 'adaptive automation' (Parasuraman et al., 1999). Both concepts of workload manager and adaptive automation actually address only one common human factor issue, the adaptation of task allocation according to operator's functional states (OFS; Wilson & Russell, 1999, 2003a, 2007; Lorenz & Parasuraman, 2003), to wit, adaptive task allocation (ATA; Parasuraman et al., 1999), which has been proposed for the human-centered design for decades.

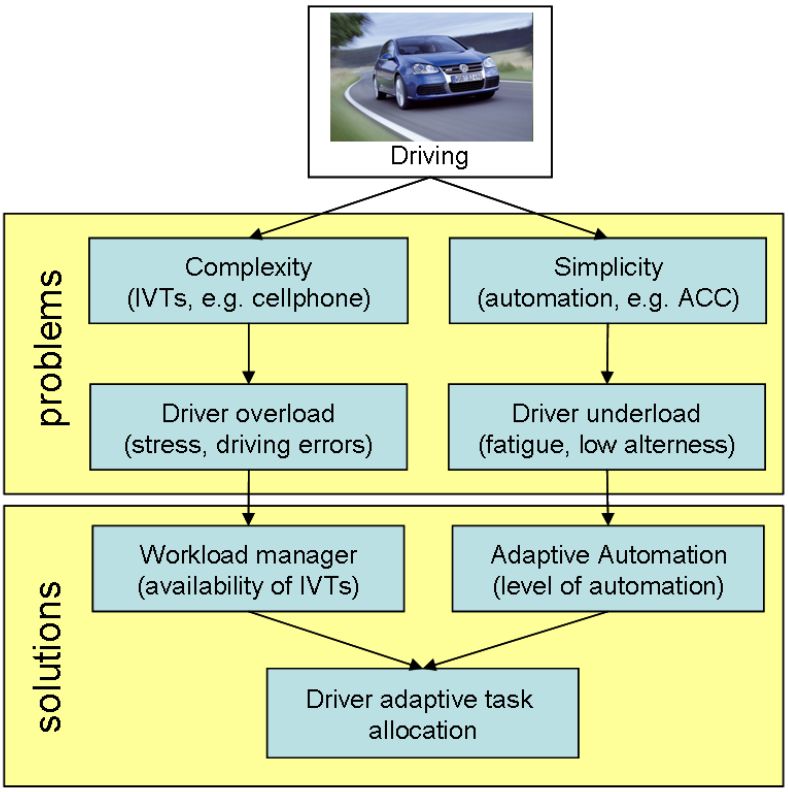


Figure 1. 1 The problems of driving induced by increased complexity and simplicity, for which driver adaptive task allocation is proposed as the solution

Research on ATA focuses primarily on the relationship between human performance and mental workload (Parasuraman et al., 1992, 1999) and addresses the need for a human-centered adaptive system, in which the tasks for the human operator can be dynamically allocated according to the operator's workload states. The needs of ATA have been theoretically and empirically addressed. Parasuraman et al. (1999) proposed an important theoretical model for ATA which interprets the optimal stages for adaptive task allocation between the human and machine according to the operator's mental workload. This model has served as a theoretical framework for later empirical work (e.g. Parasuraman, Consenzo & DeVisser, 2009). Parasuraman and his colleagues (2009) reported two experiments which examined the efficacy of such adaptive automation in a simulated high workload reconnaissance mission using uninhabited air and ground vehicles

(UAVs and UGVs). The results indicated that adaptive automation is beneficial for supporting the human operator tasked with the supervision of multiple uninhabited vehicles under high workload conditions.

Taking such a concept into consideration, especially, to the driving context, two questions rise. Firstly, how can the driver's workload states be instantaneously and reliably assessed? Secondly, can ATA be an effective solution for human factor issues induced by the increasing complexity of in-vehicle technologies as well as the monotony caused by automation? To find answers to these important questions, it is inevitable that a reliable and fast assessment tool should be developed to detect mental workload states. Moreover, it would also be important to demonstrate the empirical evidence validating the feasibility of ATA in the driving context.

In last decades, the mental workload has been examined using various methods including subjective measurement (e.g. NASA Task Load Index, cf. Hart and Staveland, 1988; Subjective Workload Assessment Technique, cf. Reid and Nygren 1988; performance measurement, cf. De Waard, 1996); and psychophysiological signals, cf. Gevins et al., 1997, 1998; Wilson & Russell, 2003a, b; Baldwin et al., 2004; Baldwin & Coyne, 2005; Lei et al., 2009a, b, c; Lin & Cai, 2009; Ting, 2010). However, performance measurement demonstrates occasional insensitivity (Lenneman & Backs, 2009), while subjective evaluation does not measure time-varying qualities and is often influenced by events towards the end of immersion at the time of questionnaire administration (Insko, 2003).

Psychophysiological measures, e.g. electroencephalogram (EEG), electrocardiogram (ECG), pupil dilation, do address shortcomings in performance and subjective measurements and are found to be robust candidates for operators' workload evaluation (Kramer, 1991; Gevins et al., 1997, 1998; Wilson & Russell, 2003a;

Brookhuis & De Waard, 2010). Unlike performance and subjective measurements, psychophysiological measures offer continuous observation in high time resolution (e.g. in milliseconds) and could be collected without intruding into the operator's task (Kramer, 1991; Wilson & Russell, 2007; Ting et al., 2010).

In this thesis, we concentrate in particular on using EEG technology for workload assessment. The EEG signal is a representation of the brain's electrical activity recorded from electrodes placed on the scalp. The advantages which EEG possesses include: higher temporal resolution (e.g. on the oder of milliseconds), intrusiveness and a wide application environment[3]. It has been used to assess the operator's workload for many years in both laboratory (Gundel and Wilson, 1992; Berka et al., 2007; Lei et al., 2009 a, b, c; Fu et al., 2010 a, b) and applied settings (Wilson, 2001, 2002; Kohlmorgen et al., 2007). Both time-locked event-related potential components, e.g. P300 (a positive deflection occurs around 300 ms after stimulus onset) and EEG spectral components, e.g. theta (4-8 Hz) and alpha (8-12 Hz) have been revealed to be sensitive to changes in the task load. For instance, the amplitude of P300 was reported to attenuate with the task load (Wickens et al., 1977; Schultheis and Jameson, 2004; Brookhuis & De Waard, 2010); the increased task load may lead to increased frontal theta activity and decreased parietal alpha activity (Yamamoto & Matsuoka, 1990; Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Smith et al., 2001). However, such findings are largely task- or individual- dependent. So far, there is still a remarkable lack of well-accepted methods for workload representation that could be customized for both tasks and individuals. In particular, the driving task is actually a combination of various cognitive processes and driver's mental workload which may be simultaneously attributed to various factors. In this perspective, establishing a general model for driver mental workload assessment

[3] Comparing with other brain imaging technologies, such as PET and fMRI, EEG is less demanding on the experiment settings and can be used not only in laboratory environment but also applied settings.

would be particularly challenging albeit helpful.

The psychophysiology-driven adaptive system has also been investigated for decades. Pope et al. (1995) were among the first to report the usefulness of psychophysiological indexes for the task allocation with feedback controls. They found that three indices - beta/alpha, beta/(alpha plus theta), and theta/alpha - were able to distinguish the feedback conditions, but the best discriminator was the index beta/(alpha plus theta). The results from their study supported the notion that the psychophysiologically driven closed-loop system provided a method for regulating operator attention, arousal, and workload. Similar conclusions can be found in other studies (Prinzel et al., 2000; Wilson & Russell, 2007; Parasuraman et al., 2009; Ting et al., 2010). In terms of the driving context, only few studies have reported that such system is useful for improving the driver-vehicle interaction. Kohlmorgen and colleagues (2007) reported an EEG-based system for detecting driver mental workload under real traffic conditions. They showed that the EEG driven adaptive aiding is beneficial for improving drivers' overall task performance. However, for investigating the usefulness and feasibility of ATA in the driving context, a single study is insufficient. Both theoretical and empirical evidence are needed.

In this dissertation, we concentrate more intensively on these two points: a robust workload quantification model using EEG signals, and the validation of the feasibility of ATA in the driving context. We begin with the theoretical background of the ATA, the workload and its assessment, the EEG technology, the driving task and driver workload, etc. Particularly, a new demand-workload-matched model for ATA (DWM-ATA) is proposed as an extension of the Parasuraman's model (1999). Following this, a study aiming to find the optimal variables in the EEG signals for the workload representation was conducted and the approval for using EEG spectrum components (theta and alpha) to index the workload was concluded.

Later, a study to further investigate the variations of EEG spectrum parameters with the task load was conducted. This study focuses for the most part on theta and alpha power changes induced by task combinations, in which the driver's workload is simultaneously attributed to different factors. Furthermore, a computational model, viz. logistic function model (LFM), for online driver mental workload was proposed. This model projects the EEG parameters into unified scales for workload representation with the logistic function. Specifically, the p-quantiles were used to customize the individual variations. Such an approach provides a generalized method for workload quantification even across tasks. Finally, a study investigating the EEG-driven closed-loop for driver adaptive task allocation in a simulated driving environment was also conducted. This study applied the new theoretical DWM-ATA model and LFM into practice in order to demonstrate the efficacy of adaptive task allocation in the driving context.

Chapter 2. Theoretical Background

2.1 Adaptive task allocation

2.1.1 The concept of adaptive task allocation

The concept of adaptive task allocation focuses primarily on human factor issues associated with a highly automated system. The automation technology is fast-developing ever since its inception. Such technology attempts to allocate the performance of activities to a machine or system rather than to the human operator (Parsons, 1985). Numerous merits associated with automation have been addressed. Automation could increase the flexibility of operations or permit to control more complex systems (Woods, 1994), since automation enables machines to take over more and more the performing of activities in situations where the human operator is overloaded with activities. Furthermore, the automation can attenuate the variability associated with human performance and therefore significantly reduce errors, making the system more efficient (Scerbo, 1996). For instance, Wiener (1988) demonstrated that in the aviation industry, automation helps to reduce flight times, increase fuel efficiency, and improve the pilot's perceptual and cognitive capabilities. Further advantages attributed to automation include increased productivity, better product quality, more efficient use of materials, shorter workweeks for labour, and improved safety.

However, human factor issues of highly automated systems have also been addressed by researchers (Sarter & Woods, 1992; Woods, 1994; Sheridan & Parasuraman, 2006; Lee & Seppelt, 2009). Lee and Seppelt (2009) pointed out that over-trusting in automation as a way to increase system efficiency and safety often leads to disappointment since the role of people becomes more, not less,

important as automation becomes more powerful and prevalent. Such a technology-centred approach, instead of a human-centred design, might lead to new and more catastrophic failures. Funk et al. (1999) systematically investigated human factor issues associated with flight deck automation through a survey of a wider variety of sources (e.g. accident reports, incidents reports, pilots, aviation experts, etc). They concluded that three issues, namely, misunderstanding of the automation, non-apparentness of automation behaviour, and overconfidence in automation, are ranked at the top of human factor issues.

Another concern is that automation results in a shifting of the human operator's role in the human-machine systems. In a highly automated system, there are only few activities for the operator to perform and the role of the operator shifts from an active participant to a passive monitor (Scerbo, 1996). Consequently, such a role-shifting and the operator's overreliance on automation might excessively lower the operator's workload (i.e. mental underload) which can cause a decrement in vigilance, sustained alertness or situation awareness, and can thus lead to boredom (Manly et al., 1999). Scerbo (1996) stated that such an overreliance on automation may make the operator less aware of what the system is doing at any given moment. It is also believed that humans are not well suited for the monitoring task in extended periods of time (Parasuraman, 1986; Scerbo, 1996). Parasuraman and colleagues revealed that the ability to monitor automation failures is reduced under automatic conditions compared with manual operating conditions (Parasuraman et al., 1993, 1994).

Therefore, the automation involves both benefits and costs. In order to avoid the costs of automation, humans would need to change their view of automation. Sheridan & Parasuraman (2006) pointed out that automation does not mean that humans are replaced. Oppositely, humans are asked to interact with automation in a so-called human-automation interaction (Parasuraman et al., 2000; Sheridan &

Parasuraman, 2006). The concept of human-centred automation has been proposed to address the shortcomings of automation (Sheridan, 1995). One of the basic solutions for such a concept is adaptive automation (AA; Parasuraman, 1987; Scerbo, 1996; Kaber & Riley, 1999). Adaptive automation is an alternative approach to automation integration compared to static automation (Hancock et al., 1985; Parasuraman et al., 1999). The basic principle of adaptive automation is to dynamically change the level of automation based on the operator's functional states (Kaber et al., 2005). A higher-level concept behind adaptive automation is that the human-machine systems dynamically change the task assignments, not only the level of automation, to the human according to operator's functional states, which is defined as adaptive task allocation (ATA). Such a concept suggests building a closed-loop in the human-machine systems by introducing the human's workload state as a feedback aiming to optimize the human's performance (Figure 2.1).

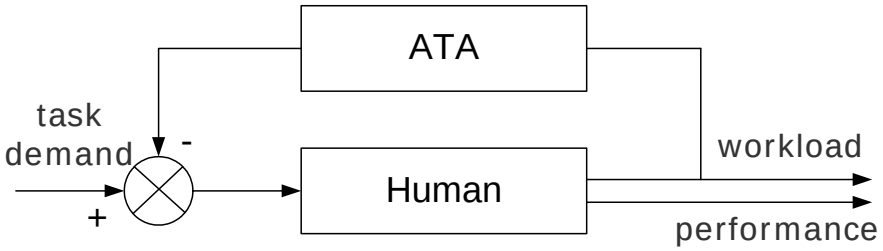


Figure 2. 1 A closed-loop for adaptive task allocation

In this disseration, the author argues that ATA can be applied not only in the automation context, but also in contexts where the complex systems are semi-automated or less developed. Such systems may impose severe requirements on the human operator's mental or information-processing capabilities. For instance, the increasing implementation of IVTs (e.g. navigation system and cellular phone) poses a new challenge to the driver on how to cope with multiple attention-demanding tasks which might not relate to the primary driving task (Lenneman & Backs, 2009). Since such devices are somehow helpful for daily

driving, it is often hoped that they be made available in the vehicle. But solutions are also expected to prevent the negative impact of these devices on the primary driving task. ATA is among these solutions and in this context it serves as a regulator of the information flow to the operator from various subsystems according to the operator's mental states.

The ATA research focuses primarily on the relationship between human performance and mental workload (Parasuraman et al., 1999). These studies attempt to determine the optimal task demand to the human operator or seek optimal stages for switching of tasks between human and machine. To understand this concept, the theory about workload, human-machine interaction and its relation to task demand, performance, is needed.

2.1.2 Mental workload

Definition of mental workload

Mental workload is a miscellaneous concept. So far, there is neither a common understanding of mental workload nor a shared methodology for measuring it. Historically, the concept of mental workload was first introduced in the 1940s within the context of optimizing human-machine systems (Bornemann, 1942). Since then, a large body of literature has appeared with various definitions of (mental) workload. A simplistic definition is that it is a demand placed upon humans (De Waard, 1996), which attributes workload exclusively to an external source. However, mental workload can be better defined as the combined effect of external and internal sources on humans, reflecting the interaction of mental demands imposed on operators by tasks they attend to. It is well-accepted that the mental workload is not only affected by the task demand from outside, but also by person-specific ones (Rouse et al., 1993), such as individual capabilities,

motivation for task performance, strategies adoption, as well as the mood and operator's functional states (De Waard, 1996, p. 15).

Various definitions referring mental workload to a relation between the task demand and personal capacity have been formulated. It is suggested that mental workload describes the "*relation between the (quantitative) demand for resources imposed by a task and the ability to supply those resources by the operator*" (Wickens, 2002). Gopher and Donchin (1986, p. 41) stated that "*mental workload may be viewed as differences between the capacities of the information processing system that are required for task performance to satisfy performance expectations and the capacity available at any given time*." Eggemeier et al. (1991, p.207) defined it this way: "*Mental workload refers to the portion of operator information processing capacity or resources that is actually required to meet system demands*" (for a similar definition, see O'Donnell & Eggemeier, 1986). Such definitions emphasize that resources for human cognition and information processing are limited (Kahneman, 1973) and that mental workload is equivalent to the amount of "*the demand imposed by tasks on the human's limited mental resources…*" (Wickens, 2008). In a different view, the mental workload is defined as "*the operator's evaluation of the attentional load margin (between their motivated capacity and the current task demands) while achieving adequate task performance in a mission-relevant context* " (Jex, 1988, p. 11).

Another group of definitions for mental workload focus on the amount of information that can be accomplished within a definite period of time, namely, the information processing rate. Colle and Reid (1999) stated that "*mental workload is considered to be the average rate of mental work…*"; Xie & Salvendy (2000 b) suggested that "*mental workload is the amount of mental work or effort necessary for a person or group to complete a task over a given period of time*". Additionally, in neuroscience, the mental workload is often referred to as the degree to which

neural resources are activated by an effectual task performance (Smith et al. 2001). Therefore, workload definitions are extremely varied within the scientific literature depending on the research interest.

To establish a standard terminology for mental workload, the international norm EN ISO 10075[4] was created. This standard was developed to provide a theoretical fundament and establish principles and requirements for mental workload measurement. According to this standard, the mental workload has been defined as the combination of the *mental stress* and *mental strain*, which were first distinguished by Rohmert (1984). The authors of ISO 10075-1 defined the term "mental" as referring to cognitive, informational, and emotional processes in the human being. Mental stress is a neutral term which has been defined by ISO 10075-1 as: "*The total of all assessable influences impinging upon a human being from external sources and affecting it mentally*". Therefore, mental stress describes the entire set of all external influences on the human being. Such influence can be attributed to any factor including aspects of task demands, work equipment, physical work environment, and social work environment (Hagemann, 2008). Mental strain is defined as: "*The immediate effect of mental stress within the individual (not the long-term effect) depending on his/her individual habitual and actual preconditions, including individual coping styles*". Such a definition depicts mental strain as the short-term effect of mental stress on the individual attributed to a variety of operator characteristics. The effects of mental strain can be both facilitating and impairing. The former includes warming-up effect and activation, whereas the later could be mental fatigue or fatigue-liked states including monotony, low vigilance, mental satiation, etc.[5]

[4]European standard EN ISO-10075 – Ergonomic principles related to mental workload – General terms and definitions. The original version is ISO 10075-1 published in 1991. The later versions ISO 10075-2 and ISO 10075-3 were published in 1996, 2004 respectively.
[5] ISO 10075

Mental workload in this dissertation is defined in line with the definition of ISO 10075 which includes both *mental stress* and *mental strain*. It reflects the cognitive, informational, and emotional processes in the human brain. In neuroscience, the mental workload can also be referred to the degree of utility in the neural (or mental) resources for task performing (Smith et al., 2001).

Following the multidimensional nature of workload, Xie & Salvendy (2000a) proposed a general theoretical framework aimed at developing context-specific[6] workload assessment methods. They divided the workload concept into five dimensions: instantaneous workload, peak workload, accumulated workload, average workload, and overall workload. The instantaneous workload measures dynamic changes in the workload values during task performance. The typical examples for such measures are the physiological markers. The peak workload is referred to as the maximal value of instantaneous workload. Accumulated workload is the total amount of instantaneous workload. The average workload is defined as the average of the instantaneous workload. Finally, the overall workload is the individual's experienced mental workload which maps instantaneous workload (or accumulated and averaged workload) in the operator's brain (Xie & Salvendy, 2000a). The technique of subjective workload belongs to this category. The authors argue that before measuring the mental workload, one should first select the relevant workload dimension(s). Xie & Salvendy's conceptual framework is particularly useful for the present dissertation, as it introduces the important concept of instantaneous workload, which can be distinguished from the other dimensions. This framework serves as the basis for the short-term workload measurement using the physiological parameters.

[6] In the above paragraphs, we introduced various mental workload definitions according to the different application contexts.

Some theoretical models related to mental workload

Human Information Processing (HIP) models: Modelling human information processing is the basic approach of human-machine interaction research to explore human cognition. Generally, these models depict the human brain as an analogy of the electronic computer and divide the human information-processing into three stages: sensory (or perception) of environmental information, central processing of perceived information, and responding to that information (Figure 2.2). However, Wickens et al. (2004) added more sophisticated elements into his model, which has been well-accepted in the human factors domain. In Wickens' model, the *feedback* was introduced to show that a human being has the ability to monitor the consequences of his or her action. Central to the model for this dissertation is the concept of attentional resources, which has been defined as "*a sort of pool of attention or 'mental effort' that is of limited availability and can be allocated to processes as required*" (Wickens et al., 2004). As shown in Figure 2.2, each of these structures relied on the attentional resources. It is suggested that multi-task performance is allowed in case the pool of attentional resources is not depleted and different tasks are not competing for the same resources. Otherwise, mental overload may occur. Another advantage of this model is integration of the *working memory* concept, which refers to the temporary storage of information linking with various cognitive tasks such as reading, problem-solving, or learning (Baddeley, 1983). Working memory in Wickens' model is distinct from the long-memory.

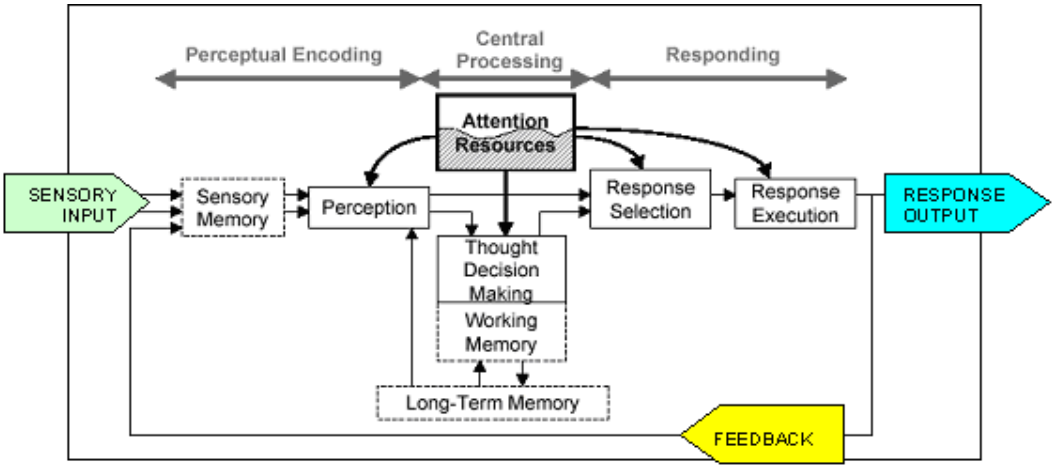


Figure 2. 2 Human Information processing model (Wickens et al., 2004, p. 122)

Multiple resources theory: Evidence from the experiment settings as well as from real life experience often shows that a successful performance of two concurrent tasks can be achieved when one task is auditory and the other visual, more than in cases where both are either auditory or visual. For addressing the underlying reasons for this, the multiple resources theory (Figure 2.3) proposed by Wickens (2002) is adopted. This model can be viewed as an extension of the concept of attentional resources in Wicken's HIP model. The multiple resource theory is linked to the resources allocation in the multiple task performance and has both practical and theoretical implications (Wickens, 2002). Wickens suggested that the resource structure can be described using four important categorical and dichotomous dimensions that account for time-sharing performance: two states of processing (*perceptual-central* and *response*), two modalities of perception (*auditory* and *visual*), two codes of processing (*visuospatial* and *verbal/phonological*), and two aspects of visual processing channels (*focal* and *ambient*). According to Wickens (2002), tasks involving shared resources might lead to a decline in task performance when resources are not sufficient to meet task demands. The multiple resources theory provides useful implications for the level of performance in cases with two or more time-shared tasks to be coped with at the same time. For instance, it can effectively interpret why the performance of

two-visual (or auditory) tasks is less successful than the case of one visual and one auditory task. The two visual processing channels theory is interpretable in that humans can successfully walk down while reading (Wickens, 2002).

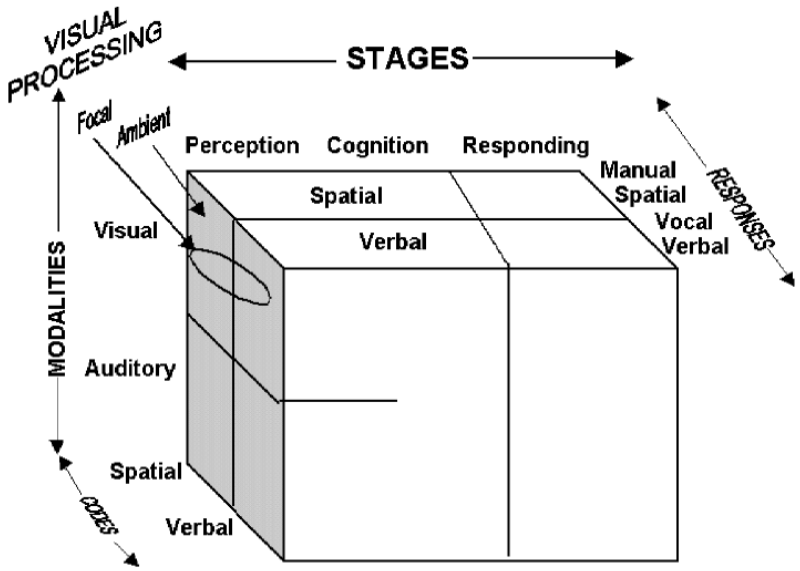


Figure 2. 3 Wickens' Multiple Resources Theory (adapted from Wickens, 2002)

2.1.3 Task demand, workload, and performance

Task demand or load is imposed by the external task and is generally distinct from the concept of workload[7], which is considered as the individual experience of the task demand. Generally, the operator's workload is expected to be linearly associated to task load, i.e. the increased task load leads to increased workload. However, as introduced in section 2.1.1, workload is a multifaceted concept and influenced not only by the task demand placed on the operator, but also by the operator's characteristics (De Waard, 1996). Parasuraman and Hancock (2001) pointed out that the workload may be driven by the task load from external environmental sources, but also mediated by the individual human response (or internal factors) to the task load. Hilburn and Jorna (2001) stated that such internal

[7] In the following text, workload refers to the mental workload.

factors include skill, training, experience, fatigue, time pressure, and motivation. Particularly, the adopted strategy plays an especially important role for the contribution of the operator's workload (Hilburn and Jorna, 2001; Parasuraman and Hancock, 2001). In other words, there may be cases where the high task load does not necessarily lead to a high level of workload, and vice versa. Such cases were termed as the 'dissociation' of the task demand and workload (Parasuraman & Hancock, 2001).

A model concerning the inter-relationships of the task load, mental workload and performance was proposed by R. Parasuraman[8] (Figure 2.4). In this model, the task load, together with the strategies, affects the operator's workload, while the workload and strategies both contribute to the operator's performance. That is, factors from the task load, performance and adaptive strategies directly or indirectly affect the operator's or system performance. Interestingly, in this model, the system performance could be given as feedback to regulate the task load exposed to the operator.

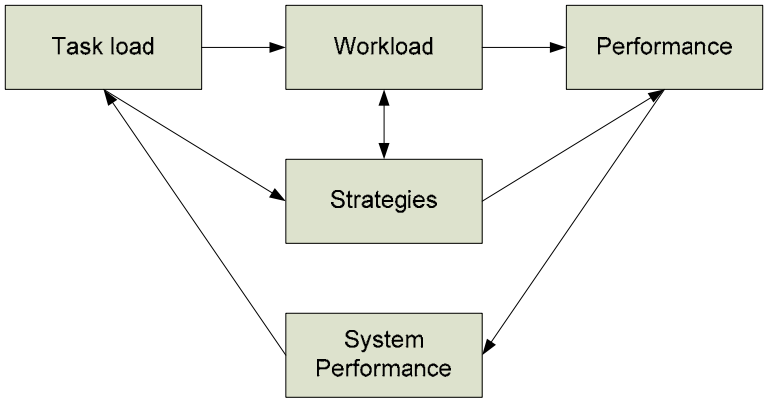


Figure 2. 4 Inter-relationships of task load, workload, and performance (modified from Parasuraman & Hancock, 2001)

More quantitatively, the relation between the workload and performance was first

[8] This model appeared originally in 'Workload and Vigilance' (p. 116) in: Flight to future human factors in air traffic control (1997), Washington, DC: National Academy Press.

described by Yerkes and Dodson (1908), called the Yerkes-Dodson law (Figure 2.5), which was originally related to arousal into performance. The law has been extended to incorporate workload as a replacement for arousal (Hancock & Warm, 1989). A relationship similar to the Yerkes-Dodson curve suggests a drop in the operator's performance during low workload, while a drop in performance during high arousal is a result of overload. This model suggests that the moderate workload level is the optimal region for operator performance.

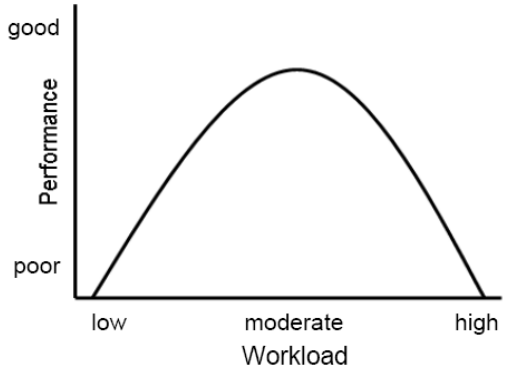


Figure 2. 5 Yerkes-Dodson law about the relation between workload and performance

The Yerkes-Dodson curve only depicts a two-dimensional relationship. A more complex model providing quantitative relationships of task demand, workload, and performance was proposed by De Waard (1996). This model suggests dividing the whole task demand terrain into 6 regions[9] (Figure 2.6). In region D (D for deactivation), the operator's state is affected by monotonous tasks. In this area, De Waard (1996) suggested that there may be a state of deactivation in this region and this could lead to an increase in workload, since the low demand tasks can result in a reduction in capacity. The performance of this region is also supposed to decrease, which is consistent with the Yerkes-Dodson curve. The region A (including A1, A2, and A3) is equivalent to the middle range of the Yerkes-Dodson curve and is suggested as the optimal region for performance. In region A2, the operator can easily cope with the task requirements and reach an

[9] These 6 regions include A1, A2, A3, B, C, D, as shown in Figure 2.6.

adequate level of performance, with the mental workload reaching the lowest. In regions A1 and A3, performance remains unaffected but the operator compensates in terms of more efforts so as to maintain an undisturbed performance level. In region B, the task demands exceed the operator's capacity and the workload becomes higher. It is no longer possible to maintain performance, even though the operator devotes more effort to the task. In region C, performance is at a minimum level and the operator is definitely overloaded.

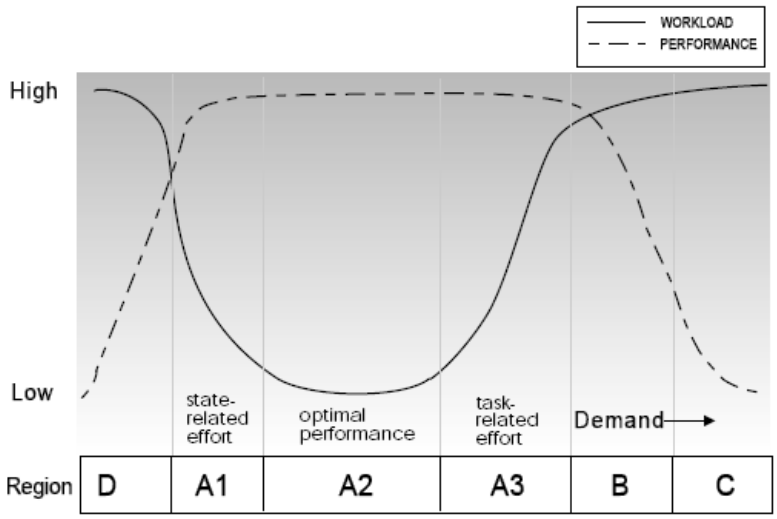


Figure 2. 6 De Waard's model concerning relations between task demand, workload and performance (adapted from De Waard, 1996)

However, the interpretation of the D region in De Waard's model seems controversial. Wilson & Schlegel (2004) presented a model (Figure 2.7, see also Veltman and Jansen, 2006; Greef et al., 2007) which is slightly different from De Waard's. In a low workload level, the performance is supposed to decline (or at least not increase) and the moderate workload level is thought to be the optimal stage for performance. This is consistent with both the Yerkes-Dodson curve as well as De Waard's model. The difference between Wilson & Schlegel's model and De Waard's lies in the relation between low task demand and workload. In Wilson & Schlegel's model, the low task demand is supposed to elicit low workload, which contradicts the D region in De Waard's model. Additionally, it

suggests that as the task demand continuously increases, the performance can be maintained by compensating the amount of effort until the task demand reaches the limit of the operator's capacity. This is in line with the A3 region in De Waard's model. Wilson & Schlegel's model portrays a generally linear association between the workload and task loads. This is acceptable when ignoring the occasional dissociation between workload and task loads.

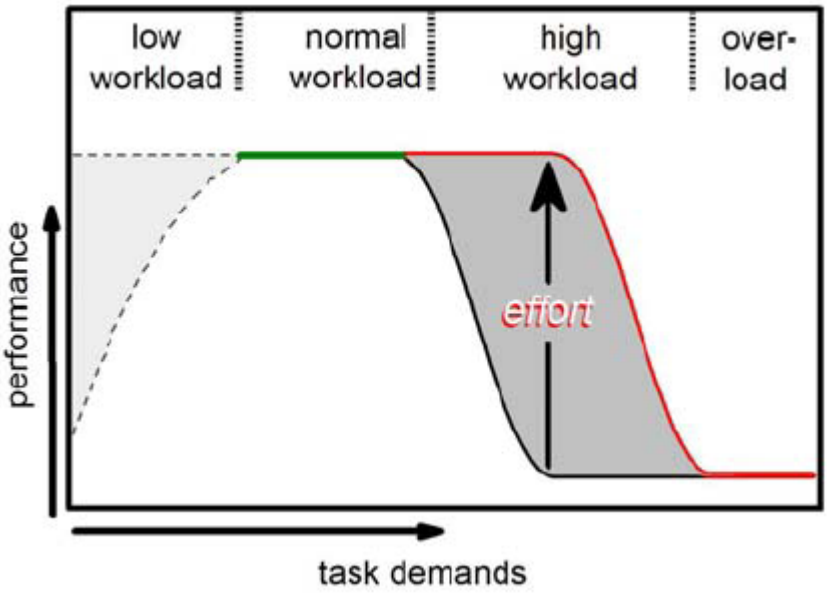


Figure 2. 7 Wilson & Schlegel's model for task demands, workload, and performance (adapted from Greef et al., 2007)

To sum up, the Yerkes-Dodson curve, De Waard's model, and Wilson & Schlegel's model consistently demonstrate that the optimal region for task performance lies in the moderate (or normal) workload level. Both high end and low end of the workload are not appropriate for the task performance. Additionally, the workload is thought to be linearly associated with the task load in general. But there are also occasional cases wherein workload is dissociated with the task demand[10]. The theoretical models introduced in this section serve as a fundamental basis for the adaptive task allocation concept, which attempts to find a solution for questions such as how the task demand can be allocated to the operator, how the operator's

[10] High (low) task demand does not necessarily induce high (low) level of workload

workload can be the maintained in a moderate level for an unattenuated performance, etc.

2.1.4 The demand-workload-matched model for adaptive task allocation (DWM-ATA)

As described in the last section, to optimize the operator's performance, the workload should be regulated in a moderate level. In this perspective, Parasuraman et al. (1999) proposed a model of workload-matched adaptive task allocation or adaptive automation (Figure 2.8). This model plots the operator workload (containing three levels: low, medium, and high) against the workload time-period. In the high workload periods, it is recommended that the task load assigned to the human operator should be reduced (e.g. increasing the level of automation), while the task load for the human operator should be increased during low workload periods (e.g. decreasing the level of automation) in order to avoid the impairing effects[11] associated with a low level of mental strain. The goal of this model is to regulate the workload into the optimal level for human performance (the medium level) by dynamically adjusting the task assignment between human and machine.

[11] Such effects often involve mental fatigue or fatigue-liked states including monotony, low vigilance, mental satiation and so forth, see Section 2.1.1.

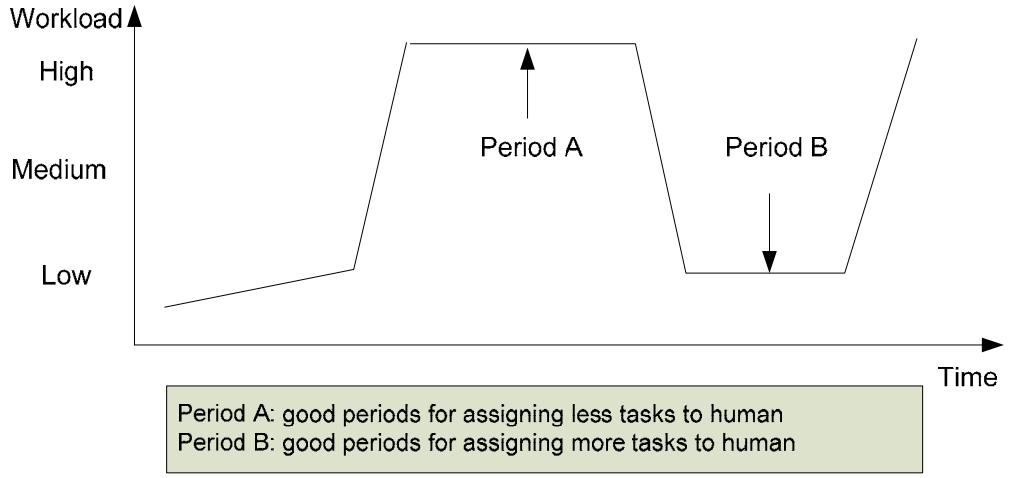


Figure 2. 8 Workload-matched adaptive task allocation (modified from Parasuraman et al., 1999)

Parasuraman's model provides a good solution for adaptive task allocation. However, taking Parasuraman's model into practice, especially in case of psychophysiology-based adaptive task allocation, some problems have to be addressed.

Lei, Zhang, and Roetting (2011) demonstrated some empirical evidence that there is occasional occurance of dissociation between the workload and task demand when the workload is evaluated by the EEG signals. For instance, it was indicated that there were some periods of high workload in low task demand condition and also some periods of low workload in high task demand condition. This finding is consistant with previous theory about task demand and workload. As discussed in section 2.1.2, task demand is not the only factor contributing to the workload. Factors such as strategy, motivation, emotional states, etc., also significantly affect the workload. Consequently, the dissociation between the task demand and workload may occur when such factors simultaneously contribute to the operator's workload. Parasuraman & Hancock (2001) pointed out that there are unusual circumstances where the external task demand maybe increasing while experienced workload is decreasing, and vice versa. These cases were also

addressed by Yeh and Wickens (1988) and Hancock (1996). Particularly, when workload is evaluated in terms of instantaneous workload (defined by Xie & Salvendy, 2000a), the dissociation cases may become more frequent since the instantaneous workload varies dynamically rather than maintains a relative constant value.

In this perspective, it is reasonable to divide the demand-workload plane into 9 areas (Figure 2.9). Region 1, 5, and 9 reflect a general linearly associated relation between task demand and workload, while the rest of the regions (2, 3, 4, 6, 7 8) reflect the dissociated demand-workload relation. In the regions 4, 7, 8, the participant is '*overpaced*', since the workload is higher than is required by the task. These situations could be the result of the operator's internal factors, e.g. poor task strategy selection, inappropriate emotional states (anxiety), etc. By contrast, in regions 2, 3, and 6, the participant exerts less effort than is required by the task. We term such regions as an '*underpaced*' state. Especially in region 3, the participant may completely give up trying to perform since the task seems too difficult. For instance, in our second experiment, the participant frequently gave up the secondary working memory task when s/he was required to drive at high speed. A break-down in the performance in this region is verified in Chapter 6 (section 6.3.6).

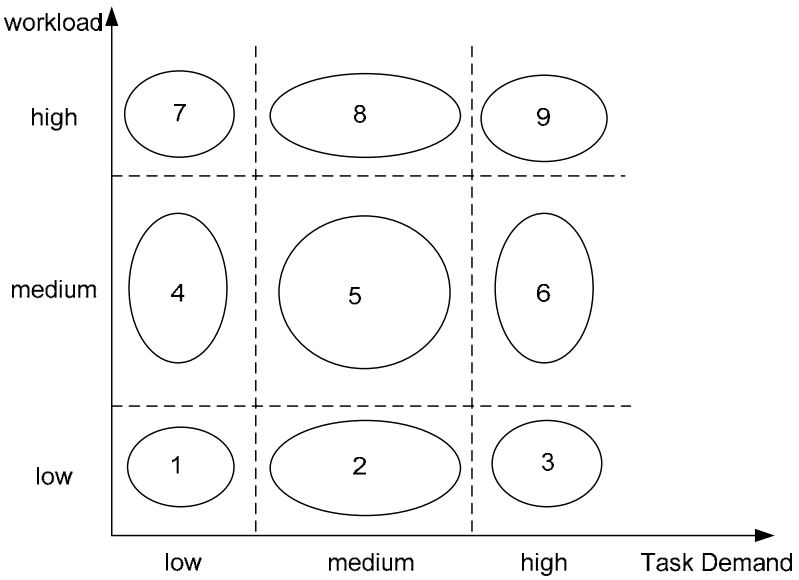


Figure 2. 9 A 3*3 metrix for the task demand and workload

In Parasuraman's model, the operator's workload is thought to be the only factor for triggering the allocation of task demand. This is principally based on the postulation that workload is linearly associated with the task demand. In the dissociation cases, the workload-only-matched adaptive task allocation might be inappropriate. For instance, according to Parasuraman's model, the task demand placed on the human operator will be increased in cases where the workload is low even if the task demand is already very high. In such situations, a continuously increased task demand may potentially lead to a failure in task performance. In region 7, it is also suggested that the task demand should be reduced according to Parasuraman's model. However, in this region the high workload may not be a result of task demand but other factors and therefore, reducing the task load to the operator may not be helpful for improving operator's workload states. Some methods which may help to reduce the participant's emotional burden or a better task performing strategy would be more beneficial in this area rather than decreasing the task demand.

Therefore, taking both association and dissociation between the demand and

workload, we proposed a new strategy (see, Figure 2.10) for adaptive task allocation as an extention of Parasuraman's model. The task load assignment in most of the regions (1, 2, 4, 5, 6, 8, 9) is consistent with Parasuraman's model (1999) for adaptive task load, increasing task load when the workload of the participant is low, and decreasing task load when the task load is high. However, in region 7, the task load level is maintained, since the high workload is unrelated to the task load in this area. We suggest that some other solutions, e.g. light music, may help relieve the operator's mental strain. In region 3, it is also necessary to lower the task load level to avoid 'giving up' due to high task demand.

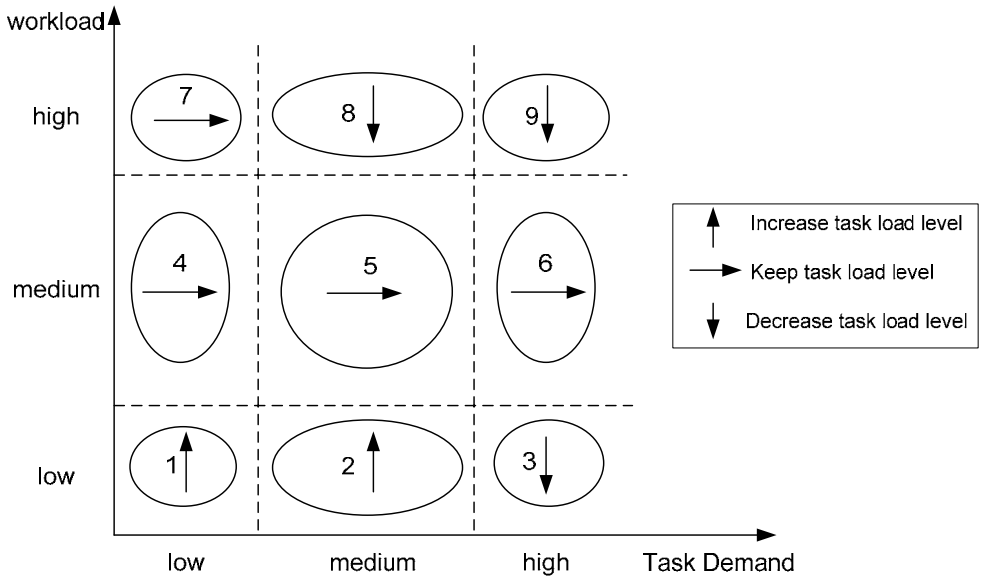


Figure 2. 10 Demand-workload-matched adaptive task allocation (DWM-ATA) for human operator

2.2 The measurement of mental workload

Since the inception of workload research, a large number of workload assessment techniques have appeared. Generally, workload measurement techniques could be divided into three categories: self-reporting or subjective ratings, performance measurements (including primary or secondary task measurements), and

psychophysiological measures. It has already been noted that different measures are sensitive to different aspects of the workload. These differences might be partially explained by the lack of any accepted definition of workload. However, generally speaking, there are some fundamental principles and criteria for the evaluation of workload. O'Donnell and Eggemeier (1986) defined a set of five criteria for the evaluation of workload assessment indices which has been frequently cited throughout the literature. A few years later, this set was expanded by Wickens (1992) and other researchers to a set of seven criteria which need careful consideration in any serious investigation within this field of research. The five main criteria are listed below:

- *Sensitivity:* Sensitivity is usually investigated by manipulating the task demands and task complexity at multiple levels. This criterion requires that the method is reliably sensitive to reflect the changes in workload at different levels of difficulty or complexity during performance of the task.

- *Diagnosticity:* Diagnosticity depicts whether an index responds to a specific source of workload and allows conclusions to be drawn about the cause(s) of a certain type or resource of mental workload.

- *Nonintrusiveness:* The method should not be intrusive and disruptive, nor should it contaminate or interfere with the operator's ongoing performance of tasks. The principle is especially important in a real world environment involving critical safety issues. The physiological and subjective ratings show the lowest level of intrusion, while secondary tasks seem likely to be the most prone to interference.

- *Acceptance:* The accuracy of a measurement may be affected by the operator's acceptance or approval of the technique. Low acceptance may lead

to poor accuracy and low face validity. Therefore, the method should be acceptable to the subjects.

- *Practical requirements:* The method should be practical and require minimal equipment for the assessment. Additionally, the constraint on the amount of practice that an operator needs in order to perform a task to reach a reasonable and stable performance is also included in this criterion.

Generally, sensitivity and diagnosticity are considered to be the most important of these criteria, while the others may be considered additional selection criteria (De Waard, 1996). In the next part of this chapter, the different workload assessment techniques will be reviewed. The summarised methods for workload assessment are shown in Figure 2.11 and the following section introduces these methods in details.

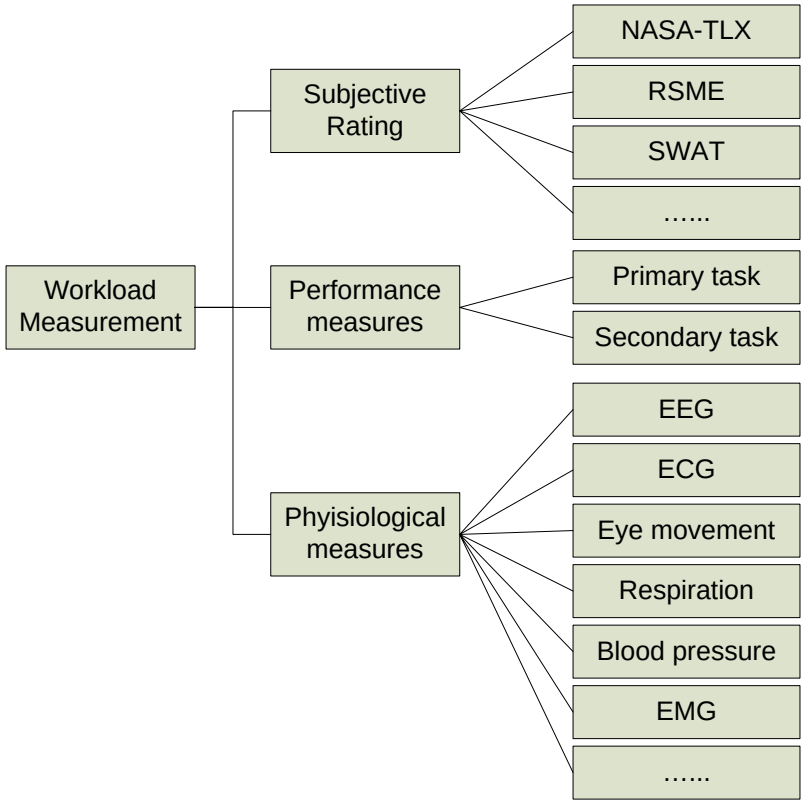


Figure 2. 11 Workload assessment methods

2.2.1 Subjective rating

Subjective rating is a psychological construct which attempts to quantify the personal interpretations and judgements (Cain, 2007). Subjective rating has been implemented in many research topics such as assessment of workload, effort, mood, fatigue, etc. It is considered to be the best measure since it comes nearest to tapping the essence of mental workload (Wickens, 1984; Hart & Staveland, 1988). In the absence of any single objective measure, the fundamental measure is the individual's subjective evaluation (Jex, 1988). Muckler and Seven (1992) stated that "*the operator's awareness of increasing effort being used, even before any performance degradation occurs, should give subjective measures a special role to play*", and then suggest that the biggest advantage of subjective measures is their subjectivity. Since its inception, a number of subjective rating techniques have been developed and these generally could be subdivided into two categories: unidimensional ratings and multidimensional ratings.

A popular one-dimensional questionnaire is the Rating Scale Mental Effort (RSME; Zijlstra and van Doorn, 1985). It consists of a vertical line from 0 mm (representing almost no effort) to 150 mm (representing the highest effort) in steps of 10 mm. Statements related to invested efforts are given at several anchor points along the line. The distance from the origin to the mark that subjects draw represents the amount of invested effort. Another unidimensional example is the Modified Cooper-Harper scale (MCH, Wierwille & Casali, 1983). It is a set of criteria originally used for the evaluation of the handling qualities of aircraft for flight tests. It consists of 10 scales, with 1 indicating the best handling characteristics and 10 being the worst.

The most commonly used multidimensional ratings are the NASA task load index (NASA-TLX; Hart & Staveland, 1988) and the subjective workload assessment

technique (SWAT; Reid & Nygren, 1988). NASA-TLX divides the workload into six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration level, with each ranging from 0 to 100 scales. In order to obtain a general workload score, the NASA-TLX makes the first aggregation with a paired-comparison weighting scheme. In contrast, SWAT consists of only three categories: time load, mental effort, and physiological stress. Unlike NASA-TLX, the performance dimension is not included in SWAT, since Reid and Nygren (1988) reckoned that performance measures cannot of themselves describe workload, because operators may compensate some effort to maintain a constant performance level when the task demand increases. Similar to NASA-TLX, the SWAT also proposes an aggregation procedure to produce an overall workload rating.

More specifically, there are also subjective techniques specialized in driver workload assessment. Schumacher (cited by Hagemann, 2008, p. 61) used a video rating to assess the level of mental workload in certain driving situations. The driving situation is recorded by video and displayed to the participant after driving. The participant uses a sliding lever to continuously rate the situations. Pauzie (2008) proposed the driver activity load index (DALI) as a subjective method to evaluate driver workload. DALI is a revised version of NASA-TLX. It divides the driver workload into six categories: effort of attention, visual demand, auditory demand, temporal demand, interference, and situational stress - followed by a weighting procedure in order to combine the six individual scales into a global workload score.

In general, subjective measurement has several advantages. It is easy to apply especially when using unidimensional scales. It is highly immediate and can be provided even before any performance degradation. However, this immediacy might be obtrusive in some continuously performing task such as driving, in

which promoting the workload assessment might disturb the driving task itself (Hagemann, 2008). In this situation, subjective evaluation does not measure the time-varying qualities of presence and is easily influenced by events towards the end of immersion around the time of questionnaire administration (Insko, 2003). Hagemann (2008) pointed out that biases in measurement might occur due to personal memory recall when the subject is asked after task completion. Additionally, Yeh and Wickens (1988) found that a number of the tasks influence the subject's judgement in such a way that the workload in dual task conditions is always reported to be higher than that in the single task even if the two tasks occupy separate resources and are not difficult.

2.2.2 Performance measures

Measuring the changes on performance such as performing accuracy, reaction time (RT), etc., could provide an index of workload. With an increasing level of workload, it is assumed that this will affect the quality of operator performance (usually degradation). Cain (2007) stated that when demands exceed the operator's capacity, the performance degrades from baseline or ideal levels. Therefore, performance measures are often used as important indicators of workload. The performance measures of workload can be generally divided into two categories: primary task measures and secondary task measures.

Primary task measures

Primary task measures attempt to assess directly the operator's performance on the task at hand. It has been suggested that an increasing level of task difficulty or complexity requires additional processing resources and may affect the quality of the operator's performance. The relationships between task demand, workload and performance have been discussed in section 2.1.3. Accuracy, speed, reaction times

and error rates are often used to assess primary task performance. Primary task measures are assumed to be globally-sensitive and provide an index of load variation across a variety of operator information processing functions (Eggemeier et al., 1991).

However, some drawbacks of the primary task measures have also been addressed. Firstly, as described in De Waard's demand-workload-performance model (De Waard, 1996), the performance is not sensitive to the change of workload in region A. In this region, the performance can be maintained in cases where the increasing task demand is still 'at the limit of one's capacity', since increased task demand can easily be compensated by effort investment (increasing of workload). Even in cases of overload condition, no additional changes might be observed when performance reaches a certain limit. Secondly, other factors such as performance strategy, the energetic states of the operator (rather than the workload) have their influence on the operator's performance. Last but not least, the performance measures are task specific. Primary task measures are unique to their task and a comparison between two different primary tasks seems impossible.

Therefore, to use primary task measures as the only resource for workload assessment is not proper. It is necessary to combine performance measures with other workload measures in order to draw a valid conclusion.

Secondary task measures

Secondary task measures provide an index of the operator's spare capacity when performing primary tasks. It is assumed that a totally undifferentiated capacity for individuals is available for all tasks. Spare capacity is defined as the unused capacity available for the secondary task in the case of unaffected primary task performance (De Waard, 1996). The secondary task measures are often used to

infer the interaction between the primary task and the secondary task. Generally, it is recommended that the secondary task be divided into two different paradigms: 'subsidiary task paradigm' and 'loading task paradigm' (see O'Donnell & Eggemeier, 1986). Within the 'subsidiary task paradigm', the performance of the primary task is maintained and consequently, the performance of the secondary task varies with the task load. Therefore, performance of the secondary task could be the index of workload. In contrast, with the 'loading task paradigm', the subject is instructed to maintain the performance of the secondary task even if performance of the primary task might decline when the task load increases, so that primary task performance measures can be used as indicators of workload.

The secondary task measures are often used for driver workload detection, especially for the evaluation of workload or inference of the in-vehicle secondary task to the primary driving task (Van Winsum et al., 1999; Mattes, 2003). Van Winsum et al. (1999) presented a peripheral detection task (PDT) for workload detection. The standard task requires simple manual responses to stimuli presented at certain places to the driver as secondary task when driving. The response times and hit rates could be used as indicators of the workload. Martens and van Winsum (2000) used the PDT in a simulator study and demonstrated that response times increased and hit rates decreased when task demands are increased. Mattes (2003) initiated a standardized Lane Change Task (LCT) for the evaluation of the workload or distraction of various secondary tasks. This LCT is widely used for the driver workload detection, combined with certain forms of the secondary task (Hagemann, 2008; Lei et al., 2009 b; Lenneman & Backs, 2009).

However, problems for secondary task measures are also often addressed. Firstly, the performance measurement demonstrated occasional insensitivity (Lenneman & Backs, 2009). This could be explained by Wickens's multiple resource theory. According to the multiple-resource theory (Wickens, 2002), the most sensitive

situation in secondary task measures is that there is an overlap in resources used. In other words, if the time-shared resources occupied by the two tasks are different, the performance may fail to indicate the workload. Secondly, the intrusion of the secondary task into the primary task limits the utility of this methodology. A high level of intrusion in the primary task might occur if the secondary task shares the same resources with the primary task. Tsang & Vidulich (2006) pointed out that the introduction of secondary task not only adds additional workload to the operator, but also changes the processing of the primary task. The operators may follow a different resource policy even though they are instructed to prioritize the primary task. Eggemeier et al. (1991) stated that the secondary task always imposes non-specific intrusion (e.g. peripheral interference). Eggemeier et al. (1991) also mentioned that secondary task measures require additional instrumentation and exhibits a lack of operator acceptance. For such reasons, the use of secondary task measures in an applied environment (e.g. car driving) requires additional cautions.

2.2.3 Physiological measures

Another group of workload measures are those derived from the operator's physiology. Various physiological parameters are reported to be differently sensitive to changes in the task load (Kramer, 1991; De Waard, 1996). In general, physiological indicators could be divided into two categories which are associated with two separate nerve structures: central nervous system (CNS) and peripheral nervous system (PNS). CNS consists of the brain and spinal cord, while PNS can be further divided into the somatic nervous system and autonomic nervous system (ANS). The somatic nervous system regulates activities that are under conscious control, while ANS acts as a control system for internal organs without consciousness. ANS is typically divided into two subsystems: the parasympathetic nervous system (PNS) and sympathetic nervous system (SNS). The SNS responds

to emergency reactions, and is responsible for the increase of heartbeat and blood pressure, whereas PNS is needed for relaxation regulation and is responsible for the constriction of the pupil, decrease in heart rate, dilation of blood vessels, etc. (see Kramer, 1991; De Waard, 1996). Physiological signals such as pupil diameter, heart rate and respiratory, electrodermal activity (EDA), and hormone level measures, belong to the ANS-measures. Electroencephalogram (EEG), electrooculargram (EOG), magnetic and metabolic activity measures of the brain all belong to CNS.

Unlike performance measures and perceptual ratings, physiological measures attempt to interpret the psychological processes through their effect on the body states (Cain, 2007). The biggest attraction of the physiological measures is that they could continuously record the operator's response in high temporal resolution without intruding into the operator's task (Kramer, 1991; Wilson & Russell, 2007). Hagemann (2008) pointed out that this approach provides maximal objectivity since the subject can hardly control the outcome of the recordings. Cain (2007) also stated that this kind of approach may be "*particularly useful when subjective methods or performance measures become insensitive due to covert changes in operator strategies, or the applied levels of effort lead to an apparent dissociation among subjective and performance measures*". Additionally, physiological signals could be recorded in a variety of operational environments such as car driving (Brookhuis & De Waard, 1993) or flight piloting (Wilson, 2001, 2002b), etc.

On the other hand, the disadvantages of physiological indicators are often addressed by researchers. Kramer (1991) mentioned that the greatest weakness of this approach at least from the analyst's perspective is a lack of strong conceptual link with the physiological measures of performance. This problem is reiterated by Eggemeier et al. (1991) in the need for a better understanding of the links between the various physiological measures and mental workload. From the perspective of

users, most negative issues of the physiological measures are associated with the technological parts. Physiological measures often entailed cumbersome, invasive equipment, unsuitable for applied settings (Cain, 2007). Moreover, some techniques (e.g. EEG) require considerable time for setting or calibration, and therefore exhibit low acceptance by the user. Nevertheless, this problem might be susceptible since advances in technology have made the equipment far more portable and capable.

The most frequently used physiological measures include: EEG, ECG, EDA, pupillatory response, blink duration, blood pressure, respiration rate, electromyogram (EMG), etc. Reviews and discussions of these techniques could be referred to Krammer (1991), De Waard (1996), and Manzey (1998). The ECG and EEG technologies were used in our experiments to index workload. The detailed introductions to these two techniques are given in sections 2.3 and 2.4. Here, we briefly introduce some other modalities including the eye movement, EDA, blood pressure, respiration, and EMG.

Eye movement

Measurements of eye activity can be unobtrusive using various equipments (e.g. eye-tracker, head mounted display, etc.) that are already in place to support these measurements. There are a number of eye activities, such as: horizontal and vertical eye movement (extent and speed), blink activity (duration, latency and frequency), gaze fixation duration, and pupil diameter. These parameters are often reported to be sensitive to human cognitive activities. Cain (2007) pointed out that ocular measures are sensitive to mental demands. However, they are also sensitive to other factors, particularly fatigue.

The pupil dilameter has for decades been researched to investigate the cognitive

activity (Beatty, 1982; Kramer, 1991; Marshall, 2000, 2002). Beatty (1982) suggested that task-evoked pupillary responses could be a consistent index of the cognitive load within tasks, across tasks, and possibly even across individuals. Kramer (1991) pointed out that pupillary responses could provide a more reliable measure for processing demand in general, compared to other measures such as ERPs[12]. Wilson (2004) noted that pupil diameter generally increases with higher cognitive processing levels and is sensitive to rapid changes in the workload. However, when overload occurs, pupil diameter can become unresponsive to changes or even reverse its response. More recently, Bailey and Iqbal (2008) used an average percentage change in pupil size (APCPS) to index workload within various tasks including route-planning, document editing, and email classification. Methodologically, a patented method (Marshall, 2000) has been introduced, the Index of Cognitive Activity, which allows precisely the locating of cognitive load in the time dimension (Marshall, 2002).

Except for the pupil diameter, the eye blink activities, in terms of blink rate, blink duration, and blink latency, contain also information about the cognitive load. Cain (2007) pointed out that blink measures can be context-dependent. A decrease in blink rate has been expected when there is an increased workload with visual stimuli processing. However, Wilson (2004) showed an increase in the blink rate when the task load increased with memory tasks. Furthermore, the connection between blink rate and workload is thought to be tenuous (Castor et al., 2003). Blink closure duration appears to decrease with increased workload, resulting from visual stimuli or gathering of data from a wide field of view, while blink latency increases with memory and response demands (Castor et al., 2003). Recently, Pedrotti et al. (2011) proposed a data-driven algorithm and identification-artifact correction (I-AC) for eyeblink detection. Such an approach allows for the performing of a clean signal analysis and provides a basis for the

[12] The introduction to ERPs is given later in section 2.4.

ongoing online eye blink detection.

Electrodermal activity (EDA)

EDA reflects electric potential and resistance of the skin as a result of ANS activity. It can be expressed in terms of Skin Resistance Response (SRR) or Skin Conduction Response (SCR), which are inversely related. EDA can be distinguished in tonic and phasic activity (Heino et al., 1996). Tonic EDA is the baseline activity, while phasic EDA is the electrodermal response (EDR) induced by the external stimulus. Additionally, Kramer (1991) adds a third category, viz. spontaneous or non-specific EDA, which is thought to be generally sensitive to certain levels of arousal.

Evidence from previous studies indicates that EDA serves as a reliable indicator of mental workload in various contexts (e.g. Wilson and O'Donnel, 1988; Wilson, 2001) including driving (Baldauf et al., 2009). Baldauf et al. (2009) reported that tonic EDA (baseline activity) varied significantly with changing driving demands when participants drove on a round course of a driving simulator consisting of three different environments with different levels of task demands. However, De Waard (1996) pointed out that spontaneous EDA, relative to unknown stimuli, has predominantly been used as an indicator of arousal or emotion, but not as a measure of workload, since it exhibits a global sensitivity. Heino et al. (1990) also stated that all behaviour, e.g. emotional, physical, etc., that affects the sympathetic nervous system can lead to a change in EDA. Furthermore, EDA is thought to be fairly slow, approximately 1.3 to 2.5 s to the occurrence of stimulation (Kramer, 1991; De Waard, 1996).

Respiration

Respiration supplies oxygen to the blood and expels carbon dioxide. De Waard (1996) stated that measures of respiration could provide an index of energy expenditure, which may be a result of increased cognitive effort (Backs & Seljos, 1994). The well-studied parameter associated with workload assessment is the respiration rate (Wilson, 1992; Wilson & Russell, 2003a). Respiration rate is consistently found to increase as the task demand increases (Backs & Seljos, 1994; Partin et al., 2006; Mehler et al., 2009; Lehrer et al., 2009; Karavidas et al., 2010). Mehler et al. (2009) reported a study investigating the impact of secondary task load on driver's physiology and an increase in respiration rate was found as well when the secondary task demands increased. More recently, Karavidas et al. (2010) revealed that the respiration rate was significantly associated with the in-flight task workload.

However, De Waard (1996) pointed out that the main problem with respiration is the measurement technique itself. The flow meters, which can provide an accurate analysis of expired gasses, somehow add dead space and resistance, and are very intrusive. Other methods such as strain gauges, impedance pneumography and air flow temperature measurement are less intrusive but are also less accurate (Porges & Byrne, 1992). Additionally, respiration measurements demonstrate global sensitivity and may be affected by not only workload, but also other factors such as speech and physical effort, emotions, personality characteristics, etc. (De Waard, 1996).

Blood pressure

Blood pressure is exerted by circulating blood upon the walls of blood vessels. It is closely related to the heart beat. An often used blood pressure parameter is the

blood pressure variability (BPV), which is thought to be linearly associated to the heart rate variability[13] (HRV; De Waard, 1996; Mulder, 1988). The blood pressure is analyzed also as an indicator of workload within various studies (Wierwille, 1979; Veltman & Gaillard, 1996; Lehrer et al., 2008). Veltman and Gaillard (1996) reported that blood pressure was affected by the different levels of task difficulty in a simulated flight task. More specifically, Hjortskov et al. (2004) speculated that heart rate-derived variables reflect a central pathway in cardiovascular control mechanisms, while the blood pressure response is more influenced by local conditions in the working muscles that partially mask the effect of changes in mental workloads. However, in several studies, the use of blood pressure to index workload was reported to be disappointing (Corwin et al., 1989).

Electromyogram (EMG)

EMG is a technique for evaluating and recording the electrical activity produced by muscle cells. Research related to mental workload and EMG focuses on two types of measurements: the limb-muscle activity and facial muscle activity with more concentration on the latter. Facial muscles are supposed to be strongly related to expressive behaviour and also function in the regulation of cerebral blood flow and temperature (De Waard, 1996). It is also suggested that the task-irrelevant or spontaneous electrical activity of muscles could be a reliable and promising index for workload (Van Boxtel & Jessurun, 1993; De Waard, 1996). Detailed description of a generation of spontaneous EMG is given by De Waard (1996). Van Boxtel & Jessurun (1993) and Waterink & Van Boxtel (1994) reported that the lateral frontalis and corrugator supercilii, as well as orbicularis inferior facial muscles all reflect mental effort. EMG was reported to be sensitive to workload changes in various later studies (Yang et al., 1997; Doyle et al., 2009). However, the EMG technique requires careful calibration and the setup is fairly

[13] Introduced in section 2.3.

time consuming. Additionally, the EMG technique requires also careful instrumentation, data manipulation, and interpretation. These may limit the use of such method during application.

2.3 Electrocardiogram (ECG)

2.3.1 ECG and ECG measures

The ECG interprets the electrical activity caused by depolarization and polarization of the heart muscle. It reflects the electrical impulses produced by heart contraction, which is innervated by both PNS and SNS. It is captured and recorded by skin electrodes. The ECG can be expressed in three methods: (a) time domain measures; (b) amplitude measures; and (c) frequency domain measures.

A typical time-domain ECG tracing of the cardiac cycle (heartbeat) consists of a P wave, a QRS complex, a T-wave, and a U-wave (Figure 2.12). The QRS complex is often used to detect peaks while the time between peaks i.e. namely, Inter-Beat-Interval (IBI), can be extracted. Typically, heart rate (HR) and heart rate variability (HRV) are widely used for the representation of the mental workload. HR is determined by the number of heart beats within a fixed period of time (usually per minute) and is non-linearly related to IBI. Compared with IBI, HR is less normally distributed in samples (Jennings et al., 1974; De Waard, 1996). Additionally, the amplitude of T-wave (TWA) is another variable in the ECG signal reflecting SNS activity (Furedy, 1987). Müller et al. (1992) reported that the amplitude of TWA decreased with increases in SNS activity.

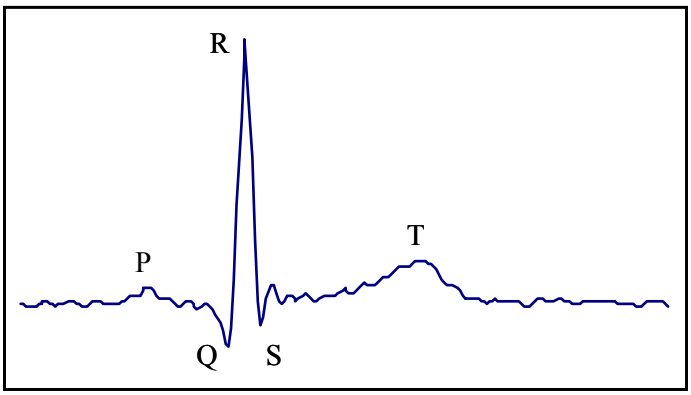


Figure 2. 12 The typical time-domain ECG tracing of the cardiac cycle[14]

Compared with HR and TWA, the analysis of HRV is more complex. HRV is usually defined as the changes in the interval between heart beats in either time or frequency domain. It reflects the irregularities in heart rate caused by a continuous feedback between the CNS and peripheral autonomic receptors (De Waard, 1996). In the time domain, HRV parameters, which are often used in clinical research, include: SDNN (the standard deviation of NN interval[15]), SDANN (standard deviation of the average NN intervals calculated over short periods), RMSSD (the square root of the mean squared differences of successive NN intervals), NN50 (the number of pairs of successive NNs that differ by more than 50 ms), etc. Another parameter (i.e. the coefficient of variability) which often appeared in the literature is defined by dividing the standard deviation of IBIs by the average IBI (De Waard, 1996). The frequency domain analysis of HRV is more sophisticated and, somehow, is preferred by researches of the ECG-workload studies, since the composed components in the HRV spectrum are assumed to link with some biological control mechanisms (Kramer, 1991). Three frequency components have been defined: a very low frequency range (VLF; 0.02-0.06 Hz), a low frequency range (LF; 0.06-0.15 Hz; also called '0.1 Hz' component), and a high frequency range (HF; 0.15-0.4 Hz). The VLF is believed to be linked to the regulation of the body temperature (De Waard, 1996); LF is assumed to be involved in the

[14] The figure is drawn from the collected ECG data (subject 9) in experiment 2.
[15] The term NN interval is used in place of RR interval (interval between successive R waves) to emphasize the fact that the processed beats are 'normal' beats.

regulation of short-term blood pressure; HF is shown to be related to respiratory fluctuations reflecting parasympathetic influences that are dependent on respiration frequency (Kramer, 1991; Grossman, 1992).

2.3.2 ECG as index of workload

Since HR is easy to obtain and less sensitive to artefacts (Kramer, 1991), it is one of the most popular physiological parameters for mental workload assessment within various environments (Backs & Seljos, 1994; Wilson, 2002a; Brookhuis & De Waard, 1993, 2001, 2010; Mehler et al., 2009). It is assumed that an increased mental workload leads to an increased cardiovascular activity, a heightened cortical energy transformation, and corresponding enhanced metabolic demands (Backs & Seljos, 1994). Therefore, the majority of previous researches consistently demonstrate that the increased workload leads to increased HR (for a detailed review, see Mulder et al., 2004). However, the sensitivity of HR for workload representation was doubted by a few studies (e.g. Wierwille et al., 1985). Lee & Park (1990) stated that physical load causes the dominant effect on HR rather than mental workload. Other researchers demonstrated that HR could be influenced by other factors such as emotion (Jorna, 1992), speech and high G-forces (Wilson, 1992), alcohol (Mascord et al., 1995), etc. Therefore, the underlying mechanism concerning the workload and HR is still unsatisfactorily explained and thus needs further investigations.

Kalsbeek and Ettema (1963) were among the first to investigate the relationship between HRV and mental workload. They demonstrated that HRV is sensitively decreased with increased mental demands in a binary choice task. Later, similar findings were found in many studies (e.g. Backs & Seljos, 1994; Lee & Park, 1990; Mehler et al., 2009). For instance, Lee and Park showed that both increased physical load and mental load could lead to decreased HRV. Brookhuis and De

Waard (2001) stated that HRV shows sensitivity to computational effort but not to compensatory effort, while HR has generally been sensitive to both.

Previous researches indicate that the increased mental effort or task demands might lead to a decrease in the power of the LF and HF components (Mulder, 1980; Mulder et al., 1988; Veltman & Gaillard, 1993). Most studies related to mental workload have focused on the 0.1-Hz frequency band (LF). Mulder (1980) demonstrated that the LF frequency band was the most sensitive to cognitive task demands. Nachreiner et al. (2003) revealed that decreases in the LF reflect increased time pressure or emotional strain. Additionally, another parameter, the ratio LF/HF, could also be sensitive to changes in the operator's workload (Kamada et al., 1992; Wilson, 2002; Murai et al., 2004). However, De Waard (1996) suggested that the spectrum changes in HRV are most sensitive to relatively low workload states. For example, distinguishing task-rest and being less sensitive to differences in task load in high workload regions. Similar conclusions are also found by other studies (Jorna, 1992; Paas et al., 1994; Aasman et al., 1987). There are still a few studies reporting the insensitivity of the HRV to workload (e.g. Wierwille et al., 1985), and that the HRV could also be influenced by other factors such as fatigue (Mascord & Heath, 1992), alcohol (Gonzalez Gonzalez et al., 1992; Mascord et al., 1995), etc.

Finally, the amplitude information from the ECG signal can be used for workload representation. The amplitude of the T-wave (TWA) is shown to decrease with increases in effort (Furedy, 1987; De Waard, 1996). Furedy (1987) concluded that TWA provides a promising psychophysiological path for using cardiac performance measures to aid in the measurement and understanding of the psychological process of mental effort. Nevertheless, such a method is rarely a concern among the workload related literatures.

2.4 Electroencephalogram (EEG)

The EEG is a recording of the brain's electrical activity, in most cases, made from electrodes over the surface of the scalp or from needle electrodes inserted into the brain. One of the first ever reports about EEG was by Richard Caton (1875), who recorded the EEG oscillations from monkeys and rabbits. In 1929, Hans Berger reported the first reliable recording of the EEG from a human scalp and a first categorization of EEG oscillation into alpha (8-13 Hz) and beta waves (14-30 Hz). Here, we refer EEG only to that measured from the head surface.

Generally, the EEG recordings could be categorized into two types: spontaneous activity and the event-related potentials or evoked potential. Spontaneous activity is often referred to the unprovoked occurrence of brain activity, in terms of the absence of an identifiable stimulus, with or without behaviour manifestation. The bandwidth of this signal is from under 1 Hz to over 100 Hz. *Event-related potentials* are time-locked components in the EEG that arise in response to a stimulus, which may be electric, visual, auditory, imaginary, etc. Such signals are often evaluated by averaging a number of trials to improve the signal-to-noise ratio.

EEG recording is completely non-invasive and can be applied repeatedly to patients, normal adults, and children with no risk or limitation (Teplan, 2002). Since the EEG technology holds important advantages such as speed and sensitivity, it has been widely used in medicine, neuro- and psycho-physiological measures, and human factor research (Kramer, 1991). In the following sections, the background to the mechanism of EEG generation, the types of EEG measures, and EEG measures for workload will be systematically introduced.

2.4.1 Mechanism of EEG generation: the brain as a bioelectric generator

It has been estimated that the brain consists of over 10^{11} interconnected neuron cells, each covered with 1,000-100,000 synapses on the surface (Nunez, 1981). The information delivery between the neurons is from dendrites (acting as the input channels) to axons (acting as output channels). The communication between the neuron cells can be realized through permeability of Na+ and K+ ions through the cell wall. Such permeability is a function of the electric potential of the cell wall and the propagation of the ion currents can be unattenuated (Nykopp, 2001).

In a normal situation, the outside of the cell is dominated by the Na+ ions, whereas the inside of the cell is dominated by the K+ ions. Therefore, the ion concentrations outside and inside the cell are different, which enables the cell to relay information between the inside and outside. The potential on the outside of the cell is about 70mV higher than the potential on the inside. The process of decreasing this potential is called depolarization and the increasing process is called hyperpolarization. When depolarization is big enough (i.e. a certain threshold is reached), more Na+ ions will flood into the cell, creating a negative potential outside the cell and a positive potential inside the cell, what is called contrapolarization. This initiates the voltage exchange between the contrapolarization regions and high depolarization regions, and consequently an action potential is delivered. Meanwhile, in the contrapolarization region, K+ ions flood out of the cell and turn electric charge to positive again. Although the action potential can propagate as fast as 100m/s, there is no attenuation in the amplitude of this potential.

The action potential propagation among the neurons is based on the connection between the synapses (located at surfaces of the dendrites) and axons. When action potential reaches a synapse, a chemical transmitter substance is released to

a gap in the synapse. The transmitter diffuses to the cell wall of the postsynaptic cell. This potential is called a postsynaptic potential. The affection between the cells can be either excitatory (increasing the action) or inhibitory (inhibiting the action). In an excitatory synapse, positive ions are flooded into the postsynaptic and the cell wall is depolarized, while in an inhibitory synapse, the potential of the cell wall is maintained below the threshold (Nykopp, 2001).

EEG is measured using scalp electrodes which record the difference in the electric potential between an electrode with an active neural signal and an electrode placed over a supposedly inactive region that serves as a reference. These recordings are the resultant field potentials containing many active neurones. However, the action potential in axons is revealed to contribute little to the scalp surface records as they are asynchronous while the axons run in many different directions. Surface records are thought to be the net effect of local postsynaptic potentials of the cortical cells. EEG measures mostly the currents that flow during synaptic excitations of the dendrites of many pyramidal neurons, a type of neuron found in areas of the brain including the cerebral cortex (Teplan, 2002).

2.4.2 EEG measurement and parameters

Although there are various EEG recording systems in the market, such systems conventionally include four parts: electrodes with conductive media, amplifiers with filters, A/D converter, and recording device. Electrodes are used to read the signal from the scalp; amplifiers increase the magnitude of the microvolt signals into a range which can be digitalized accurately; the converter changes the signals from analog to digital form; and the recorder system (normally personal computer) stores and displays the obtained data (Teplan, 2002). Additionally, a 10-20 system (Figure 2.13) EEG measurement has been adopted by the International Federation in Electroencephalography and Clinical Neurophysiology

(Jasper, 1958). Such a system provides the standardized physical placement of electrodes on the scalp. The electrodes are labelled according to adjacent brain areas: F (frontal), C (central), T (temporal), P (posterior), and O (occipital), with odd numbers on the left side and even numbers on the right side.

Two basic approaches are commonly used for the EEG analysis: (1) the analysis of event-related potential (ERP); and (2) the power spectrum analysis. These two methods have been applied in various experimental or field researches into human cognitive activities.

Event-related potentials

Event-related potentials (ERPs) are the voltage fluctuations that are associated in time with some physical or mental occurrence (Picton et al., 2000). ERP is a complex potential consisting of both time-locked fast and slow components which could both precede an event or follow it (Kotchoubey, 2006). The long-lasting potential shifts, termed slow cortical potentials (SCPs), are usually recorded either prior to an event or after it. The two most well-studied SCP components are contingent negative variation (CNV; Walter et al., 1964) and the Bereitsschaftspotential or readiness potential (RP; Kornhuber and Deecke, 1965). CNVs are normally used to describe the SCP that occurs during the warned foreperiod preceding a motor or mental task, whereas the RP is distinct in the sense that it appears prior to a self-paced voluntary response with the presence of an elicitation or stimulus (Donchin et.al., 1978).

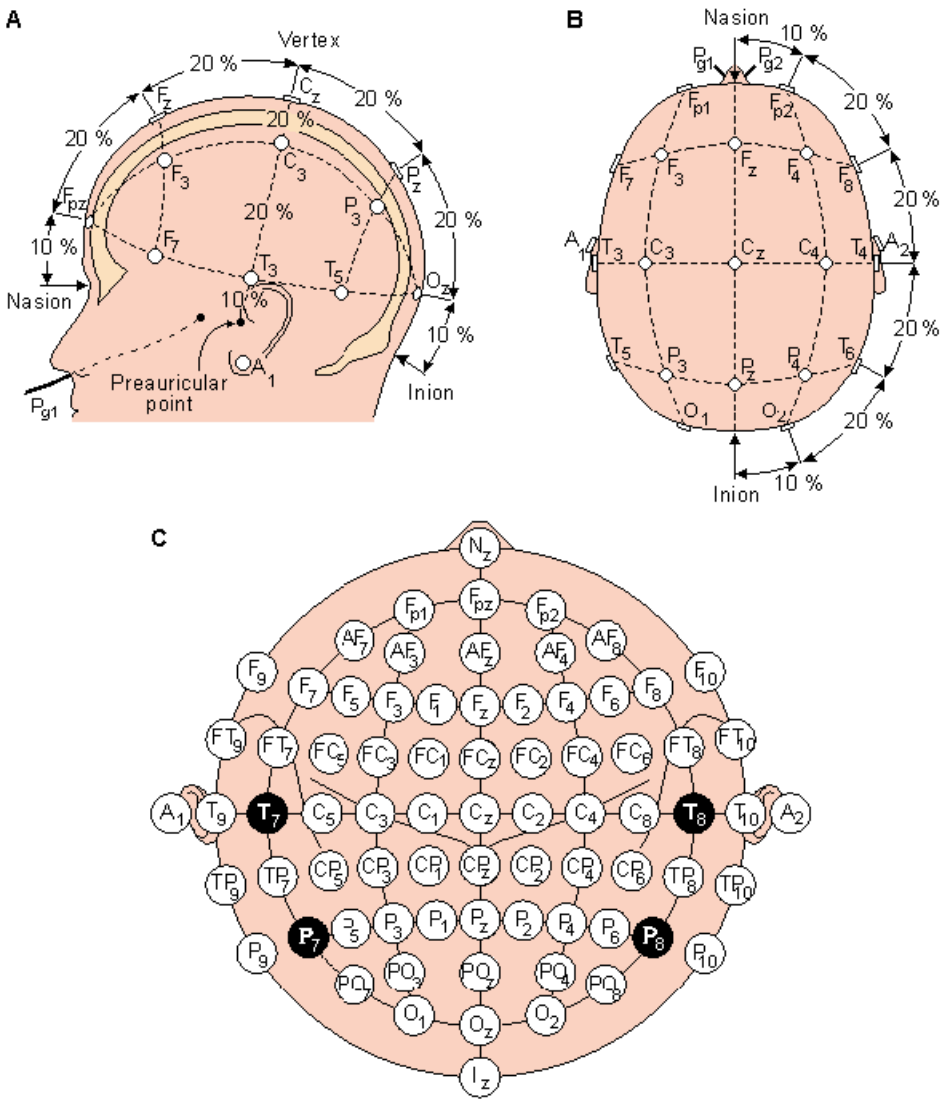


Figure 2. 13 10-20 system for the standardized electrode placement[16]

However, in a narrow and most frequently used sense, ERPs are sequences of relatively fast positive and negative waves following a stimulus, which might last for about half a second (Kotchoubey, 2006). As shown in Figure 2.14, the ERP complex contains some of the typical components such as N1, P1, N2, P2, P3, etc., which are referred to by the typical latency in steps of 100 milliseconds. For instance, N1 refers to a negative deflection which occurs approximately 100 milliseconds after the stimulus onset, while P3 denotes a positive peak at around 300 milliseconds. The occurrences of these components and their latencies in the

[16] Retrieved from: http://www.bci2000.org/wiki/images/1/15/ElectrodePositions1020.PNG

ERP depend on the kind of stimulus. For example, it has been well accepted that P2 is a very prominent and robust ERP peak, reaching its maximum of about 170-180 ms in the auditory and 180-250 ms in visual ERP (Kubová, 2002; Kotchoubey, 2006).

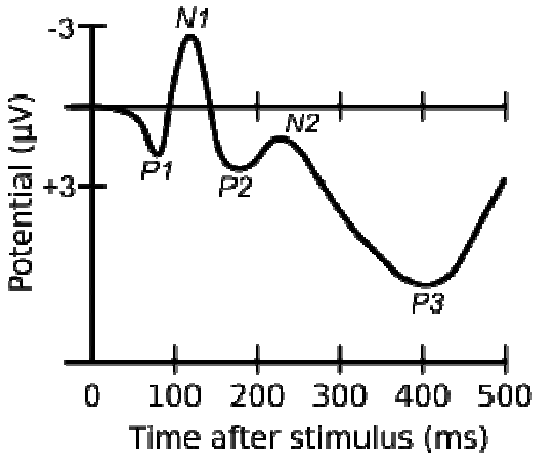


Figure 2. 14 Event-related potentials after stimulus[17]

Since the late 1950s, ERP has been established as a psychophysiological approach to provide information about the cognitive processing of an event or a stimulus in the brain. ERP components are supposed to allow for the possibility to obtain information about how the intact human brain processes signals and prepares actions. Kotchoubey (2006) stated that negative SCP shifts reflect provision of cortical resources for an anticipated activity, and that positive SCP shifts are related to consumption of these resources in actual activity. Similarly, in-post stimulus ERPs, a negative wave reflects depolarization of superficial dendrites and the preparation of resources for the information to come, whereas the consumption of these resources is reflected in the following positive wave related to an increasing activity in the deeper cortical layers. Therefore, it can be supposed that, for both pre- and post-stimulus, the negativities are attention-related reflecting the feedforward anticipatory activity (see Herrmann &

[17] Retrieved from: http://en.wikipedia.org/wiki/Event-related_potential

Knight, 2001; Kotchoubey, 2006), while positive ERP components reflect feedback information pick up (Banquet & Grossberg, 1987; Kotchoubey, 2006).

More specifically, Hillyard (1973) suggested that CNVs are responses to four types of cognitive processes: holding a motor response in readiness; preparing for a perceptual judgement; anticipation of a reinforcer, positive and negative; preparing for a cognitive decision. For the post-stimulus ERPs, for instance, N100 shows link to a person's arousal (Nash and Williams, 1982) and selective attention (Hillyard et al., 1978). The P100 effect was supposed to be a reflection of the "cost of attention" (Luck et al., 1994); N200 was reported to reflect the detection of stimulus omission with a modality-specific property (Donchin et al., 1978); P300 was argued to be related to stimulus evaluation processes and independent of response selection and execution processes (Donchin et al., 1978).

Although ERPs serve as important adjuncts to studies of human information processing, a fundamental problem with this method is the signal-noise ratio. The magnitude of the ERP signal is around 5-10 μV, which is far smaller than the amplitude of the background EEG (0-200 μV; Schandry, 2006). Therefore, the classic approach for ERP extraction is to average the signal over a number of trials in order to obtain a stable response with a sufficient signal-to-noise ratio.

EEG spectrum analysis

The oscillatory activity of the spontaneous EEG is typically categorized into five different frequency bands: delta (0-4 Hz), theta (4-8), alpha (8-12), beta (12-30) and gamma (30-100 Hz), as shown in Figure 2.15. These frequency bands are suggested to be a result of different cognitive functions.

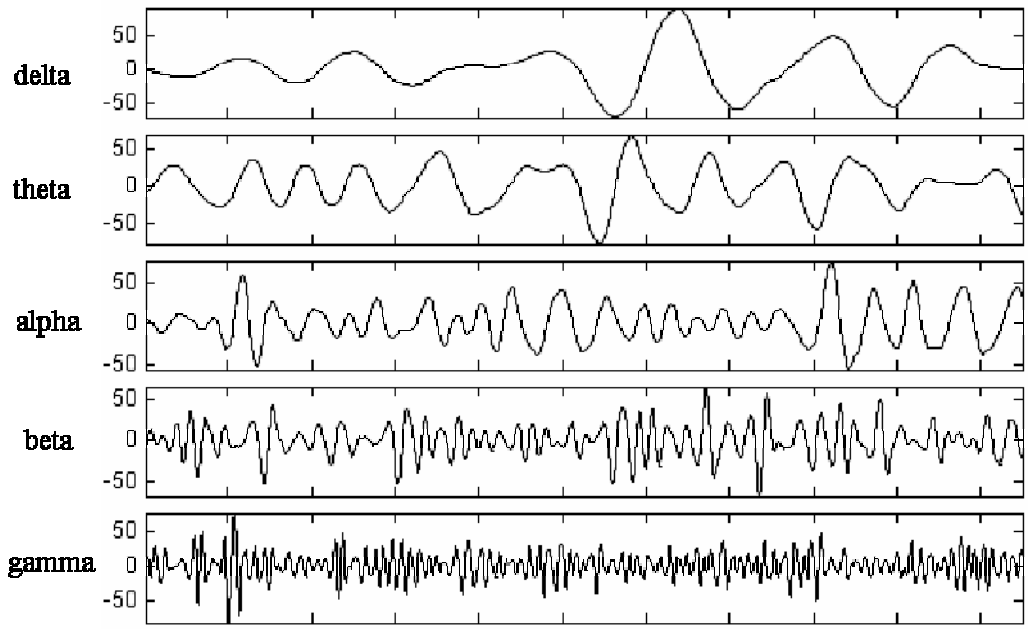


Figure 2. 15 EEG frequency components delta, theta, alpha, beta, and gamma[18]

- *Delta (0 -4 Hz):* Delta activity is characterized as high amplitude and low frequency. It is usually associated with the slow-wave sleep in the sleep research. It is suggested that delta waves represent the onset of deep sleep phases (stage 3 and stage 4) in healthy adults (Rechtschaffen and Kales, 1968). In addition, contamination of the eye activity is mostly represented in the delta frequency band (Hagemann, 2008).

- *Theta (4-8Hz):* The generation of theta power is associated with the hippocampus (Buzsáki, 2002) as well as neocortex (Cantero et al., 2003). The theta band is thought to be associated with deep relaxation or meditation (e.g. Hebert & Lehmann, 1977; Kubota et al., 2001) and it has been observed at the transition stage between wake and sleep (Hagemann, 2008). However, theta rhythms are suggested to be important for learning and memory functions (Maclver, 2003, p. 156; Sammer et al., 2007), encoding and retrieval (Ward, 2003) which involve high concentration (Hagemann, 2008). It has also been

[18] This figure is generated through the wavelet transform (Daubenchies D6 wavelet, 6-level) based on experiment data of subject 4 in experiment 1. The delta, theta, alpha, beta, and gamma are correspondent to the wavelet decompositions (see Lei, Welke, Roetting, 2009 c).

suggested that theta oscillations are associated with the attentional control mechanism in the anterior cingulated cortex (Kubota et al., 2001; Smith et al., 2001) and is often shown to increase with a higher cognitive task demand (e.g. Gundel & Wilson, 1992; Gevins et al., 1997).

- *Alpha (8-12Hz):* Alpha band activity is found at the visual cortex (occipital lobe) during periods of relaxation or idling (eyes closed but awake). It is characterized by high amplitude and regular oscillations with a maximum over parietal and occipital electrodes in the continuous EEG. The modulation of alpha activity is thought to be a result of resonation or oscillation of the neuron groups (Lopes da Silva et al., 1980; Smith et al., 2001). High alpha power has been assumed to reflect a state of relaxation or cortical idling. However, when the operator devotes more effort to the task, different regions of the cortex may be recruited in the transient function network leading to passive oscillation of the local alpha generators in synchrony with a reduction in alpha power (Smith et al., 2001). Recent results suggested that alpha is involved in auditory attention processes and the inhibition of task irrelevant areas to enhance signal-to-noise ratio (Cooper et al., 2006; Klimesch et al., 2007; Hagemann, 2008). Additionally, some researchers divide the alpha activity further into sub-bands to achieve a finer grained description of its functionality (e.g. Klimesch et al., 1999). For instance, the “mu” band (10-12 Hz) occurs with actual motor movement and intent to move with an associated diminished activation of the motor cortex (Dooley, 2009).

- *Beta (13-30Hz):* The beta wave is predominant when the human is awake. Spatially, it predominates in the fontal and central area of the brain. It has been described that the high power in the beta band is associated with the increased arousal and activity (Nykopp, 2001). Dooley (2009) pointed out that the beta wave represents cognitive consciousness and an active, busy, or

anxious thinking. Furthermore, it has been revealed to reflect visual concentration and the orienting of attention (Birbaumer & Schmidt, 1996). The beta band can be further divided into several sub-bands: low beta wave (12.5-15 Hz); middle beta wave (15-18 Hz); high beta wave (> 18 Hz). These three subbands are associated with separate physiologic processes. For instance, the high beta waves are suggested to be linked with the dopaminergic system (Gruzelier et al., 1990; Hagemann, 2008), while the low beta activities are thought to reflect the inhibition of phasic movements during sleep (Sterman, 1981; Hagemann, 2008).

- *Gamma (>30Hz):* The gamma band is the fastest activity in EEG and is thought to be infrequent during waking states of consciousness (Dooley, 2009). It is reported that gamma waves are associated with perceptual blinding problem (Gray et al., 1989; Singer & Gray, 1995). More specifically, Tallon-Baudry et al. (2005) revealed that areas of lateral occipital cortex and fusiform gyrus play an important role in visual stimulus encoding and show large gamma oscillations differently affected by attentional modulation. Recent studies reveal that gamma is linked with many other cognitive functions such as attention, learning, memory (Jensen, et al., 2007), and language perception (Eulitz et al., 1996). Additionally, Sederberg et al. (2007) demonstrated that verbal memory formation led to an increase in gamma oscillations when analyzing intracranial recording data from epilepsy patients.

2.4.3 EEG as an index of mental workload

Characteristic changes in the EEG reflecting levels of mental workload have been identified in the last decades. In general, two categories of information can be extracted for the representation of the human operator mental workload: ERPs and EEG spectrum modulations.

ERPs measurement of mental workload

As discussed in the session number, ERPs provide sensitive information for indexing the cognitive processing of the human brain. In the last decades, a large number of studies using ERPs to evaluate the mental workload have been conducted which proved that the amplitude and latency of P300 provide powerful and effective tools for the assessment of mental workload (Johnson, 1986; Schultheis & Jameson, 2004; Raabe et al., 2005). The P300 component was discovered by Sutton et al. (1965) and is characterized by a slow positive deflection with a mean latency of 300 ms after the stimulus onset depending on the stimulus type and complexity (Polich, 1987). The P300 typically has a centroparietal scalp distribution, with a maximum over parietal cortex. For workload assessment, two features from the P300, the latency and amplitude, are frequently used.

P300 latency: It is suggested that P300 latency provides a chronometric index for the duration of perceptual processing (Leuthold & Sommer, 1998). Kramer and Parasuraman (2007) also pointed out that the latency of P300 reflects the timing of stimulus identification and categorization processes. Previous research indicated that increasing the mental workload may lead to an extension of the P300 latency. Kutas et al. (1977) stated that increasing the difficulty of identifying the target stimulus also increased the latency of the P300 wave. Such conclusion was confirmed also by Fowler (1994). However, increases in the difficulty of response selection do no affect P300 latency (Magliero, 1984). This led to a discussion on whether the latency of the P300 provides a relatively pure measurement of perceptual processing and categorization time, independent of response selection and execution stages (Kutas et al., 1977; McCarthy and Donchin, 1981).

P300 amplitude: It has also been assumed that the amplitude of P300 is proportional to the amount of attentional resource allocation for the task performance (Johnson, 1986). This assumption is in line with the findings in the oddball paradigm that the amplitude of P300 is sensitive to the probability of the presentation of stimulus. For instance, in an oddball paradigm, in which the occurrence of one stimulus is much higher than another (e.g. 80% vs. 20%), the P300 amplitude was found to be much higher for the rare stimulus than that for the frequent stimulus (Polich, 1990; Duncan-Johnson & Donchin, 1977). Gopher and Donchin (1986) suggested that the P300 amplitude could index the perceptual/central processing load, until the moment performance declines, in which case the amplitude remains unaffected.

It is assumed that the amplitude of P300 may show different changes in the single and dual task performance. In a primary-task-only-condition, it was suggested that the P300 amplitude increases with task complexity (De Waard, 1996). In a dual-task paradigm, the diversion of processing resources away from target discrimination leads to a reduction in P300 amplitude (Kramer & Parasuraman, 2007). For example, Wickens et al. (1977) showed that the amplitude of P300 decreased when a primary task (tone counting) was combined with a secondary task (visual tracking). Schultheis and Jameson (2004) conducted a study of the difficulty of text presented in hypermedia systems, with a view to investigating whether the text difficulty could be adaptively varied depending on the cognitive load imposed on the operator. They used an auditory oddball task with easy and difficult versions of a text, and then measured the pupil diameter and the P300 to the oddball task. They found that the P300 amplitude, not the pupil diameter, was significantly reduced for difficult hypermedia conditions. In this way, these findings support the view that P300 reflects processing resources allocation associated with perceptual processing and stimulus categorization (Fu and Parasuraman, 2007). However, Fu and Parasuraman (2007) also stated that the

response-related contributions to workload are not reflected in the P300 amplitude.

The different changes in P300 for both the single task and dual task may seem contradictory. But in fact, these observations are in line with the theory that the amplitude of P300 reflects the amount of attentional resources allocation. In a single task, the P300 is actually elicited by the primary-only-task. The increasing task load may lead to an increased attentional resources allocation for the task performance, which may produce higher P300 amplitude. However, in a dual-task condition, the resource allocation to one task may lead to a withdrawal of resources to another task. Wickens et al. (1983) demonstrated that while P300 to a secondary task decreased with increase in the task difficulty, the P300 to an event embedded within the primary task increased. Therefore, it is assumed that there is a reciprocal relation between the P300 amplitude and resource allocation between primary and secondary tasks. This pattern of P300 reciprocity is consistent with resources tradeoffs predicted by the multiple resources theory of workload proposed by Wickens (2002).

Other components: Besides the P300, other earlier ERP components, e.g. N1, P1, P2, have also been shown to be sensitive to the workload in a few studies (Hink et al., 1977; Parasuraman, 1985; Fu & Parasuramann, 2007; Fu et al., 2010b). Parasuraman (1985) conducted an experiment which required participants to perform both visual and auditory tasks concurrently, and systematically varied the priority to be placed on one task relative to the other. The effect of information load was modulated by changing the presentation rates (slow vs. fast). The results indicated that the amplitudes of visual N160 and auditory N100 components decreased directly with task priority, suggesting that N1 is involved in the attentional resource allocation between visual and auditory perception channels. However, these reductions were only observed in high information load conditions

(fast presentation rate). This suggests that the occurrence of such attentional resource allocation requires a certain amount of workload (Fu & Parasuraman, 2007). More recently, Fu et al. (2010b) reported a spatially-cued study investigating the effect of attentional load on the visual C1[19] and P1 components. Their results revealed that the amplitude of P1 in the parietal area is affected by the attentional load in a decreasing manner when the load is increased. However, there was no attentional load effect on C1. Similar findings that the P1 component can be modulated by visual attention are found in other studies (Hillyard & Münte, 1984; Fu & Parasuramann, 2007).

EEG spectrum modulation with mental workload

An extensive body of literatures exists concerning the EEG spectrum modulation according to the variation of cognitive workload and the allocation of mental effort in both laboratory (Gundel and Wilson, 1992; Berka et al., 2007; Lei et al., 2009a, b, c; Lei & Roetting, 2011) and applied settings (Wilson, 2001, 2002b; Kohlmorgen et al., 2007).

Quantitative changes in the EEG spectrum parameters with the task loading have been derived after a period of observation. Gevins and Smith (2000) investigated the spectrum compositions of the ongoing EEG which displays regular patterns of load-related modulation during n-back task performance. They showed the power spectrum density in the 4-14 Hz range at the frontal midline (Fz) and parietal midline (Pz) scalp location computed from the continuous EEG recorded from a low-load (0-back) and high-load (2-back) versions of a spatial n-back task. Significant differences in the spectral power as a function of task load were observed. Specifically, at the Fz site the theta power was increased during high-load task relative to low-load task, whereas alpha power tended to be

[19] The first visual cortical response originating in the striate cortex.

attenuated in the high-load task compared to low-load tasks. Consistent results have been found not only in similar working memory (WM) task (Gundel and Wilson, 1992; Gevins et al., 1997), but also in more complex cognitive tasks (Smith et al., 2001; Wilson, 2002b; Wilson & Russell, 2003a). Smith et al. (2001) recorded continuous EEG while 16 participants performed versions of the compute based flight simulation task, the Multiple-Attribute Task Battery (MATB; Comstock & Arnegard, 1992), in low, moderate and high difficulty. As task difficulty increased, frontal midline theta EEG activity increased while parietal midline alpha decreased. In field research, Wilson (2002b) reported a study involving ten pilots who flew an approximately 90-minute scenario containing both visual and instrument flight conditions. Multiple variables including EEG parameters were analyzed. Wilson found that parietal alpha band showed significant reduction in high workload condition, but increase of theta power could only be observed at a few scattered electrode sites. However, disputing voices on theta power can be also heard. Decreases in theta activity were found with transitions from single- to dual-tasks (Sirevaag et al., 1988). Pigeau et al. (1987) revealed that theta power initially increases with increments in the task difficulty of an additional task and then decreases at high levels of difficulty. These observations were addressed by Kramer (1991).

Compared with theta and alpha, studies investigating the variations of other EEG spectral compositions (such as delta, beta and gamma) with the task load are relatively rare. These components generally show no systematic changes when the task load changes. For instance, in Wilson's field flight-task study (2002b), a significant increase of delta band power in the central electrodes C3, Cz, C4 as well as in parietal electrode P4 was found as the task difficulty increased. Additionally, significant reductions in beta band activity were also observed from Wilson's study in some task conditions, though not in all conditions. However, the EEG beta power modulation with workload seems task dependent. Brookings el

al. (1996) found increased beta activity during the complexity manipulations. But Dussault et al. (2005) revealed that rest sequences would result in higher beta power than active segments in a simulated flight task. More recently, Doyle et al. (2009) showed that gamma power was positively associated with increased levels of workload in a satellite management decision-training task. Unfortunately, there is no systematic support for such a conclusion among the literature.

From a neurophysiological perspective, the changes in theta power and alpha power with the workload match the underlying generation or functionality mechanism. Sammer et al. (2007) suggested that theta activity is associated with numerous processes including working memory, problem solving, encoding, or self monitoring. Sammer and his colleagues revealed that EEG-theta was enhanced in the frontal area by mental arithmetic-induced workload in an EEG-fMRI study. However, they also found that several other brain regions (superior temporal areas, superior parietal) were also involved in the activation of theta activity, which implies that theta represents comprehensive functional brain states rather than specific processes in the brain. Topographic analysis has indicated that this task-load-related theta activity tends to have a sharply defined potential field with a focus on the anterior midline region of the scalp (Inouye et al., 1994; Gevins et al., 1997). The anterior cingular cortex has been indicated as the generating sources of theta oscillations in both EEG (Gevins et al., 1997; Smith et al., 2001) and magnetoencephalographic (MEG; Ishii et al., 1999) studies. This region is thought to be part of an anterior brain network that is critical to attention control mechanisms and is activated by the performance of complex cognitive tasks (Posner et al., 1990, 1992; Smith et al., 2001). Paus et al. (1998) found that the major source of variance affecting the activation of the anterior cingular cortex was associated with changes in task difficulty after reviewing 100 PET studies. It seems that any increased complexity or difficulty in

tasks involving enhanced attentional demand may lead to the activation of theta activity in the frontal area.

Alpha oscillation was found to systematically decrease in power as the task load increases. This inverse proportion has been found in numerous earlier studies (Sterman et al., 1994; Gevins et al., 1997, 1998) and is consistent with current understanding of the underlying neural mechanisms in the generation of the alpha rhythm. Alpha activity is presumed to be generated by coherent local groups, in which local coherence is probably driven through intracortical connections between nearby neurons with individual cells acting as either resonators or oscillators (Smith et al., 2001). When the brain is in relative rest, a high proportion of alpha generators come to oscillate in phases yielding a large alpha rhythm (Lopes da Silva, 1980). As task demands increase, different regions of the cortex may be recruited in the transient function network, with a decline in the overall proportion of local alpha generators that passively oscillate in synchrony with the reduction in alpha power (Smith et al., 2001). In this perspective, alpha power can be assumed to reflect the recruitment of neural resources for task performance.

Unlike the underlying mechanism of theta and alpha rhythm, unfortunately, there are no well-accepted hypotheses concerning the underlying neural mechanisms for the modulation of other band powers induced by workload.

2.5 Psychophysiology-driven adaptive aiding design

Research into adaptive aiding in the human-machine-system based on the psychophysiology is a particular interest in recent years. This emerging area focuses on development in the human cognitive states recognition and how that knowledge can be used to improve performance in real-world environments. Parasuraman (2003) pointed out that a better understanding of the brain perception

and processing of information can lead to better designs of equipment, systems, and tasks by enabling a tighter match between task demands and the underlying brain process. One application is to integrate the psychophysiology-driven interface into an adaptive system design (e.g. adaptive automation), which could help reduce the operator's task demands and workload and could make work and life easier.

This work has been attributed to theoretical bases (see Parasuraman & Riley, 1997; Parasuraman et al., 1999), model development (Parasuraman et al., 1999; Parasuraman, Sheridan & Wickens, 2000), empirical knowledge (i.e. Byrne & Parasuraman, 1996). Parasuraman and Riley (1997) discussed important issues such as overreliance on automation and the potential for skill degradation that may accompany automation. Parasuraman et al. (1999) reported a model for finding the optimal task demand on the human operator to determine optimal stages for switching of tasks between human and machine. This model has served as a theoretical framework for subsequent empirical work (e.g. Parasuraman, Consenzo & DeVisser, 2009).

In the last section, it was discussed that EEG spectral components vary in a predictable pattern according to the variation in workload. Many studies attempted to combine the EEG parameters for a reliable index of neural activity, for example, using the ratio of the different band powers (Brookhuis & De Waard, 1993; Pope el al., 1995; Prinzel et al., 2000, 2001). Pope et al. (1995), who reported the first brain-based adaptive system, established a system to index the task engagement based upon ratios of EEG power bands (alpha, beta, theta, etc.). They recorded the EEG data from several locations on the scalp and sent them to a Labview Virtual Instrument to calculate the spectral components every 2 seconds. This helped to infer a task engagement index, which could be used to change a tracking task between automatic and manual models. Pope and his colleagues

evaluated several engagement indices and found that the index beta/(alpha and theta) showed the greatest difference between the positive and negative feedback conditions. As an extension of Pope's study, Prinzel et al. (2000) built a closed-loop system to examine the psychophysiological measures for adaptive task allocation based on the EEG index beta/ (alpha and theta) which also showed some optimistic results. Similar results have also been reported in other studies (Freeman et al., 1999; Scerbo et al., 2003).

Another approach towards real-time assessment of mental workload, instead of the EEG spectral components, is to use Brain-Computer-Interaction (BCI) technology, e.g. linear discriminate analysis (LDA), support vector machine (SVM), artificial neural network (ANN), etc. These studies classified workload into several levels (e.g. low, moderate and high) using the various EEG parameters in either a simple, single-task (Wilson & Fisher, 1995; Gevins et al., 1998; Nikolaev et al., 1998; Gevins & Smith, 1999) or complex tasks with skilled operators (Russell, et al., 1996; Russell & Wilson, 1998; Wilson & Russell, 2003a; Grimes et al., 2008; Honal & Schulz, 2008; Heger et al., 2010; Putze et al., 2010). For instance, Wilson and Russell (2003b) employed artificial neural networks (ANN) to classify mental workload using five log spectrum power of five frequency bands (delta, theta, alpha, beta, gamma) plus several EOG features in a Multi-Attribute Task Battery. The mean classification accuracies were respectively 85%, 82%, and 86% for the baseline, low task difficulty, and high task difficulty conditions. Later, Honal and Schulz (2008) implemented learning approaches that derive task demand from the user's brain activity using EEG. By using SVM, they distinguished between high versus low task demand with an accuracy of 92.2% in session dependent experiments, 87.1% in session independent experiments, and 80.0% in subject independent experiments.

There are also other examples integrating such an approach in the human-machine-system for adaptive design. Wilson and Russell (2004) reported a study with augmented cognitive system design to moderate workload. Operators were asked to perform a target identification task with a simulated uninhabited combat air vehicle under different levels of workload. Multiple variables including EEG, heart beat, blink, and respiration rates were recorded. The ANN was used to filter the physiological data to distinguish between low and high workload in real time. And the results were used to trigger some changes in the task to modulate workload. Comparisons among adaptive aiding, no aiding, or random aiding revealed some performance benefits but with lower ratings of subjective workload for the adaptive aiding condition under the more difficult conditions. More recently, Ting et al. (2010) proposed a framework for the online monitoring and adaptive control of automation in complex and safety-critical human-machine systems using psychophysiological markers. An adaptive fuzzy model was used to link the heart rate variability and task load index with operator states in an automation-enhanced Cabin Air Management System. Such model was used as the basis for an online control system to modify the level of automation under which the system may operate. A series of human volunteer studies showed that such a system seems promising for the improvement of performance.

2.6 Driving task and driver task load

Although car driving is one of the basic skills of daily life in modern society, it is a complex behavioural task. The traffic system as a whole consists of three interactive parts: vehicle, driver, and road environment; while the driving task actually requires a driver-vehicle-environment (DVE) interaction. For a comprehensive understanding of the features of driver mental workload, the analysis of psychological mechanisms underlying the driving task and driver

behaviour is essential.

2.6.1 Driving task and driver mental workload

Driving task analysis and driver behaviour modelling in the field of human factors have been conducted for over 60 years, and various theoretical models have been developed, aiming to improve driving safety and driver education. Carsten (2007) pointed out that the variety of models of driving task is almost as numerous as the number of authors who have contributed to the models. Generally, a hierarchical model proposed by Michon (1985) serves as the basic framework for the other models and theories (Weller et al., 2006).

The driving task can be described using a hierarchical three-level model. Michon (1985) divided the driving task into three-levels: a strategic level, a tactical level (or manoeuvring level, see De Waard, 1996), and a control level. The top level is the strategic level, in which strategic decisions are made, such as the choice of transport tools, the goals of the trips, selection of the routes, etc. The middle level, the tactical level, includes immediate decisions for a driving manoeuvre according to local situations, e.g. choice of driving lane and driving speed, decisions for obstacle avoidance and behaviours of other traffic participants, etc. The lowest level is the control level consisting of the basic vehicle-control processes, such as lateral control, speed adjustment, steering, etc. In this level, the driver is mostly in an automatic mode (Hagemann, 2008). De Waard (1996) stated that driving performance measures could be connected to these three levels. For instance, steering performance is linked with the lowest level; car following performance and mirror looking behaviour are associated with the tactical level; errors in route choice reflect performance at the strategic level.

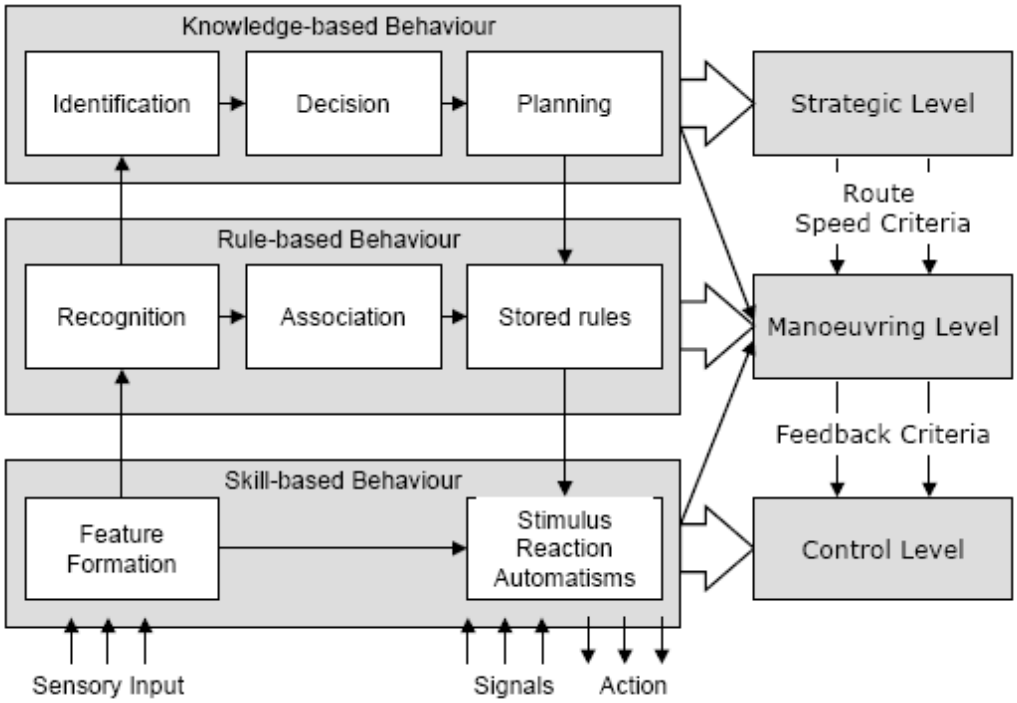


Figure 2. 16 Linking Michon's hierarchical model with Rasmussen's (1983) knowledge-rule-skill model (adapted from Weller et al., 2006)

Donges (1999) linked Michon's hierarchical model with Rasmussen's (1983) knowledge-rule-skill model for human performance. As shown in Figure 2.16, this connection was reinforced by Weller et al. (2006). The left part of this figure shows the task levels according to Rasmussen's (1983) theory, in which human behaviour is classified into knowledge-based, rule-based, and skill-based behaviour. Knowledge-based behaviour is often caused by complex demand situations, which require the analysis of various behaviour alternatives with respect to their suitability for reaching the goal. This analysis undergoes mental processes based on the available knowledge. The level of rule-based behaviour is related to situational conditions, which have already frequently occurred during earlier opportunities. The human driver has already prior experience and possesses certain rules for handling these situations. The skill-based behaviour is characterised by automated and unconscious stimulus-response mechanisms acquired from well-learned procedures (Hagemann, 2008).

In associating Rasmussen's model with Michon's model, the skill-based level is mainly applied to the control level in the form of automatic action patterns (Ranney, 1994). Rule-based behaviour dominates the manoeuvring level, in which standard interactions with other road users frequently occur. Knowledge-based level dominates the strategic level for the experienced driver even when s/he is confronted with unfamiliar traffic networks, while it is applied as well in the control-level for the novice driver whose driving skills are not fully developed (Ranney, 1994).

De Waard (1996) pointed out that demands from all three levels in Michon's model can exceed capacity, and may affect the driver's mental workload and performance. Individual factors, such as experience and age, may affect the driver workload at each level. For instance, a student driver lacks the basic skills at the control-level tasks, and the workload related to vehicle control is high. This may result in performance decline at higher level tasks, such as mirror-checking (De Waard, 1996). Additionally, the environment factors, e.g. unfamiliar road conditions or driving in heavy traffic in a city abroad, may pose high demands on visual and central sources when driver performs manoeuvre-level tasks that affect the performance of other tasks.

Therefore, sources both inside and outside the vehicle may affect the driver's workload. De Waard (1996) categorized the factors affecting workload into three categories: driver state factors, driver trait factors, and environmental demands. Driver state factors include various driver functional states, e.g. monotony, fatigue, alcohol, etc. Such states are supposed to increase workload by a reduction in capacity (Schneider et al., 1984; Wierwille & Eggemeier, 1993). Driver trait factors consist of experience, age and strategy. Experience and an appropriate strategy may reduce workload (Patten et al., 2006; Parasuraman and Hancock, 2001). However, an old driver may experience higher workload due to decreased

capacity (Verwey, 2000; Son et al., 2011). Both state factors and trait factors address the internal workload sources of the driver.

In comparison, environmental factors reflect the external workload sources including road environment demands, traffic demands, vehicle ergonomics, in-vehicle technologies, automation, etc. Environment parameters, such as road curvature, traffic density, weather, etc., may contribute to the external demand on the driver (Baldwin & Coyne, 2003, 2005; Baldwin et al., 2004; Musa and Navin, 2005; Horrey and Lesch, 2009). Baldwin & Coyne (2005) revealed that driver workload increased when driving in the presence of fog versus clear visibility. More recently, Horrey and Lesch (2009) scored a narrow road as the highest demanding condition for driving within various road situations.

Another important factor for driver workload, which is not included in the De Waard's categorization but often addressed by researchers, is the driving speed. Senders et al. (1967) and McDonald et al. (1975) both concluded that the attentional demand of the road increased with the operating speed. A similar result was also obtained in Cnossen's study wherein the mental workload increased according to the speed level (Cnossen et al., 2000). More recently, Musa and Navin (2005) proposed a rate of information processing demand (RID) combining road complexity and driving speed in terms of time constraint. They proposed to regulate the driving speed according to the task demand from the driving environment. Fitzpatrick et al. (2010) also pointed out that high driving speed may lead to increased attention for vehicle control and may slow down the driver's response to hazards. Additionally, the driving speed is frequently used as an important indictor reflecting the driver's workload or task demand condition, since the driver may adjust the speed accordingly when the road and traffic situation becomes more complex (Törnros & Bolling, 2006).

With the boost in in-vehicle technologies (IVTs), e.g. navigation system and cellular phone, human factor issues concerning such devices are also being increasingly addressed. Lenneman & Backs (2009) pointed out that such IVTs pose a new challenge for drivers to cope with multiple attention-demanding tasks which might not relate to the primary driving task. In such situations, the combined use of IVTs may result in driver overload (Verwey, 1990), especially in a complex driving environment (i.e. high traffic density or poor weather). Some reviews concerning workload issues induced by IVTs could be found in Collet et al. (2010). On the other hand, the highly developed automation technology tends to bring the driver to another end, a 'low workload' state, rather than a 'high workload' state (Parasuraman et al., 1999; Parasuraman & Hancock, 2001). In a highly automated vehicle, driver functions are taken over by technology. This will lead to monotony in task performance, in which the driver may eventually become deactivated.

To summarize, driver workload is a result of the simultaneously combined effect of various factors, both internal and external to the driver. This poses a particular challenge to the quantification of driver workload using EEG signals, since the underlying neural mechanism induced by different workload factors is complex and still unknown. In the following section, some preliminary findings concerning the neural correlates of driving from previous neuroscience or neuroimaging studies are dealt with.

2.6.2 Neural correlates of driving

Driving engages a variety of cognitive functions such as perception, attention, motor control, working memory, and decision making, all of which have been integrated in a number of psychological models (Michon, 1985). It is also well known that humans, who recover from traumatic brain injury, are often thought to

be unfit for driving due to deficits in remembering, learning, and planning (Groeger, 2002). However, the underlying neural correlates of these functions involved in driving are still unknown.

Recently, a large number of neurophysiological or neuroimaging studies have been conducted to investigate neural interpretations of cognitive activities involved in driving (Calhoun et al., 2002, 2007; Graydon et al., 2004; Horikawa et al., 2005; Uchiyama et al., 2003; Walter et al., 2001; Spiers & Maguire, 2007). These studies reveal some preliminary findings concerning the neural correlates of driving. Generally, these studies could be categorized into four groups: comparison of driving with no driving, speed related neural activity, safety related driving, and driving components (e.g. starting, stopping, steering, etc.).

- ***Comparison of driving and no driving:*** The driving and no-driving paradigm comparison is interpreted as the attention difference. Schier (2000) reported an EEG study which showed greater alpha power in frontal lobes during driving than during driving replay. This could be interpreted as a reduction in attention during the replay task. Later, Walter et al. (2001) studied 12 healthy subjects with fMRI while they performed a driving simulation task. They manipulated two conditions, the active driving (participants steered the car themselves) and passive driving (a person from outside the scanner steered the car). The results of this study revealed that active driving was associated with an increased activation in the left sensorimotor cortex and cerebellar region. Additionally, in both conditions there were activations in the right occipital and parietal regions, which were thought to be linked to perceptual processes during driving. Compared to passive driving, activity in active driving was reduced in numerous brain regions including the middle temporal area/medial superior temporal area. Walter and his colleagues concluded that simulated driving requires mainly perceptual-motor integration while the

limited cognitive capacity model of driving is inadequate to describe the complex neurophysilogical processes in driving. Calhoun et al. (2002) employed an independent components analysis to investigate brain activity during driving using fMRI. They manipulated three task conditions: (a) an asterisk fixation task; (b) active simulated driving; and (c) watching a simulated driving. Their results demonstrated some common patterns. Firstly, the most active areas during driving were in primary visual and higher order visual/cerebellar areas. These areas demonstrated less activation in the watching task. The primary visual areas have been interpreted as sensory acquisition (Bower, 1997) as well as attention/anticipation (Akshoomoff et al., 1997). Meanwhile, the visual/cerebellar areas activated by driving are thought to be associated with orientation and complex scene interpretation or memory processing. Secondly, the anterior cingulated, median frontal, and other frontal areas demonstrated exponential decrements during driving and rebounded during fixation, which is consistently in line with Walter's findings (Walter et al., 2001). Finally, Calhoun and his colleagues demonstrated that the cerebellar/motor increased during driving whereas the frontoparietal areas decreased during driving. The activation of cerebelar/motor areas was consistent as well with Walter's findings.

- ***Speed related neural activity:*** The speed related neural activity is mainly associated with the attention modulation (Calhoun et al., 2002). Calhoun's study investigated the speed associated with neural activities. The faster group elicited decreased activity in both frontoparietal and anterior cingulated areas. The former area is thought to be associated with vigilance, whereas the latter is implicated with error monitoring and inhibition. Calhoun and his colleagues suggested that during fast driving conditions, the vigilance component changes more so that the error correction and disinhibition component decreases at a faster rate.

- ***Performance or safety related neural activity:*** Horikawa et al. (2005) conducted a PET study that investigated the correlations between the regional cerebral flow and the time required to complete a course as well as the number of crashes. Horikawa and colleagues found that there were activations in the parietal-occipital areas associated with visual processing, and activations in the cerebellum linking with motor control. Interestingly, they showed that impaired driving was related with increased regional cerebral blood flow (rCBF) in the thalamus, midbrain and posterior cingulated gyrus, which is consistent with the interpretation that such areas are involved in processes of attention, high vigilance and internal timing. Therefore, Horikawa and colleagues concluded that these areas might be associated with driving performance maintenance. Spiers and Maguire (2007) provided some evidence for the neural basis of crashes/near collisions. Their results showed that a number of brain regions were associated with unprepared actions, e.g. hazard responding or collision avoiding. For instance, there was an increased activity in the supplementary motor area (SMA) and lateral dorsal premotor regions. Meanwhile, increased activity in the posterior parietal cortex, medial occipital cortex and lateral posterior temporal gyrus was observed when avoiding collisions, which might be related to the increased attention allocated to the approaching object. Additionally, increased activity in the mid cingulated and insula was also found. This may be related to the particularly arousing nature of urgent events. A more recent study investigated the relationship between activation in the dorsolateral prefrontal cortex (DLPFC) and the driver's risk behaviour (Beeli et al., 2008). This study revealed that the stimulation of DLPFC influenced driving behavior while excitation of the DLPFC leads to a more careful driving style in virtual scenarios.

- ***Driving components and neural activity:*** Spiers and Maguire (2007) systematically examined the brain regions supporting the prepared action of

different driving components based on a second by second analysis using fMRI. All types of prepared action activated a core network of brain regions including cerebellum, pre-SMA/SMA, posterior cingulated, medial parietal, medial and lateral occipital cortex. However, there were also certain regions which became more active only during specific prepared actions: significant increased activation in the left dorsolateral precentral gyrus and postcentral gyrus was observed when starting to move; turning activated an extended area from occipital cortex dorsally to superior parietal cortex and laterally in the right hemisphere to the posterior middle temporal gyrus; reversing activation was prominent in the lateral precentral gyrus and anterior insula/ventrolateral prefrontal cortex; stopping involved a more restricted activation and focused more on the anterior part of the pre-SMA. The neural activations in other higher level of driving components (e.g. strategy level or tactical level) were also reported in Spiers and Maguire's study: monitoring actions from other drivers showed extensive activation in the precuneus and superior parietal cortices; traffic rule related thoughts were associated with significant activation of the right lateral PFC.

To sum up, previous studies have revealed activations in a network of brain regions (including the parieto-occipital cortices, cerebellum and cortical regions) associated with perception and motor control. Activation in these regions was generally attributed to increased demands on vision, motor skills and visuomotor integration. In addition, activity in frontal, parietal, occipital and thalamic regions was found to correlate with average driving speed (Calhoun et al., 2002; Horikawa et al., 2005). Horikawa et al. (2005) also found that the number of crashes was negatively correlated with activity in the posterior cingulate, while others observed that the ability to maintain a safe driving distance was negatively correlated with activity in the anterior cingulated (Uchiyama et al., 2003).

2.6.3 State-of-the-art driver workload assessment using psychophysiological signals

The assessment of driver mental workload using psychophysiological signal has been investigated for decades. Most studies focus on the offline quantitative analysis of psychophysiology.

Brookhuis and his colleagues were among the first to investigate the ECG response in the context of driving. Brookhuis et al. (1991) showed that dual task on-road, driving a verbal serial addition task presented via cell phone, elicited significant increases in HR and decreases in HRV. Later, Brookhuis and De Waard (1993) revealed that IBI decreased and HR increased in ring-road driving situation compared with motorway driving situations. In a more recent study, Mehler et al. (2009) revealed that increased cognitive workload leads to increased HR and reinforced that HR can be a sensitive measurement of cognitive workload in the driving environment (for further support, see also Lenneman et al., 2005, 2009; Healey & Picard, 2005). More recently, HRVs were used to evaluate the effectiveness of the new driver assistant system in the context of mental workload. For instance, Brookhuis et al. (2009) reported a simulator study which demonstrated that the IBI was higher with the Congestion Assistant than without the system in situations of traffic jam. Other researchers attempted to establish an on-line workload detector using ECG parameters. Lin & Cai (2009) proposed a k-means cluster-based methodology for real-time workload detection using various ECG parameters. Their validation of the driving application indicated that such approach has good agreement with RSME.

Since the 1980s, the EEG has also been introduced to investigate driver workload when the task load is induced by various factors including the combination of in-vehicle secondary task (Raabe et al., 2005; Lei et al., 2009 b, c; Kohlmorgen et

al., 2007; Hagemann, 2008), the road conditions (Brookhuis and De Waard, 1993), and environmental factors (Baldwin & Coyne, 2005; Baldwin et al., 2004). A majority of these studies investigated driver workload using either ERP components or spectrum decompositions in an offline analysis.

The P300 has been introduced to evaluate the driver mental workload. Baldwin et al. (2004) investigated the changes in P300 evoked by a visual color-discrimination oddball task under conditions of normal and reduced visibility induced by fog density during simulated driving. P300 amplitude was reduced when participants drove in fog compared to driving in normal visibility. However, P300 was not sensitive to changes in traffic density, though the behavioural measures were sensitive to this manipulation of driving difficulty. Raabe et al. (2005) revealed a decline in the amplitude of P300 evoked by a secondary oddball task when the primary driving task difficulty was increased. However, Baldwin et al. (2004) pointed out that neither neural nor behavioural measures alone would be sufficient for assessing cognitive workload during different driving scenarios. And they speculated that an ERP based method could be used to trigger adaptive aiding for the driver by engaging or disengaging the in-vehicle systems such as cellular phones or entertainment devices.

Although the EEG spectrum components are commonly used to assess the mental workload, the investigation of the modulations of such parameters with workload in-driving context does not frequently appear in the literature. Brookhuis and De Waard (1993) used an energy parameter ([theta+alpha]/beta) to index subjects' activation during on-the-road driving experiments. This parameter was observed to increase with an increased task load induced by different road conditions. Recently, Hagemann (2008) systematically investigated changes in alpha power with task load in driving contexts. A significant attenuation of alpha was found when participants performed simultaneously LCT and a secondary word or tone

detection tasks, compared with conditions of only word or tone detection.

There are also a large number of studies using machine learning methods to classify psychophysiological signals for driver mental states assessment. Such studies also combined other psychophysiological variables such as pupil diameter, skin conduction, EMG, and reparation rate (Healey and Picard, 2005; Putze et al., 2010). Healey and Picard (2005) developed a classifier to monitor the stress levels in daily life car driving tasks. They collected multiple psychophysiological data including electromyography, electrocardiography and skin conductance from the drivers. LDA was used to distinguish three levels of driver stress during rest, highway, and city driving conditions. An accuracy of over 97% across multiple drivers and driving days was reported. More recently, Putze et al. (2010) reported a study on the development and evaluation of a recognizer for different levels of cognitive workload in the car. They collected multiple psychophysiological information including skin conductance, pulse, respiration, and EEG during an experiment in a driving simulator in which drivers performed a primary driving task and several secondary tasks (visual and cognitive) of varying difficulties. The SVM was used to classify driver cognitive load based on one-minute windows of data. Such a system showed reliable detection of a stable level of the driver's workload.

However, the fact remains that only few studies integrate the EEG-based interface in the driver-vehicle-system to initiate adaptive aiding for the driver. As far as we know, only Kohlmorgen and his colleagues (2007) conducted such a study, in which an EEG-based system was used to detect driver mental workload under real traffic conditions. They classified driver workload into high and low conditions under various EEG spectrum features. The result was used immediately to modulate the workload induced by the influx of information from the car's electronic systems. They showed that a system as such was beneficial for

improving drivers' overall task performance.

2.7 Limitations of the current EEG-workload research

Although a decent proportion of the literature is concerned with the observation of EEG activity for workload assessment, there is still a lot of work to be done for the application of such technology into the complex human-machine systems. Several issues concerning this topic ought to be addressed.

First of all, the majority of these results were based on the average of a relatively long period of measurement (e.g. 3 or 5 minutes or longer) or directly used the short term EEG features to classify the workload levels (Wilson & Russell, 2003a, b; Lee & Tan, 2006; Grimes et al., 2008). These studies failed to provide a detailed observation of the short term (e.g. 5 seconds) changes to these EEG parameters. Since the integration of this technology into a human-machine-system often requires real-time processing which means a fast and short time window for the estimation, the observation and analysis of the short term changes in EEG parameters is deemed necessary.

Secondly, most of the quantitative analyses of previous studies investigated the modulation of the EEG spectrum based on some simple tasks in which workload levels were manipulated in a single dimension (e.g. working memory load). And there is still a lack of evidence as to how these EEG parameters quantitatively change in the complex situation wherein workload levels might be a result of various combined factors. Indeed, some earlier studies employed relatively complex tasks such as Multi-Attribute Task Battery (Smith et al., 2001; Wilson & Russell, 2003b) or some in-flight tasks (Wilson & Russell, 2001). But most of these studies employed the machine learning methods to classify the workload levels using multiple EEG variables with the result that they failed to provide a

clear comparison of the similarities and differences in the modulation of EEG spectrum induced by different dimensions of workload. Questions such as how these sub-resources of workload contribute to the general workload and what effect would be had after these sub-resources are combined, remain to be sorted out. Driving is a complex task and the driver mental workload could be affected by various factors. This specific feature of driving task challenges the establishment of a universal driver workload detector using the EEG signals.

Thirdly, many studies attempted to combine the EEG parameters for a reliable index of neural activity such as using the ratio of the different band powers (Brookhuis & De Waard, 1993; Pope et al., 1995; Prinzel et al., 2000, 2001). However, as far as we know, such combined values are largely task dependent and involve plenty of individual variations. Different participants might demonstrate different values. It is still hard to draw the 'redline' for mental overload using a general predefined threshold. At present, there is still a lack of methods which could project the EEG parameters into unified workload indices, for instance, scales ranging from 0-1.

Finally, a large body of studies used the pattern recognition method to classify the workload into several levels. Another approach for real-time assessment of mental workload is to use the classic machine learning methods, which are typically used in the Brain-Computer-Interface (BCI) research (Wilson and Russell, 2003b; Kohlmorgen et al., 2007). Such approach includes various machine learning algorithms, e.g. linear discriminate analysis (LDA, Kohlmorgen et al., 2007), Neural Network (NN, Wilson and Russell, 2003b), etc. Using these methods, workload states are classified into several levels (e.g. low, moderate and high) using various EEG parameters in either simple, single-task or complex tasks. However, questions such as "are we really classifying the mental workload", or "does classification involve some other task components rather than workload"

are often raised. To our knowledge, using a machine learning method to classify workload seems problematic. Firstly, the traditional machine learning method trains the dataset recorded in a specific task load condition with a unique training label. However, such training method is problematic, since the operator's functional states dynamically change in a broad range during the performance of the task and it is impossible to keep the workload at a constant level. Secondly, as was usually done in previous research, the classification was based on a large feature vector consisting of different EEG parameters selected from different recording sites on the scalp. However, this kind of classification is sensitive to the selection of features and might easily involve cognitive components rather than workload. This effect is especially prominent when the task involves multiple subtasks which occupy different neural resources. The third prominent drawback of a BCI-based approach is that these classifiers are task dependent. The task-specific classifier is not applicable to other tasks, which dimmed the possibility of establishing a general and reliable workload detection matrix.

Therefore, earlier studies indeed demonstrate that using EEG to index the driver's mental states is feasible. However, to integrate such psychophysiological interfaces in the driver-vehicle system for adaptive aiding design, both theoretical and empirical evidence is definitely necessary.

2.8 Summary of the theoretical background

This chapter starts with the theoretical background of the adaptive task allocation (ATA), which focuses primarily on the relation between mental workload and task performance during the human-machine interaction. For understanding the fundamental theory of ATA, the definitions of workload within the literature as well as in the ISO scope are discussed. The mental workload in this dissertation is referred to as a combination of *mental stress* and *mental strain*. The necessity of

ATA is addressed by discussing the relationships of task demand, workload, and task performance within the literature, which concluded that a moderate workload level is the optimal stage for task performance. Following this, a new model, *demand-workload-matched adaptive task allocation (DWM-ATA)*, is originally proposed by the author as an extension of the Parasuraman's workload-matched ATA model. The biggest merit of the new proposed model is that it provides an ATA solution for cases in which workload is dissociated with task demand[20].

Within the scientific literature, numerous techniques including subjective reporting, performance measures, and psychophysiology measures, have been widely used. But to date, it still can be seen that there is no agree-upon technique that would allow a valid and reliable measurement of mental workload. Compared to the subjective methods and performance measures, psychophysiological measures show clear merits in the measures' directness, unobtrusiveness, and high time resolution. EEG technique is the main concern of this dissertation. In section 2.4, we particularly focus on a discussion of the principles for using the EEG technique in mental workload assessment. Various parameters in the EEG signal (e.g. ERP components, EEG frequency bands) have demonstrated an extra sensitivity to mental workload representaiton. Following this, in section 2.5, earlier studies investigating the psychophysiology-driven adaptive aiding were reviewed. Such studies show that integrating the psychophysiological interface into the human-machine-system is beneficial for improving the operator's overall performance.

Driving is a complex skill. In section 2.6, the driving task analysis as well as the neural correlates of driving provides the underlying psychological and physiological theory behind the complex task of driving. Driver mental workload

[20] The dissociation of workload and task demand has been addressed in previous literature (Parasuraman & Hancock, 2001). But there is actually no concern for dissociation of cases in the adaptive task allocation within the literature.

is a result of various internal and external factors and the multi-factorial property makes it particularly difficult to establish a general workload detector allowing for driver workload assessment across tasks and individuals. However, the state-of-the-art literature shows that psychophysiological markers, particularly the EEG signal, hold potentially an adequate ability for indexing driver mental workload.

Nevertheless, some limitations of current EEG-workload are addressed in section 2.7. A lack of reliable workload quantification model using EEG signal and the inadequate evidence for supporting the feasibility of adaptive task allocation in a driving context are specifically addressed. These two points are the principal foci of the whole dissertation.

Chapter 3. Representation of driver of workload in EEG: ERP or Band Powers?

3.1 Motivation

As introduced in Chapter 2, both ERP and EEG spectrum components could be used to index the mental workload states. For instance, the amplitude and latency of P300 may differ when workload varies (Polich, 1990; Kramer & Parasuraman, 2007); EEG theta activity may increase and alpha may decrease as task load increases (Gevins and Smith, 2000; Smith et al., 2001; Wilson, 2002b). In the driving context, there are also studies investigating the modulations of ERPs (Baldwin et al., 2004; Baldwin & Coyne, 2005; Brookhuis & De Waard, 1993) and EEG spectrum parameters with workload (Hagemann, 2008). The results from these studies generally demonstrate that EEG technology could effectively assess the human operator's mental workload states.

Exploring the sensitivity and robustness of the ERPs and EEG band powers is meaningful for establishing a reliable workload detector. However, the earlier studies showed that EEG parameters may demonstrate context dependent sensitivity as the workload is varied. As far as we know, no study has systematically compared the robustness of ERPs and band powers for workload assessment.

Two objectives are envisaged in this chapter. Firstly, we aim to investigate if the previous findings can be reproduced in the context of driving, mainly in the Lane Change Task (Mattes, 2003; ISO, 2010), which is a standardized stimulated driving task. Secondly, the comparison of the robustness of these two methods (ERPs and Band Powers) for workload representation was also a specific interest

of ours. According to these objectives, both single task and dual task paradigms were employed to provide an insight into the comparison. In the single task paradigm, the LCT was used and the driving task load was modulated by the speed settings, which have proven to be an important contributor to driver workload (see section 2.6.1). In the dual task condition, a secondary task, the Paced Auditory Serial Addition task (PASAT; Gronwall, 1977), was used to investigate the variation in the EEG parameters induced by the task combination. Additionally, a classification was also performed using ERP and band power features to provide an insight into the robustness of these two types of information for instantaneous workload assessment. The following section describes the details of the methods, the data analysis and results, to be rounded off with a multifaceted discussion.

3.2 Introduction of the tasks

3.2.1 Lane Change Task

The Lane Change Task (LCT) was initiated by the project ADAM (Advanced Driver Attention Metrics) as an easy-to-implement, low-cost, and standardized methodology for the evaluation of the attention associated with performing in-vehicle tasks while driving (Mattes, 2003). Recently, this method has become an ISO standard[21] for in-vehicle secondary task demand assessment (ISO, 2010). The LCT can be implemented in various experimental contexts (laboratory, driving simulator, or production vehicle) for all types of interactions with in-vehicle information.

In LCT, the driving road consists of three lanes, and participants are asked to repeatedly perform lane changes among the three lanes when prompted by road

[21] ISO 26022 in road vehicles-ergonomics aspects of transport information and control system.

signs without other vehicles or pedestrians present (see Figure 3.1). The LCT consists of 10 tracks and each track is approximately 3 km (around 3 minutes if driving at 60 km/h) including 18 randomly sequenced road signs (indicating to which lane the participant should change) plus a start sign. The participant could control the driving speed using a simulated gas pedal and stop the vehicle with a brake pedal.

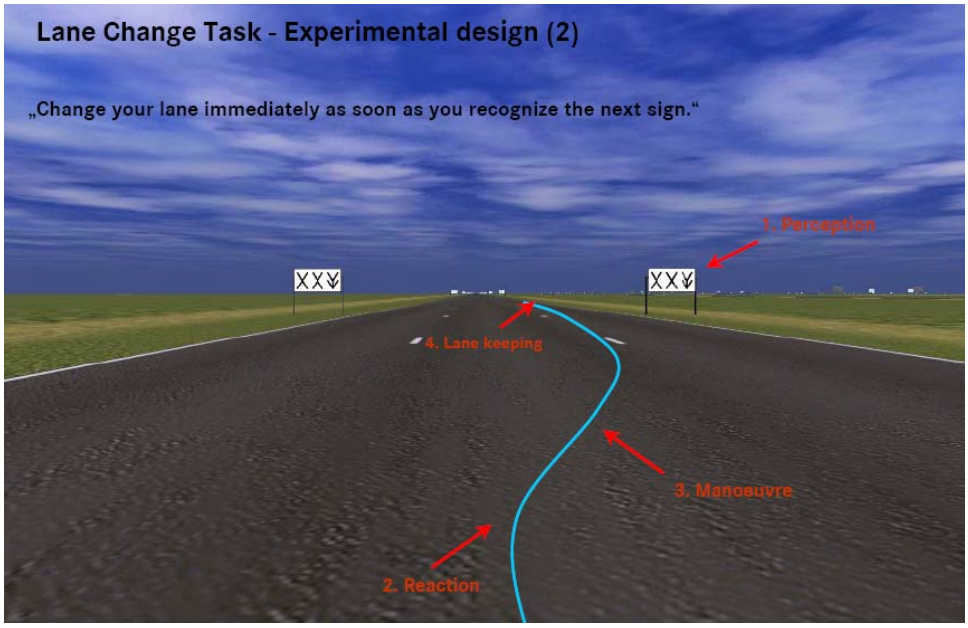


Figure 3. 1 Lane Change Task: lane changes are promoted by the road signs

In LCT, the participant is instructed to repeatedly perform lane changes as soon as possible when s/he recognizes the road signs. Basically, the LCT requires skills in control level, e.g. lane keeping, steering, speeding, and needs also some skills in manoeuvring level, e.g choice of lane according to road signs[22]. The driving data (vehicle positions, driving speed, steering angle, etc) can be recorded in a log file for the post-driving analysis. The quality of these lane changes can be evaluated by the difference (mainly based on Mean Deviation) between a normative model and the driver's behavioural data (see Figure 3.2), which is influenced by the driver's performance in detecting and responding to the road signs as well as their

[22] See, Michon's hierarchical model in page 66

lateral control maintenance. The lane change performance analysis enables the evaluation of the extent of distraction due to the secondary task.

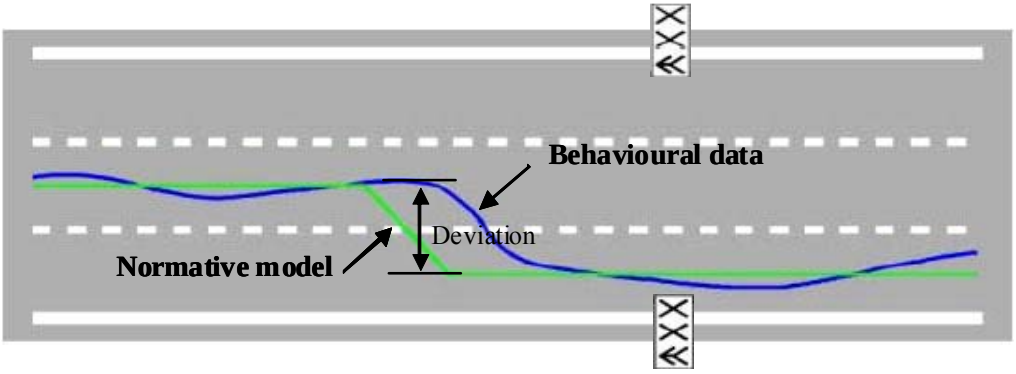


Figure 3. 2 Deviation between a normative model and the driver's behavioural data for performance assessment in LCT

3.2.2 Paced Auditory Serial Addition Task (PASAT)

The PASAT is a method for triggering cognitive function that assesses the auditory information processing speed and flexibility, as well as calculation ability. It was developed by Gronwell in 1977. The PASAT is presented using audio cassette tape or compact disk to ensure standardization[23] in the rate of stimulus presentation. Single digits (from 1 to 9) are presented at certain pace (e.g. every 2.4 seconds) and the participant has to add each new digit to the one immediately prior to it. Shorter inter-stimulus intervals may increase the difficulty of the task. The digit is randomly arranged to minimize possible familiarity with the stimulus items when the PASAT is repeated over more than one occasion.

3.3 Pre-study: Manipulating workload in Lane Change Task

The pre-study aimed at determining whether speed settings lead to changes in the driver's workload. As introduced in the first chapter, the speed is an important

[23] The original version presented the numbers every 2.4 seconds with 0.4 decrements for subsequent trials. But in this study, the PASAT was combined with a driving task. The standard paces are too fast for the participant. Therefore, we used two slower paces, every 5 and 3 seconds.

contributor to the driver's workload (Cnossen et al., 2000; Musa and Navin, 2005). However, evidence to support this theory is essential especially in a simulated driving environment.

18 participants whose age ranged from 23 to 37 (M=30, S=4.35) participated in the pre-study. The experiment setup involved a PC, a projector, a comfortable chair, and the Logitech driving interfaces including steering wheel, gas pedal and brake pedal (see Figure 3.3). The driving scene was projected onto a wall around one meter in front of the driver. The subjective load questionnaire NASA-TLX was used in the pre-study to investigate the speed-related task load.

Figure 3. 3 Experiment setup for the pre-study

Firstly, participants were asked to read an instruction about this experiment. Then, they had a 10-minute practice which included all three task conditions, driving at three speed levels 60km/h, 90km/h and 120 km/h, representing low, moderate and high task load respectively. After the practice session, participants were instructed to evaluate the NASA-TLX weight questionnaires. Later, they were required to perform the LCT in three 20-minute blocks, including one trial per each three task

load conditions (low, moderate and high). The trials for speeds were randomly organized. After each trial, participants reported their subjective workload by answering the NASA-TLX questionnaires.

The data analysis involved the driving performance (mainly the mean deviation between the predefined normal lane and driver's behavioural data) and the NASA-TLX subjective load. Figure 3.4 shows the results of the pre-study. Although the scores of workload and mean deviation differed significantly among individuals, there were significant speed associated differences in either subjective task load or in the performance. The subjective load significantly increased with the speed level ($F(2, 51) = 27.35$, $p<0.001$). Pairwise comparison indicated that the differences between each pair of these three speed conditions were significant ($p<0.001$). The increase in mean deviation ($F(2, 51) = 84.96$, $p<0.001$) indicated that there was a decline in performance with high speed. Pairwise comparison indicated that the differences between each pair of these three speed conditions were significant ($p<0.001$). These results demonstrated that the manipulation of task load by changing the speed settings was successful.

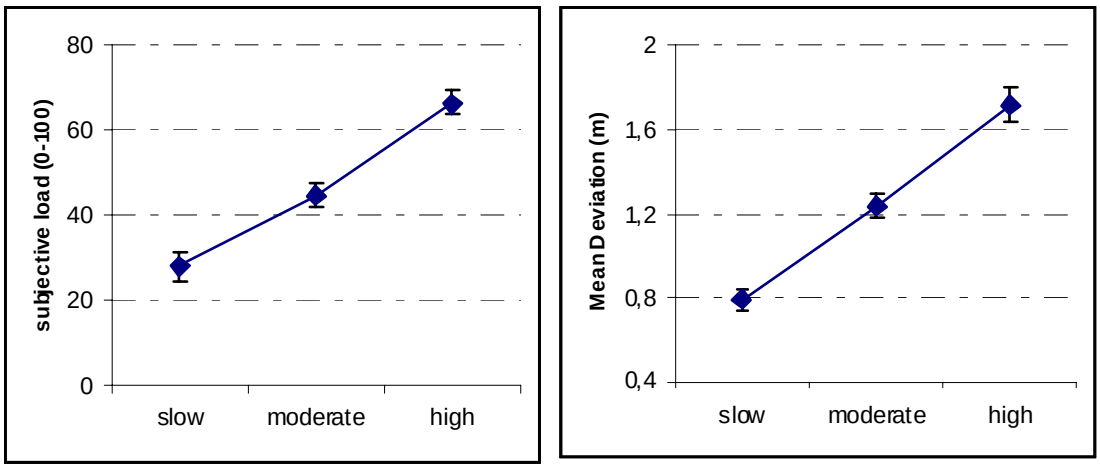


Figure 3. 4 Correlation of subjective workload (left) and LCT performance (right) to speed levels in the pre-study

3.4 Assessment of driver's mental workload with EEG

3.4.1 Participants

Overall, 30 participants between the ages of 20-34 were assessed (mean=26.1, standard deviation=2.79), comprising 18 males and 12 females. However, the data from 6 participants were incomplete because of system failure or simulator sickness. Additionally, 2 participants were not able to complete the PASAT. Therefore, 24 participants produced validated experimental data for the single task condition, and 22 validated data samples recorded for the dual task condition. All individuals were reported to be free of illness and medication. Participants also received a cash payment for their participation.

3.4.2 Experiment apparatus

The experiment setup consisted of a driving box (with steering wheel, brake, gas pedal, and an adjustable driving seat, see Figure 3.5), a PC (with LCT software), a projector, a loud speaker, and the EEG system. The LCT program ran in the PC, and was connected to the driving input devices (steering, gas pedal, etc.). Using the projector, the driving scene was projected onto the wall about a metre in front of the driving box. The auditory digits for PASAT were presented to participants using a loud speaker. The PASAT required participants to respond verbally and the results were recorded manually by the experimenter. The LCT program automatically generated a log file to save various driving data including the vehicle position, speed, steering angle, etc.

Figure 3. 5 LCT equipped with a driving box

Brain activity was recorded with 32 Ag/AgCl impedance-optimized electrodes (ActiCap, Brain Products, Germany), referenced at the nasion, sampled at 1000 Hz and wide-band filtered (0.5-70 Hz), and placed according to the international 10-20 system. Electromyogram (EMG) was recorded from both forearms using two bipolar electrodes. The horizontal and vertical eye-movement was recorded using the Electrooculogram (EOG). Both EMG and EOG were sampled at 1000 Hz, and were recorded together with the EEG dataset as the extended channels using the Brain Vision Recorder (Brain Products).

The signals from LCT and EEG were synchronized using a self-developed tool based on Labview (National Instruments, USA). This tool could read the LCT log data in real-time through TCP/IP and automatically detect the LCT road signs. Markers linking with these road signs were sent to the Brain Vision Record for synchronization with the EEG signal.

3.4.3 Experiment procedure

The experiment involved four blocks. The first block was the primary driving task. Participants were requested to perform the LCT under three different speeds 60km/h, 80km/h, 100km/h, which represented three task load levels respectively (low, moderate and high). The sequence of three speed conditions was randomly ordered. (In the formal study, we transferred the experiment from the original LCT device employed in the pre-study to a driving simulator and the steering wheel was more sensitive to participants' input. Thus, we slightly changed the speed settings). The second block was PASAT consisting of two randomly ordered pace conditions: slow and fast (the numbers were presented at every 5 and 3 seconds, namely p5 and p3). Participants were requested to calculate the numbers and report the results. The third block was a combination of the primary and secondary tasks. Participants were requested to do the calculation at two paces while performing the LCT with a fixed speed of 80 km/h (driving+p5 and driving+p3). Participants were requested to do the calculating while performing the LCT. However, they were informed that the primary task was more important. The results of the PASAT were manually recorded.

3.4.4 Data analysis

Various variables, including LCT mean deviation, PASAT accuracy, and EEG parameters, were analyzed. LCT performance was evaluated in terms of the mean deviation between the driver's behavioural path and a predefined normal lane change path. The normal lane change path was defined as follows: view distance: 40 meters; distance for two lane changes: 20 meters; distance for one lane change: 10 meters. This normal path was applied to all subjects to ensure the same standard for the performance evaluation.

Both ERPs and EEG band powers were extracted to investigate the workload associated variations. The statistical analysis method (i.e. Analysis of Variance or ANOVA) was used to verify the significance of the differences in these parameters. A Post Hoc Test was then used to offer a detailed pairwise comparison when the differences among different task conditions showed any general significance. For further comparison of the robustness of ERPs and band powers for workload representation, a machine learning method, i.e. the Adaptive Boosting (Adaboost), was used to classify the workload using features extracted from ERPs and band powers.

The data analysis was performed using EEGLAB 6.03, a freely available open source toolbox running under Matlab (Mathworks, USA). A detailed description of EEGLAB is provided by Delorme and Makeig (2004). For pre-processing, the data was down-sampled to 500 Hz to save computation time, and was then digitally filtered using band pass filter (pass band 0.5 to 40 Hz) to minimize drifts and line noise. Later, the data was processed differently for extracting the ERPs and band powers.

Event-related potentials (ERPs)

After pre-processing, the data was averaged and referenced to avoid the influence of any arbitrary chosen local reference (Nunez, 1981). Then, data epochs were extracted from 2000ms before stimulus, the command for lane change direction, until 2000ms after the stimulus onset, and the average in time range [-2000ms, -1000] as the baseline was removed from every epoch.

EEG recordings involve various artefacts, such as eye movements, muscle noises, cardiac signals and line noises. For artefacts removal, firstly, we visually inspected the data before the epochs. Abnormal contaminations or outliers were excluded.

Later, the Independent Components Analysis (ICA) was used to improve the data quality and to identify the ERP components. ICA decomposes EEG data into temporally independent and spatially fixed components, which account for artefacts, stimulus, response-locked events, and spontaneous EEG activities. Recently, this has been considered a powerful tool for EEG components identification and artefacts removal (Makeig et al., 1996; Delorme & Makeig, 2004).

After the calculation of independent components (ICs), we followed the following procedures to remove artefacts. (1) The correlations coefficient of ICs and EOG, ICs and EMG, were calculated and the highest correlated ICs were removed. (2) The correlations of the single-trial ERP and the mean ERP for each electrode site, condition and subject were calculated and the number of trials which correlated highly to the mean ERP was recorded (correlation coefficient was higher than the threshold 0.2). The ICs, which had less than 20% trials with high correlation to the mean ERP, were removed. (3) The datasets were visually checked again, and the ICs induced by the artefacts were excluded.

After artefacts correction, ERPs were averaged over conditions for each participant. To analyze the differences in the amplitude of ERP components, the average amplitude over time range 200-300 ms in the frontal area and the average amplitude over time range 450-650 ms in the parietal-occipital were extracted for each condition and participant. To analyze the differences in the latency of ERP components, the latency of the peak point over time range 200-300 ms in the frontal area and the latency of the peak point over time range 450-650 ms in the parietal-occipital area were extracted for each condition and participant as well.

Frequency Band Powers

After pre-processing, the data from the start point to the end (the last stimulus of this track) for each track in LCT was extracted (approximately 2 minutes), i.e. the data during the LCT curve was excluded. Then, ICA was used for artefacts removal and ICs induced by ocular artefacts were removed. An FFT analysis was then used to estimate the power spectrum density for each trial (driving track). The relative power density spectrum (percentage to total power of frequency range 4-30 Hz) was calculated for each trial. Frequency bands, theta, alpha and beta were extracted by accumulating the power of frequency bands to 4-8Hz, 8-12Hz, 13-30 Hz respectively. This way, around 40 data points for each EEG spectrum parameter were extracted for each participant including all task conditions. Finally, these data points were averaged over task conditions for each participant.

Feature extraction and classification

The ERPs feature extraction was followed by the ERPs quantitative analysis in section 3.4.4. After artefacts correction, a small laplacian filter (Hjorth, 1975) was used to re-reference an electrode to the mean of its four nearest neighbouring electrodes. The edge electrodes and some abnormal electrodes were excluded and still 18-21 electrodes remained. ERPs were averaged every 50ms from [-1000ms, 2000ms], which yielded 60 features. Therefore, a vector of more than 1000 features for each trial was extracted.

The band power feature extraction was different from the quantitative analysis of band power. Firstly, the data was segmented into 2s epochs with a 50% overlap. This way, an average of 300 epochs was obtained for each electrode, condition and subject. Wavelet transformation was used to decompose the EEG signal into

different frequency bands. Daubenchies D6 wavelet function, which has been reported to be similar to the sum of the neuron action potentials (Daubechies, 1992), was used to perform a 6-level analysis, and the output coefficients were reconstructed. In this way, five EEG frequency bands, delta (1-4Hz), theta (4-8Hz), alpha (8-16Hz), beta (16-32Hz), and gamma (32-64Hz), were extracted. After wavelet decomposition, the EEG band power was calculated as the sum of the square of data points for each frequency band and normalized as the ratio of the power of frequency band to the total power of these five frequency bands. As such, these five normalized bands formed a vector of more than 100 features from the electrodes which were also used for ERP feature extraction.

For classification, a powerful classifier (Adaboost) was used. AdaBoost, i.e. Adaptive Boosting, is a machine-learning algorithm originally proposed by Freund and Shapire (1997). It offers numerous advantages. It is fast, simple and easy to program. It has no parameters to tune, requires no prior knowledge and can be flexibly combined with any method. Recently, Adaboost has been introduced for the classification of EEG signals (Pei et al., 2005). In the present study, we used decision trees based on a logistic regression version of Adaboost (Friedman et al., 2000; Collins et al., 2002). Decision trees make good weak learners because they provide automatic feature selection and limited modelling of the joint data statistics. Each decision tree provides a partitioning of the data and outputs a confidence-weighted decision that is the class-condition log-likelihood ratio for the current weighted distribution. In order to guarantee the reliability of the results, a 10-fold cross-validation was used.

3.5 Results

3.5.1 Task performance

There were significant increases in the mean deviation when the driving speed increases ($F(2,24)=111.2$, $p<0.001$). Post-hoc tests indicated that differences between each pair of three speed conditions were significant ($p<0.001$). The mean deviation also exhibited significant differences among three dual task conditions ($F(2,22)=3.9$, $p<0.05$) (see Figure 3.6). Pairwise comparison indicated that only the difference between pairing of driving only and driving + p3 condition was significant ($p<0.05$).

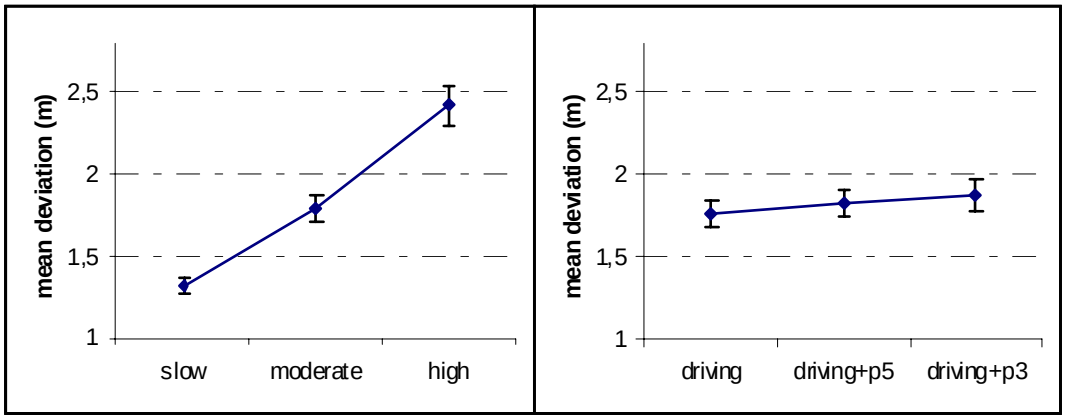


Figure 3. 6 Comparison of mean deviation in both single task (left) and dual task (right) conditions

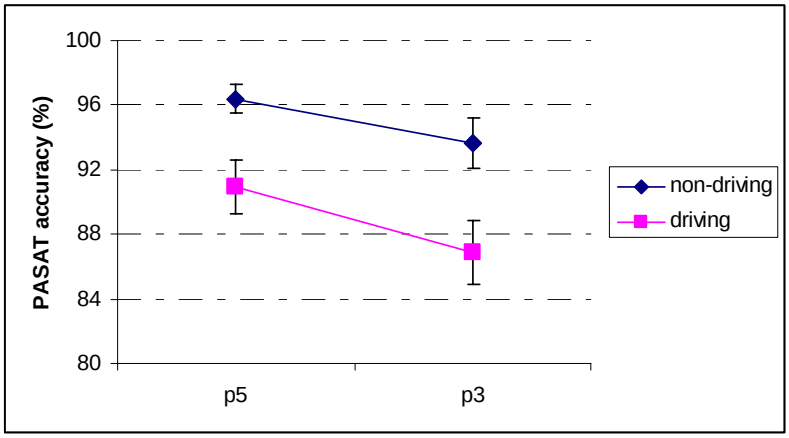


Figure 3. 7 Comparison of the PASAT performance in single and dual task conditions

The PASAT accuracy demonstrated that the participants' performance in PASAT decreased either when the PASAT load was increased from p5 to p3 ($F(1,22)=10.3$, $p<0.05$) or when the PASAT was combined with the driving task ($F(1,24)=26.5$, $p<0.001$) (see Figure 3.7). But there was no significant interaction effect when the driving task and additional tasks were combined ($F(1,22)=0.75$, $p=0.39$).

3.5.2 ERP in LCT

Independent Components Analysis

ICA decomposes the EEG into temporally independent and spatially locked components. These components could either be the result of task events/stimulus or on account of the artefacts. Figure 3.8 shows some typical IC examples evoked by LCT.

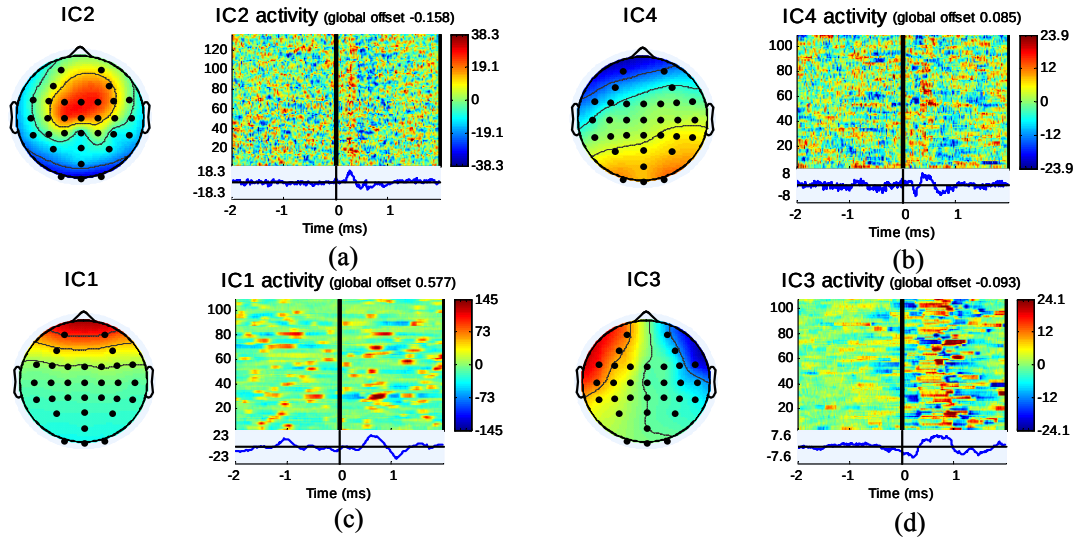


Figure 3. 8 Independent components which demonstrate the typical components obtained in the present study (left: topography of the IC and right: the corresponding averaged signal).

As shown in Figure 3.8, some of these components showed a high coherence across all trials and the peaks seemed to be time locked, e.g. (a) and (b), and were likely evoked by the involved events. Empirically, we could conclude that Figure 3.8 (a) shows a P2 or an early P3a-liked positive peak, since it happened at around 250mm and was located in the frontal-central area, while Figure 3.8(b) demonstrates a P3b-liked positive peak located in the posterior area with a latency of 500ms. Both showed a high coherence across all trials. However, some of these presented a low coherence across all trials and the peaks were randomly located at the time course, e.g. (c) and (d), and may also account for artefacts. Figure 3.8 (c) shows the eye blink artefact and Figure 3.8 (d) shows the eye horizontal movement artefact, which is characterized by a location in the frontal area of the head and a strong power activity in the low frequency.

ERP evoked by LCT

Artefacts-corrected data showed a prominent positive peak around 250ms (P250) after the stimulus onset at the frontal-central area of the brain. This peak had its maximum ($21.6\,\mu V$) at FCz and was still active at the parietal area but with a smaller amplitude and earlier latency (200ms) (see Figure 3.9). Another prominent positive peak occurred at the latency of 550ms (P550) at the parietal and occipital area (see Figure 3.9) and it had a maximum ($20.2\,\mu V$) at POz. The steering angle recorded in LCT indicated that both of these two components happened before the participants took the action of a lane change, which happened around 800ms after the stimulus onset. Additionally, it seemed that other components were also involved, such as the contingent negative variation (CNV), an increasing negative shift associated with an expected "go-signal", and N2, a negative maximum around 250 ms at the parietal and occipital area.

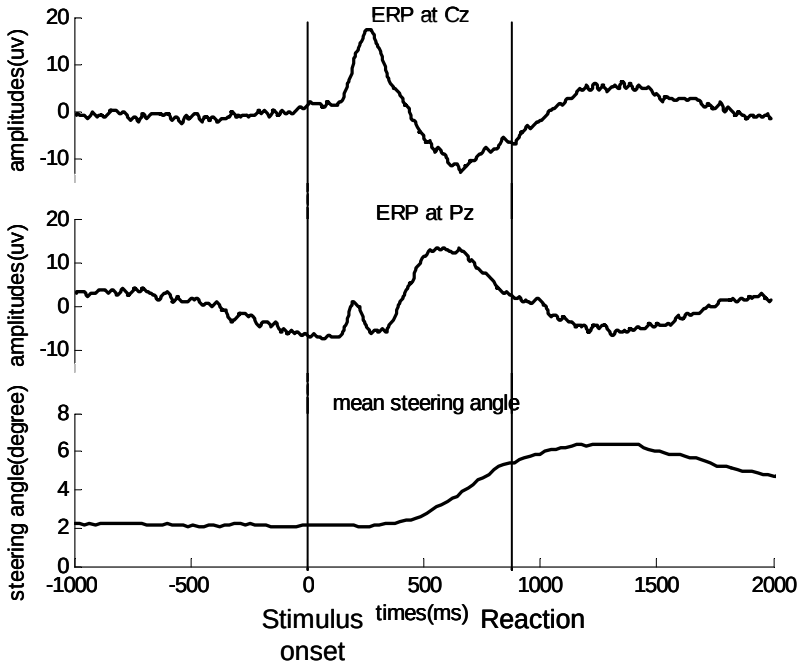


Figure 3. 9 Artefacts-corrected ERP and steering angle in LCT

3.5.3 ERP and workload

Single task: Speed related changes

Figure 3.10 shows the spatial patterns of ERPs at different latencies in the single task condition. As shown in Figure 3.10, generally, the change of speed elicited no obvious changes in the early components (around 200-300ms at the frontal-central area). However, the late component (around 500-600ms at the parietal-occipital area) showed slight decreases when the speed is increased.

Figure 3.11 shows the artefacts-corrected ERP under three speed conditions at different electrode sites, which demonstrated the most prominent components seen from the spatial pattern. For P250, statistic analysis demonstrated no differences in its amplitudes ($F(2,24)=0.68$, $p=0.509$ at FCz; $F(2,24)=1.36$, $p=0.264$ at Cz).

However, the differences in amplitudes of P550ms under three speed conditions were significant at the parietal-occipital area with maximum differences at P4: ($F(2,24)=12.83$, $p<0.001$ at P4; $F(2,24)=6.66$, $p<0.01$ at POz; $F(2,24)=6.25$, $p<0.01$ at Oz).

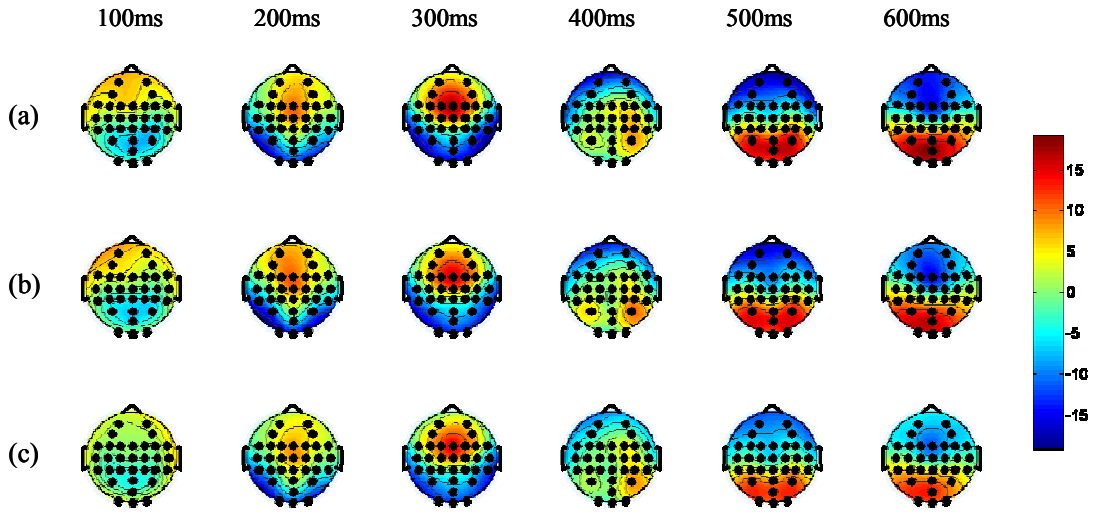


Figure 3. 10 Spatial pattern of the ERPs in single task condition: (a) slow driving (60km/h), (b) moderate driving (80km/h), and (c) fast driving (100km/h). These results were based on the average of 24 participants.

Post hoc tests indicated that there were significant differences between the pairing of slow driving and moderate driving, slow driving and fast driving, but no difference was indicated between conditions of moderate driving and fast driving at the location of POz and Oz. But at the location of P4, the differences between the three pairs of speed conditions were significant. Furthermore, statistic analysis showed no differences in the latency of these two components: ($F(2,24)=0.14$, $p=0.869$ at Cz; $F(2,24)=0.08$, $p=0.922$ at Pz; $F(2,24)=0.009$, $p=0.916$ at Oz).

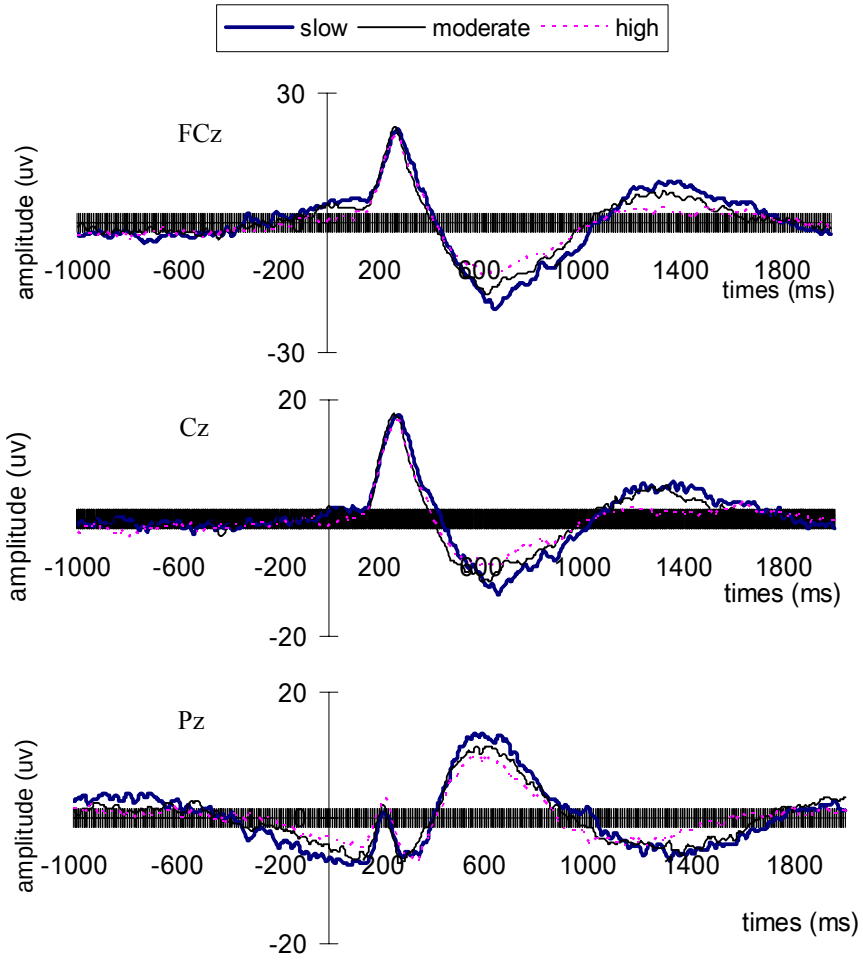


Figure 3. 11 Artefacts-corrected ERPs at different electrodes in single task conditions (N=24)

Dual tasks: Task combination related changes

Figure 3.12 shows the spatial patterns of ERPs at different latencies in the dual task condition. As shown in Figure 3.12, generally, the task combination elicited no obvious change in the early components (around 200-300ms at the frontal-central area). However, the late component P500 or P600 showed prominent changes at the parietal area.

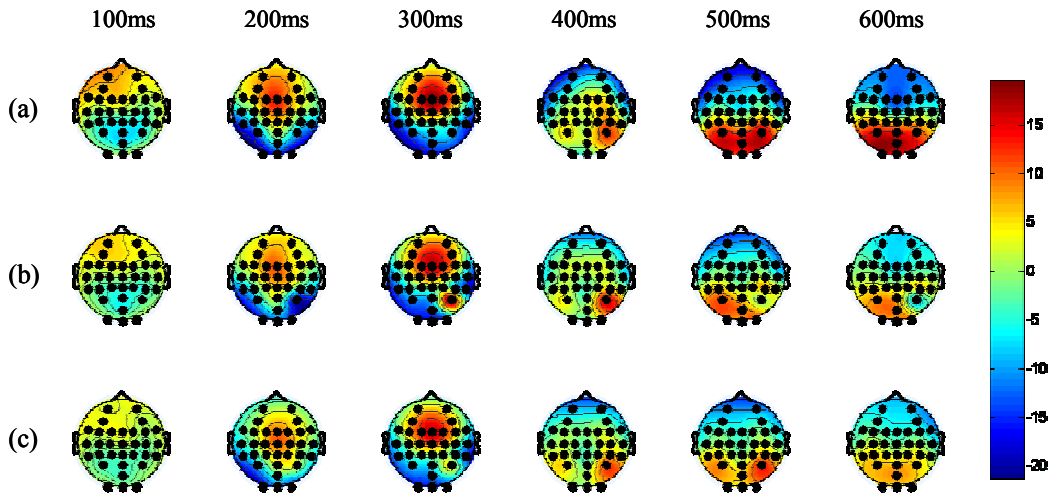


Figure 3. 12 Spatial pattern of the ERPs in dual task conditions: (a) driving only at 80 km/h, (b) driving at 80 km/h plus p5, and (c) driving at 80 km/h plus p3. These results were based on the average of 22 participants.

Figure 3.13 shows the comparison of the ERPs under three task conditions at different locations. Statistical analysis ANOVA revealed that there were no obvious differences in amplitude of P2 among the three conditions at the frontal-central area (at FCz $F(2,22)=1.99$, $p=0.14$; at Cz, $F(2,22)=2.96$, $p=0.06$), whereas obvious changes in amplitude of P550 were obtained at the central-parietal-occipital area and the difference reached its maxima at Pz (at CPz, $F(2,22)=3.69$, $p<0.05$; at Pz, $F(2,22)=8.75$, $p<0.001$; at POz, $F(2,22)=7.13$, $p<0.01$; at Oz, $F(2,22)=3.65$, $p<0.05$). Post hoc tests revealed that the differences between pairing driving only and driving+p5, driving only and driving+p3, were significant but no obvious difference between pairing driving+p5 and driving+p3 at Pz and POz was observed. Furthermore, the difference between driving only and driving+p5 disappeared at the CPz and Oz. Nevertheless, the differences in amplitude of P550 between driving only and driving+p3 existed for the whole central-parietal-occipital area. Additionally, there were no differences in the latency when the task load changes: ($F(2,22)=0.56$, $p=0.42$ at Cz; $F(2,22)=0.97$, $p=0.24$ at Pz; $F(2,22)=1.2$, $p=0.19$ at Oz).

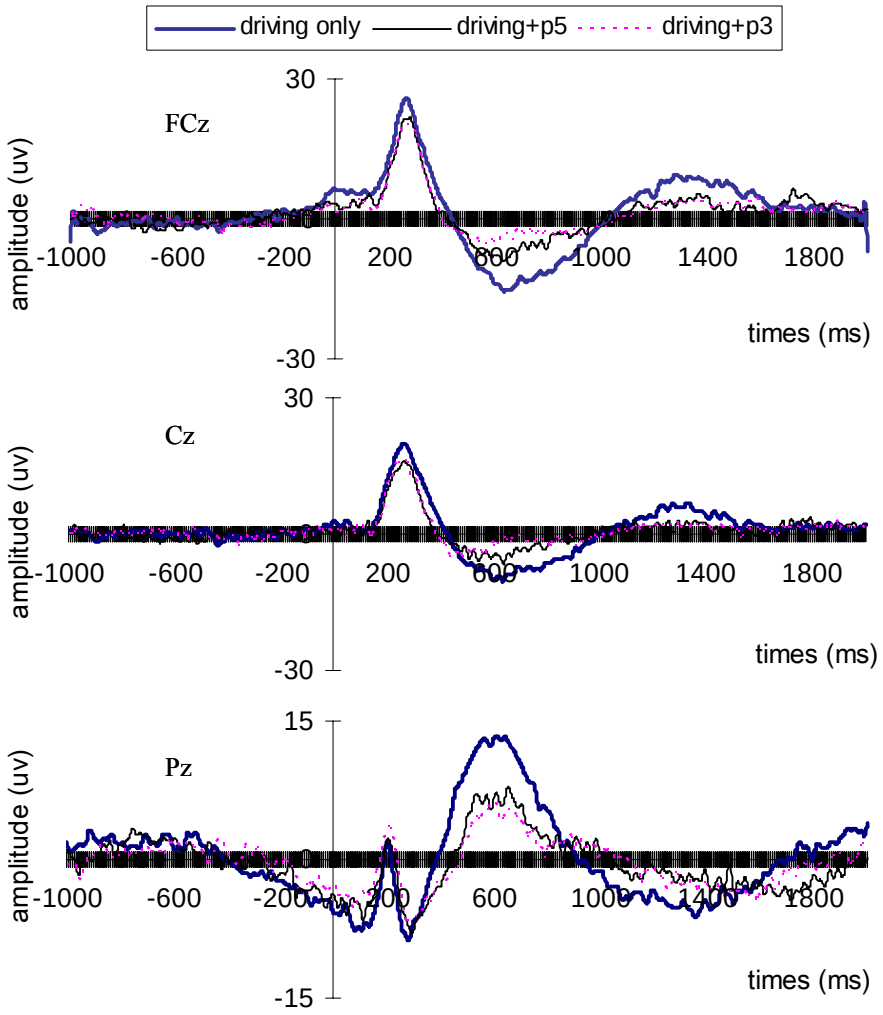


Figure 3. 13 Artefacts-corrected ERPs at different electrodes in dual task conditions (N=24)

3.5.4 Band Powers and workload

Single task: speed related changes in EEG band powers

As indicated in the data analysis section, three band powers (theta, alpha and beta) were extracted to investigate the workload induced changes in these parameters in terms of either spatial patterns or frequency powers. Figure 3.14 shows the spatial patterns of the theta, alpha and beta power induced by the speed changes. Generally, theta power showed a prominent increase in the frontal middle area as

the speed increased. There were also changes in the partial middle area. But these changes were less severe than those in the frontal middle area. In comparison, alpha power showed a general decrease over the whole head as the speed increased. But these changes were dominant in the parietal middle area. The changes in beta power seemed more complex and had no systematic changes with the task load.

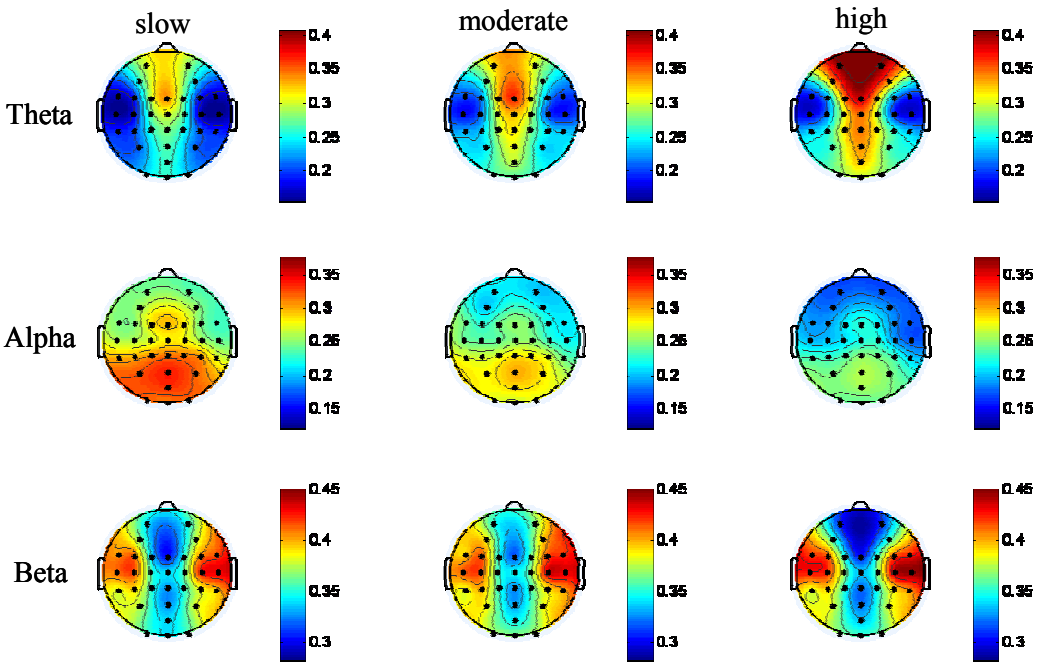
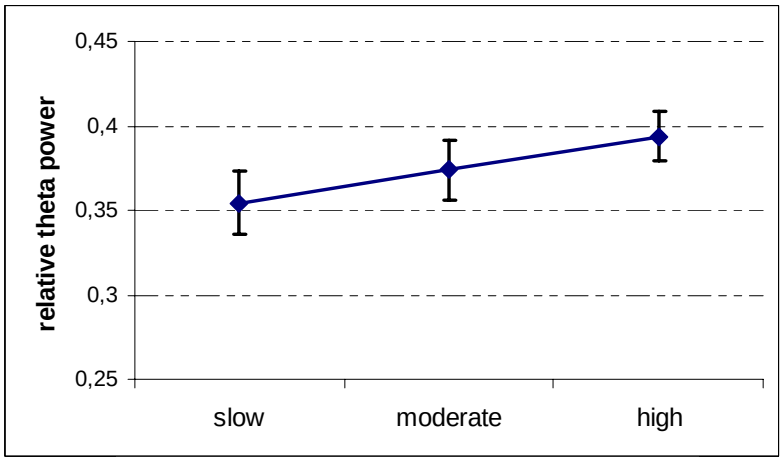
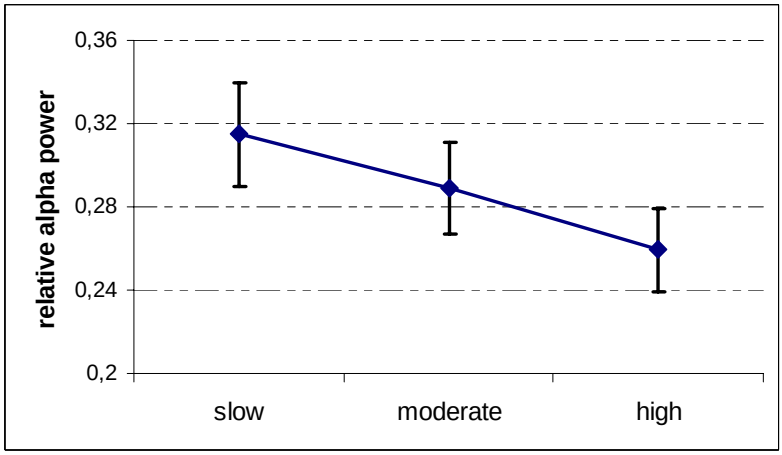


Figure 3. 14 Spatial patterns of theta, alpha, and beta in the single task condition

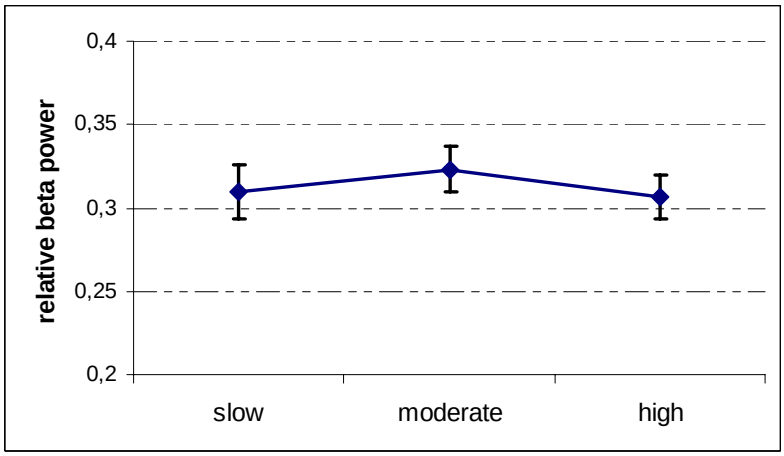
Figure 3.15 shows the changes in theta (at FCz), alpha (at Pz) and beta (FCz) parameters at the electrodes demonstrating the severest changes observed from the spatial pattern. As driving speed is increased, statistical analysis shows that the increases of theta power were significant at the frontal-central area: ($F(2,24)=8.68$, $P<0.01$; at Cz, $F(2,24)=6.3$,$p<0.01$). Pairwise comparison showed that at FCz, the differences between the slow and moderate as well as slow and fast were significant, but there was no difference between moderate and fast condition. Still, at Cz, only the difference between slow and fast condition was significant.



(a) Relative theta power at electrode sides FCz



(b) Relative alpha power at electrode sides Pz



(c) Relative beta power at electrode sides FCz

Figure 3. 15 The changes in theta, alpha and beta powers with speed (N=24)

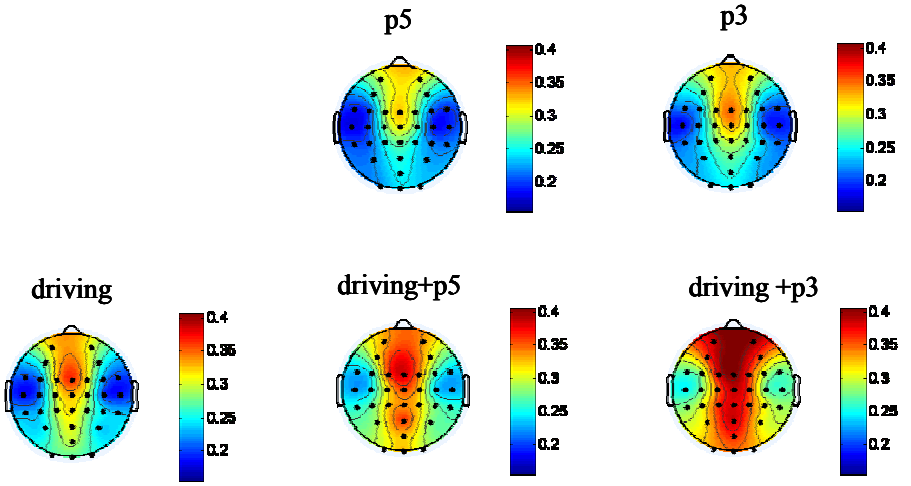
The decreases in alpha power at parietal electrode sites were significant: (at Pz, $F(2,24)=17.6$, $p<0.001$; at POz, $F(2,24)=17.8$, $p<0.001$; at Oz, $F(2, 24)=13.4$,

p<0.001). Pairwise comparison showed that the differences between each pair of these three driving speeds were significant at Pz and POz. But at Oz, the difference between moderate and fast condition was not significant. For beta power, no significant differences were found at FCz ($F(2,24)=1.43$, $p=0.25$).

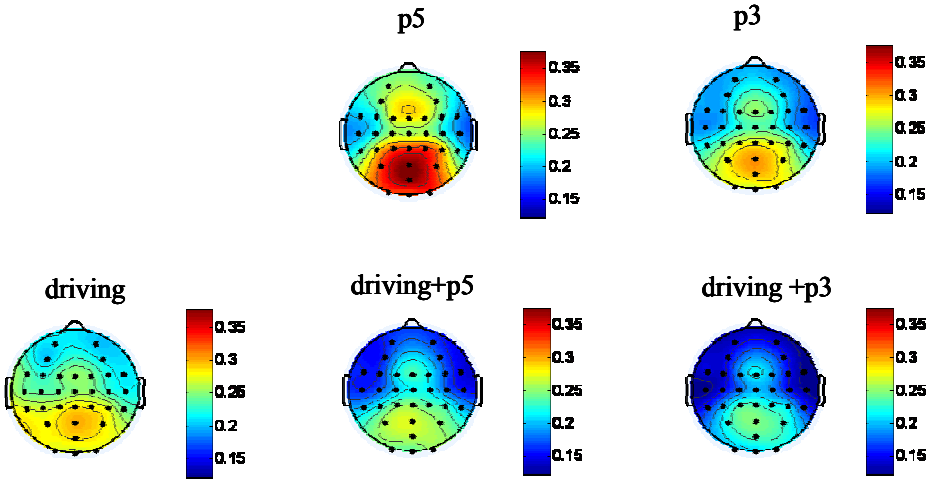
Dual task: task combination related changes in EEG band powers

Figure 3.16 shows the spatial pattern of theta, alpha, and beta in different task conditions. As shown in Figure 3.16, the changes in theta power start from the frontal middle area in the single task condition and spread to the parietal middle area when the combined task load became higher in dual task conditions. In comparison, the alpha changes predominated in the parietal middle area, particularly at electrode sites Pz and POz. Changes in beta power seemed to be focused on both frontal and parietal areas.

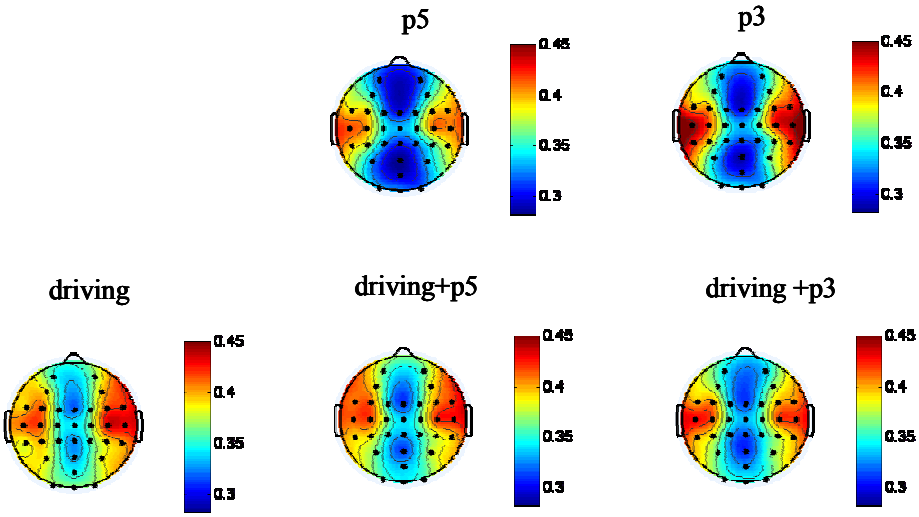
Figure 3.17 shows the changes in theta (FCz), alpha (Pz), and beta (Pz) parameters at the electrodes which demonstrated the severest changes as observed from the spatial pattern. When comparing the driving only condition with two combined conditions (driving, driving+p5, and driving+p3), statistical analysis showed that the increases in theta power were not significant at both frontal and central areas (at FCz, $F(2,22)=1.0$, $p=0.35$; at Cz, $F(2,22)=1.60$, $p=0.21$). However, the decreases in alpha power at parietal electrode site Pz were significant (at Pz, $F(2,22)=11.9$, $p<0.001$; at POz, $F(2,22)=12.2$, $p<0.001$; at Oz, $F(2, 22)=13.3$, $p<0.001$). Pairwise comparison indicated that at Pz and POz, there were significant differences between driving only and driving + p5, and between driving only and driving + p3. However, no difference was found between driving + p5 and driving + p3. But at Oz, the differences between each pair of these three



(a) theta

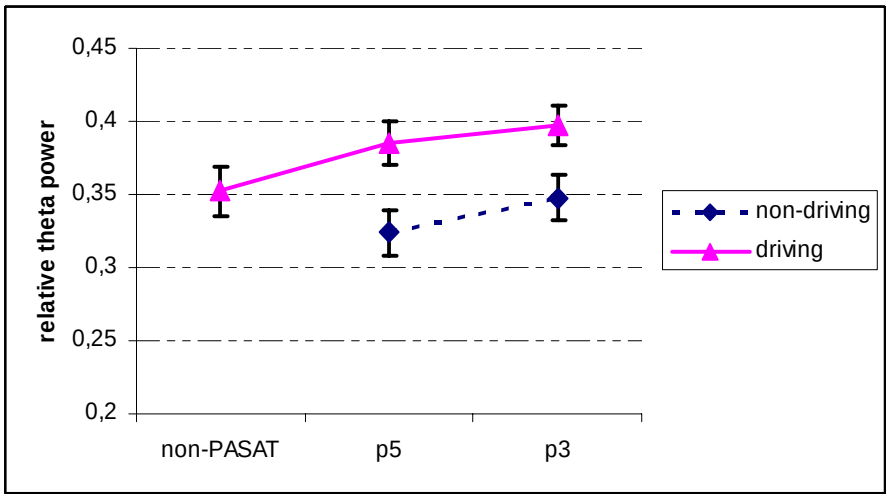


(b) alpha

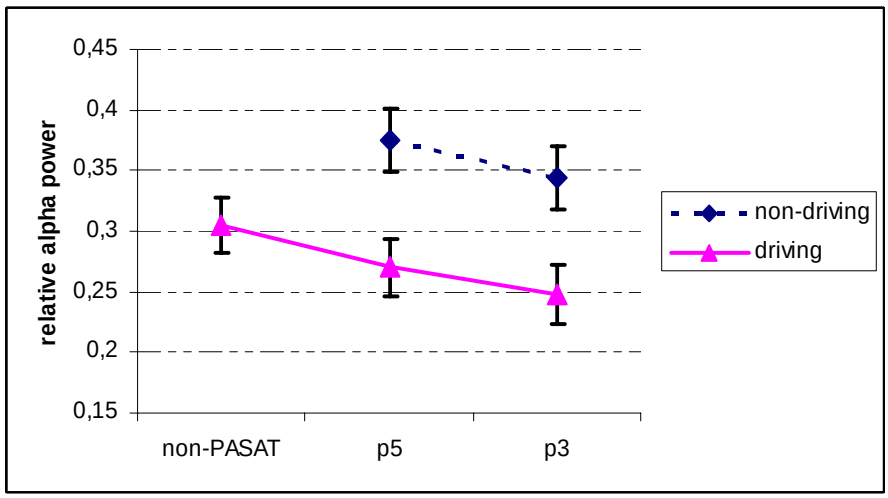


(c) beta

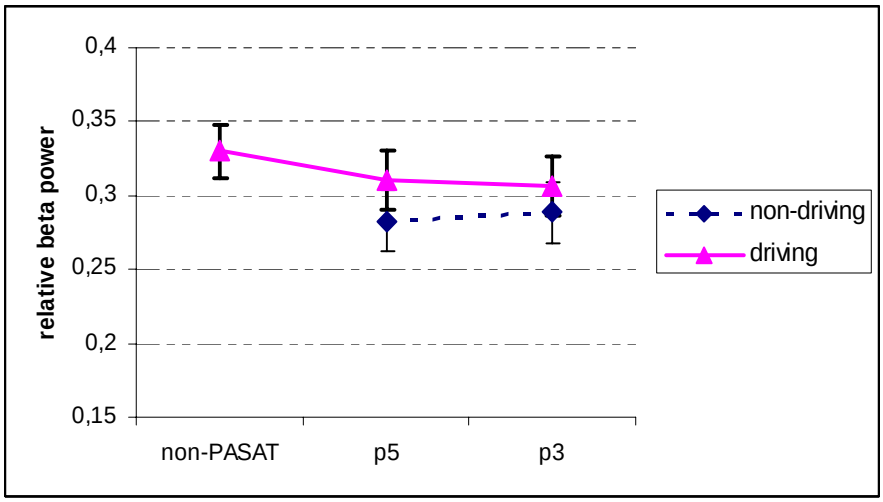
Figure 3. 16 Spatial pattern of theta, alpha and beta in dual task conditions (N=24)



(a) Changes in relative theta power at FCz



(b) Changes in relative alpha power at Pz



(c) Changes in relative beta power at Pz

Figure 3. 17 Comparison of changes in theta, alpha, and beta in dual conditions

conditions were significant. For beta power, no significant differences were found in both frontal and parietal areas (at FCz, $F(2,22)=1.4$, $p=0.26$; at Pz, $F(2,22)=1.66$, $p=0.21$).

When comparing the single PASAT tasks with conditions combining driving tasks (p5, p3, driving+p5, and driving+p3)[24], the combining of driving tasks led to a significant increase in the theta power at frontal area (at, FCz, $F(1,22)=11.7$, $p<0.01$), but the increase in the PASAT load (from p5 to p3) induced no significant difference in theta power (at FCz, $F(1,22)= 0.25$, $p=0.63$). The combining of driving tasks to the PASAT induced a significant decrease in alpha power at the parietal area (at Pz, $F(1, 22)=88.9$, $p<0.001$; at POz, $F(1, 22)=50.3$, $p<0.001$; at Oz, $F(1, 22)=29.8$, $p<0.001$). But again, the increase in PASAT load (from p5 to p3) induced no significant difference in alpha power (at Pz, $F(1,22)=0.75$, $p=0.39$; at POz, $F(1,22)=1.1$, $p<0.31$; at Oz, $F(1,22)=1.3$, $p=0.27$). For beta power, the combination of driving task to PASAT elicited significant differences at both frontal and parietal area (at FCz, $F(1,22)=6.3$, $p<0.05$; Pz, $F(1,22)=4.4$, $p<0.05$). But the increased PASAT load led to no difference in the beta power at both frontal and parietal area (at FCz, $F(1,22)=0.2$, $p=0.66$; Pz, $F(1,22)=0.58$, $p=0.45$).

3.5.5 Classification accuracy

The average results from 20 participants are shown with the error bar chart in Figure 3.18. For the single task condition, ERP and BP based classification demonstrated an average accuracy of 63% (SD=14.9) and 85% (SD=10.0) respectively; for the dual task condition, the averaged classification accuracies were 66% (SD=8.1) by ERP and 85% (SD=8.5) by BP. This indicates that the BP based classifier showed a much higher accuracy in both the single task and dual

[24] Here, we used the two-way ANOVA to investigate the effect of two factors, non-driving and driving, p5 and p3.

task conditions compared with ERPs (for single task, $F(1,20)=28.95$, $p<0.001$; for dual task, $F(1,20)=53.6$, $p<0.001$) .

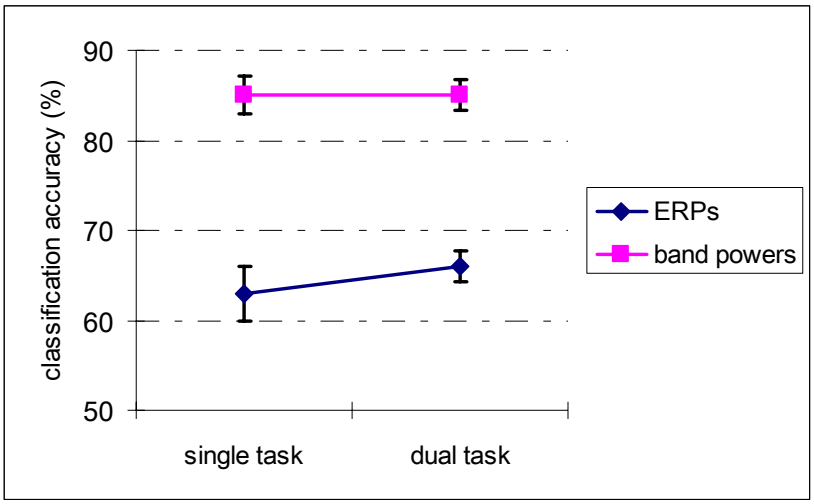


Figure 3. 18 Mean classification accuracy of ERPs and BPs for the single and dual task conditions

3.6 Discussion

3.6.1 What are these components in ERP: A Task Analysis

Lane Change Task is a typical two-stimulus task paradigm. In Lane Change Task, a white sign without any information is first presented to the participant and could be considered as a warning stimulus S1, which draws the participant's attention towards preparation for command perception. Later, the road sign is presented with some crosses and arrows, which indicate which lane the participant should change to, and this could be the second stimulus S2 allowing for the participant's perception of lane change direction and readiness for taking action. These two stimuli are both visually presented.

So far, a good number of researchers have investigated the visual stimulus evoked by EPRs. Generally, there is a late wave complex consisting of an enhanced P2 (200-250ms), a negative peak (designated in literature as N200 or N250, N2) and a fairly large prominent positive peak (P300) (Halgren et al., 1995; Giger-Mateeva et al., 1999; Kubová et al., 2002). P300 has been reported to consist of two sub-components - P3a and P3b. P3a exhibits a frontal/central maximum peak at the latency of 300 ms, whereas P3b is located in central-posterior area with the maximum around the latency of 500ms.

In the present study, both ICA and artefacts corrected ERP demonstrate the existence of a frontal/central positive maximum around 250ms and a positive central-posterior maximum around 550ms. The frontal-central P250 could be either a P2 or an early P3a (200-300ms). P3a seems involved in the non-target and involuntary attention, whereas P2 is usually observed after any visual stimulus, and usually has its negative aspect at the parietal and occipital area (N2). This is highly consistent with our finding in the present study. Thus, we rather conclude it as a P2 wave. The central-posterior peak around 550ms is a later P3b, which is reported to be involved in the participant's voluntary attention. Additionally, CNV, which is a common cognitive response in the two-stimulus paradigm, is also obtained in the present study.

3.6.2 Effect of task load on the amplitude and latency of P300

As discussed in section 2.4.3, the amplitude of P300 is thought to be sensitive to the variation of perceptual and processing load. It has been suggested that the amplitude of P300 reflects the amount of attentional resources allocated for the task performance (Johnson, 1986; Polich, 2004). This theory could explain the experimental findings in both single task and dual task paradigms. For instance, a large number of Oddball paradigm researchers would suggest that the amplitude

of P300 for the target is much larger than that of the non-target P300, since the rarely presented stimuli may attract a higher attention than those frequently presented. De Waard (1996) suggested that in a primary-only-task, the P300 amplitude increases with task complexity. In a dual task, Raabe et al. (2005) reported that the amplitude of P300 decreased with the task load in an oddball task. The P300 was evoked as the secondary auditory task including 96 frequent tones and 24 seldom tones. The primary task was set by two task load conditions, self-paced driving and car-following. Admittedly, the primary task in the higher workload level (car following) draws the attention of the subject more than in the lower workload level (self-paced driving), which leads to the fact that less attention is allocated to the secondary task so that the amplitude of P300 is decreased.

In the present study, the amplitude of P300 showed a decrease when the driving task was combined with a secondary task. This finding is consistent with the previous findings (Wickens et al., 1977; Kramer & Parasuraman, 2007; Schultheis and Jameson, 2004; Raabe et al., 2005). However, in previous studies the P300 is often evoked by a secondary task and the increase of the primary task load leads to a decrease in the amplitude of P300. This is different from what we did in the present study. The P300 we analyzed is evoked by the primary LCT rather than by the secondary task. Nevertheless, the theory about the P300 and workload mentioned above can also interpret this result. Wickens (1983) suggested that the resource allocation to one task might lead to a withdrawal of resources from another task. Therefore, in our study it is reasonable to interpret that the decrease in the P300 evoked by LCT is due to the distracted attentional resources for the PASAT performance.

In the single task paradigm, we found that the amplitude of P300 also decreased with the increased task load induced by the speed settings. It appears that this

finding is contradictory with previous suggestion that in a primary-only-task the P300 amplitude increases with task complexity (De Waard, 1996). However, a detailed analysis of the subcomponents of the LCT makes it reasonable to assume that the LCT itself could be considered as a dual-task. It can be seen that the primary task is keeping the vehicle in the lane while the secondary task is to perceive the road sign for lane changes. P300 is generated in response to the visual stimuli in the secondary task. As the speed increases, more attention is paid to lane keeping and car controlling while less attention may be allocated to visual perception. Thus, the amplitude of P300 may decrease. Nevertheless, there seems to be no difference in the P300 amplitude between conditions 80 km/h and 100km/h. This might well be the case when the difference in the mental workload under these two conditions is not distinguishable.

Unfortunately, in both single task and dual task conditions, there are no differences in the latencies of P300 as the task load is increased. The latency of the P300 is thought to reflect the timing of stimulus identification and categorization processes. The results in the present study suggest that neither the changes in the speed setting nor a combination of the PASAT with the LCT change the subject's timing pattern for stimulus identification.

3.6.3 Effect of task load on the EEG spectrum parameters

The spontaneous EEG signal yielded sensitive indices for neural resources utilization and systematic variation with the task load. In the single task condition, there were generally significant increases in the theta power and decreases in the alpha power as the task load increased with speed settings. When comparing the driving only task condition with the two dual task conditions, there were significant decreases in the alpha power but without significant differences in the theta power; when comparing the PASAT only conditions with the dual task

conditions, there were significant increases in the theta power and significant decreases in the alpha power in the driving conditions when compared with non-driving ones. However, when the PASAT load is increased from p5 to p3, there were no significant differences in both theta and alpha power. In both single and dual task conditions, there were no significant differences in the beta power.

The findings that increased task load induced increases in the frontal theta are consistent with previous studies (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998). The underlying mechanism of the workload associated changes in the theta power has been investigated in various studies. As introduced in section 2.4.3, the theta power was thought to associate with the attention control mechanism, which is suggested as one of the functions of the anterior brain network (Posner et al., 1990, 1992; Smith et al., 2001). Kubota et al. (2001) also argued that frontal midline theta occurs with the activation of attentional systems in prefrontal circuitry involving the anterior cingulate cortex. In the present study, as the driving speed increases, it is supposed that the attentional demand for the driver would also increase (Senders et al., 1967; McDonald et al., 1975; Cnossen et al., 2000) and consequently, the theta power in the frontal area increases. However, the discovery that a combination of the PASAT with the driving task elicited no significant differences did not match our expectation, since various studies had indicated that increased working memory load may produce an increase in the frontal theta power (Sirevaag et al., 1988; Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998). Nevertheless, there are also findings indicating that the theta power would decrease with transitions from the single- to dual-task conditions (Sirevaag et al., 1988). It seems that the modulation of theta power with the task load might be context dependent, especially when comparing single tasks with dual tasks.

The discovery of decreased alpha power with task load in both single task and

dual task conditions was consistent with previous findings (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Wilson, 2001, 2002b; Hagemann, 2008). For instance, Gevins and Smith (2000) found that alpha power tends to attenuate in the high-load task compared to low-load task in a single n-back task; Sirevaag et al. (1988) showed that there were decreases in alpha power in the dual task compared to the single task. These findings match the concept that alpha oscillations reflect the degree of neural resources utility during task performance[25]. When the task demand increases, more neural groups may be recruited in the function network, and the number of alpha generators decreases. This leads to a reduction in alpha power (Smith et al., 2001).

In the present study, we found no systematic changes in the beta activity in both single and dual task conditions. It is suggested that the beta wave represents cognitive consciousness, and active, busy or anxious thinking (Dooley, 2009). It is also found that the task load also modulates beta activity (Wilson and O'Donnell, 1998). However, previous results concerning the EEG beta power modulation with workload are controversial. For instance, it has been found that there is increased beta activity during complexity manipulations (Wilson and O'Donnell, 1998; Brookings et al., 1996). While Dussault et al. (2005) have shown that rest sequences resulted in higher beta power than active segments in a simulated flight task (for similar results, see also Doyle et al., 2009). Therefore, the sensitivity of beta power to the task load might well be context dependent.

It is also necessary to point out that the data processing procedures may have great influence on the results. Lei et al. (2009c) used a wavelet transform[26] to decompose the EEG frequency activities (delta, theta, alpha, beta, and gamma)

[25] see section 2.4.3, p. 59

[26] Daubenchies D6 wavelet function was used to perform a 6-level decomposition of the EEG signal. The decomposition coefficients a6, d6, d5, d4, and d3 represent delta (1-4Hz), theta (4-8Hz), alpha (8-16Hz), beta (16-32Hz), and gamma (32-64Hz) respectively.

with the same datasets from this experiment. Band powers were normalized by the ratio of the power of specific frequency band to the total power of these five frequency bands for each epoch. The results showed that no changes in five frequency bands including theta, alpha, and beta were found in the single task conditions when the task load levels were manipulated by the speed settings. However, there were indeed some changes, such as the increase in parietal theta, decrease in beta in both the central and parietal area, and decrease in parietal alpha, when the task load levels were modulated by the combination of two-paced PASAT with the driving task. These findings were distinct from the results presented in this dissertation, in which different data processing methods were used. Therefore, establishing standard data analysis procedures for EEG spectrum parameters is definitely essential for obtaining comparable results and arriving at a more convincing conclusion.

3.6.4 Which is robust for workload representation: ERPs or band powers?

There are both advantages and disadvantages for using ERP components and band powers to track the dynamic variation of workload. A unique feature of the P300 or other ERP components is that it is simultaneously sensitive to the allocation of attentional resources (amplitude) and to the timing of stimulus identification and categorization processes (latency). De Waard (1996) also pointed out that the main advantage of the ERP-technique is its high diagnosticity to perceptual/cognitive processing, and its insensitivity to response factors. While the obvious merit for using band powers in workload representation is that it could provide a continuous assessment without intruding into task performance.

As presented in the results section, both ERPs (mainly P300) and band powers demonstrated the ability for workload assessment. Generally, band powers (theta and alpha power, but not beta power) showed a higher statistical significance level

of differences in EEG parameters induced by the task load in either single task condition or dual task condition. In a further study, we used a machine learning methods, Adaboost, to classify the instantaneous workload. In this study, the ERP features were extracted from the single trial ERPs while band power features were extracted every 2s with 1s overlay. The band power based classification showed a much higher classification accuracy than the ERP based classification. This evidence supported the notion that band powers may be more robust in distinguishing the task loads, at least in the present study. There are indeed several reasons for the inefficiency of ERPs to index the operator's mental workload states.

Firstly, the use of ERP to distinguish workload levels is traditionally based on the average of a large number of trials. In the present study, the analysis of ERPs is based on the average of more than 70 trials for each condition. Therefore, it is doubtful whether the P300 amplitude computed on the basis of a few trials or even a single trial can reliably discriminate between different levels of workload. Although Humphrey and Kramer (1994) provided some relevant evidence to show that a high accuracy (approximate 90%) could be achieved which discriminates low- and high- workload conditions based on a few samples (5-10 trials, approximate 25-50 seconds), and Welke et al. (2009) also showed that the human brain contains information preceding the movement execution extracted for a single trial detectability, such evidence is still not enough for real-time workload classification and there is still no adequate evidence for such approaches. The results are in line with the lingering suspicion regarding the feasibility of using the single trial ERPs for workload detection. For instance, Baldwin et al. (2004) speculated that an ERP (event-related potential extracted from EEG data) based method could be used to trigger adaptive automation during high workload by engaging driver aiding systems and disengaging secondary systems such as cellular phones or entertainment devices.

Furthermore, the variation of workload is continuous. However, ERP components fail to provide a continuous monitoring. ERP components are time-locking to an external or maybe also internal event. But, in reality, there might be no constant generation of this event. This might lead to an information-gap in workload states during the non-event period. Some studies attempted to introduce an artificial secondary task to produce the needed events or stimuli (Raabe et al., 2005). However, such methods might produce unnecessary distraction to the operator from the primary task and thus introduce potential impairment for the task performance. In contrast, the EEG band powers could provide a continuous assessment (e.g. every 1 or few seconds) without introducing any other performance-intruding secondary tasks.

Therefore, we prefer to use the EEG band powers for the workload assessment. However, there is also a certain inefficiency when using band powers. The band powers are prone to be contaminated by the artefacts, including both ocular and muscle artefacts. Therefore, careful data cleaning processing is important for drawing any validated conclusion from the band power results. In the present study, the PASAT was presented auditorily, which might induce less ocular artefacts. However, it required a verbal reaction, i.e. participant had to speak out the results. This caused a lot of muscle artefacts. Worse, the speaking is often accompanied by involuntary biting and chewing, which induced muscle artefacts as well. In the data processing, we spent a lot of time on cleaning the data, which was almost always done by visual inspection with the help of ICA. Therefore, for an EEG study, the artefacts-free task design would be preferred. Of course, a better solution for cleaning the artefacts would be beneficial in terms of obtaining more reliable results.

Although we used the machine learning methods (Adaboost) to classify workload, there were still some issues that had to be addressed. Using the traditional

machine learning methods for workload classification is problematic. The selection of features is prone to involve non-workload components. Many studies used a large feature vector for classification, for instance, delta, theta, alpha, beta, and gamma from different electrode sites (Lei et al., 2009d; Wilson & Russell, 2003b). This kind of classification has a 'black box' effect and could not guarantee the reliability of the classification results. Furthermore, the classifier training with the predefined labels responding to the predefined task load might also be problematic. As discussed in the first chapter, the workload may not be linearly associated to the task load, i.e. a high task load does not induce equally a high workload. On this point, the label assigned for the classifier training may itself be incorrect, which also leads to an incorrect classifier. Therefore, the quantification of workload using band powers requires improved mathematic algorithms.

3.7 Summary

This chapter was aimed at determining whether the previous findings concerning changes in EEG parameters with task load can be reproduced in the driving context while also finding the optimal parameters in the EEG signal for driver mental workload representation. For the manipulation of task load levels in the LCT, a pre-study investigating the influence of driving speed on the LCT in driver mental workload was conducted and the results showed that the increased driving speed contributed to increased subjective load and a decline in driving performance. This indicated that the driving speed is an important contributor to driver mental workload. The formal study consisted of a single task paradigm and a dual task paradigm. In the single task paradigm, the task load levels were manipulated by different speed settings in the LCT, while in the dual task paradigm the task load levels were generated through a combination of two paced PASAT with the LCT. Results indicated that the amplitude of the P300 in the

parietal area significantly attenuated with the task load in both single and dual task paradigms, while the EEG theta and alpha power also demonstrated significant changes with the task load. However, the classification of workload with the Adaboost using these two groups of parameters (ERPs and EEG band powers) indicated that the band powers would be more robust for the workload detection. In the discussion section, the results were discussed in the context of their consistency and inconsistency with previous studies. The discussion of ERPs and EEG band powers suggested that both methods have unique merits for the workload assessment. However, in the case of short-term detection (e.g. in seconds), the EEG band powers are thought to be more efficient.

Chapter 4. EEG spectrum modulation with task combination

4.1 Motivation

Driving task is complex and driver's workload might be influenced not only by the primary driving task, but also simultaneously by some other factors, e.g. driver sometimes has to search for some information outside the vehicle and keep it in mind, which may involve working memory load. The modulation of EEG spectrum induced by a simple task in which workload levels are manipulated in a single dimension (e.g. working memory load) has been systematically investigated in previous literatures (Gundel and Wilson, 1992; Gevins et al., 1997, 1998) as well as in our study presented in the previous chapter (see also Lei et al., 2009 a, b). There are indeed also a large number of studies involving relatively complex tasks (Smith et al., 2001; Wilson & Russell, 2001, 2003a). But most of these studies investigated the EEG spectrum modulation in a general manner or directly used the machine learning methods to classify the workload levels with various EEG variables (Kohlmorgen et al., 2007). There is still a lack of a clear comparison of the similarities and dissimilarities in the modulations of EEG parameters when they are induced simultaneously by different workload dimensions. The answers to the questions, such as how these sub-resources of workload contribute to the general workload and what kinds of effect are involved after these sub-resources are combined, should be addressed. Therefore, in this chapter, we focus on the investigation of the inter-factor effect of workload on the EEG parameters, with a particular concentration on the frontal theta and parietal alpha power, which have been concluded as robust candidates for workload detection in Chapter 3.

Additionally, the sensitivity or robustness of different methods for workload

assessment seems task-dependent and controversial. For instance, the robustness of EEG observation is also supported by Brookings et al. (1996), whereas Doyle et al. (2009) suggested that EEG measures exhibit less sensitivity than HR when distinguishing cognitive loads during a satellite management decision-training task. It has been suggested that each of the physiological parameters provides unique information concerning cognitive load (Wilson, 2002a,b). Therefore, a systematic comparison of the robustness of the different workload assessment techniques in driving context would be meaningful. Furthermore, it would be also interesting to investigate the correlations of the EEG parameters to other workload techniques.

Following these objectives, a two-factor 3*3 task design was used in this study. The task loads were manipulated in two dimensions, driving task load and working memory load. Again, the LCT task involving several speed settings was used to produce driving task loads. For working memory load, as discussed in Chapter 3, the PASAT causes many artefacts in EEG data because of the verbal response. Therefore, in this study another working memory task, the n-back task, (Kirchner, 1958; Walter et al., 2001) which requires a silent manual response was combined with the LCT. Various variables, including the EEG, ECG, performance, and subjective load were recorded to provide an insight of the physiophysiogical behaviour behind workload. The following sections describe the details of this study, including the methods, data analysis, results, discussion, etc.

4.2 Methods

4.2.1 Participants

Overall, 26 participants (19 males, 7 females) participated in the present study. They all ranged in age from 21-33 with a mean age of 27.8 (standard deviation =

2.96 years). All individuals reported to be free of illness and medication. None had any prior experience with the experimental task. Data from 2 participants were excluded because of incomplete physiological recordings due to simulation sickness during the experiment. They all received a cash payment for their participation in the study.

4.2.2 Experiment apparatus

The experiment was conducted using a driving box, which was the same as in the study described in chaper 3. Two projectors were used to project the LCT driving scene and n-back digits respectively. The driving scene was projected onto the wall in front of the driving box (about 1 metre from the driving box). The n-back digits were visually presented to the participants with projections overlaid upon the driving scene using another projector (Figure 4.1). Participants were seated on the driving seat, which could be adjusted to a comfortable position. There were also buttons embedded in the steering wheel, which allow the participant to react to the n-back task.

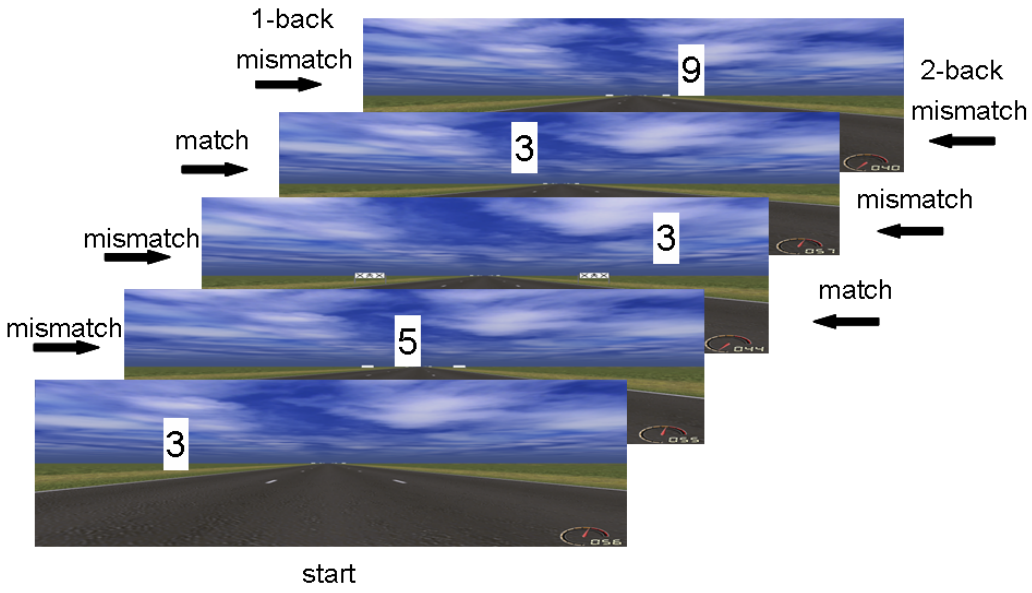


Figure 4. 1 The overlaid projection of the LCT driving scene and n-back task. The digits were presented every 3 seconds with 1 second presence.

The EEG recording system was same with that used in the first experiment (see, Chapter 3). What was different was that one channel of electrocardiogram (ECG) was used to collect the heart beat information using two bipolar electrodes. One electrode was attached at the upper breastbone and the other was attached around the lowest rib on the left side of the body. ECG was sampled at 1000 Hz and recorded together with EEG, EOG, EMG as one extended channel using the software of Brain Vision Recorder from Brain Products.

The signals from LCT, n-back, and EEG were synchronized using a self-developed tool based on Labview (National Instruments, USA). This tool automatically triggered the n-back program when it detected the start of each track of LCT and shut it down when it came to the end of each track. It also sent the synchronization markers to Brain Vision Recorder when it detected the road signs. The n-back program was developed by Christian Kothe who was one of the PhyPha team in Chair of Human-Machine-System at Berlin Institute of Technology. This program sent also marks to the Brain Vision Recorder to synchronize the data.

4.2.3 Tasks

The LCT and n-back task were used in this study. The details of the LCT were described in Chapter 3. The n-back task was firstly introduced by Kirchner in 1958 and is commonly used in neuroimaging to stimulate brain activity for working memory capacity test. The subject is presented with a sequence of stimuli (in this study, the stimuli were digits) and the task consists of indicating whether the current stimulus matches the one from n steps used earlier in the sequence. The load factor ‘n’ can be adjusted to make the task more or less difficult. In this study, the digital stimuli were presented every 3 seconds with 1-second presentation duration. Every time, the digit appeared in a random location on the

driving scene. Only 1-back and 2-back were used. Participants reacted to n-back tasks by pressing the left or right button in the steering wheel to indicate “match” or “mismatch” of the digits.

A 3*3 two-factor within-subject design was used in the present study. Task load levels were manipulated with two dimensions, the driving task load and the working memory load. The dimension of the driving task load consisted of three conditions: no driving (‘base’; the driving task was performed by the experimenter at a fixed driving speed of 75 km/h), slow driving (‘slow’; the driving speed was fixed at 75 km/h and participants used the steering wheel to perform lane changes), and fast driving (‘fast’; the driving speed was fixed at 100 km/h and participants used the steering wheel to perform lane changes). The dimension of working memory load also contained three conditions: no n-back (N0; digits were still presented to participants but required no reaction from the participants), 1-back (N1), and 2-back (N2). The combinations of the driving conditions and n-back modes manipulated in total 9 task load levels.

4.2.4 Procedures

Participants first filled out a form with their personal information (age, driving experience, illness and medication situation, etc.) before reading an experiment introduction. After the electrode preparation (around 20 minutes), they had a 20-minute practice session with all 9 task conditions. Then, they were required to evaluate the NASA-TLX weights of six workload contributors. Afterwards, they were asked to perform three randomly organized sessions (‘base’, ‘slow’, and ‘fast’) with a 5-minute break after each session. Both ‘base’ and ‘slow’ sessions comprised 9 randomized driving trials including 3 trials on each n-back condition lasting 2.5 minutes per trial. The ‘fast’ trials were about 2 minutes each. To keep the same duration for each experimental condition, the ‘fast’ session comprised 12

trials (randomized as well), 4 trials per n-back condition. Generally, each session lasted 25 minutes. During each session break, participants were asked to report NASA-TLX ratings. The whole experiment lasted around 2 hours and 40 minutes.

4.2.5 Data analysis

In total, 8 parameters, i.e. frontal theta power (fro-theta), parietal alpha power (par-alpha), subjective load (NASA-TLX), heart rate (HR), heart rate variability (HRV), LCT performance, n-back performance, and reaction time (RT), were extracted to examine variations with the task load.

EEG analysis

The EEG analysis was performed using EEGlab 6.03, a freely available open source toolbox which runs under Matlab 7.3.0 (see, Delorme & Makeig, 2004). The general processing of the EEG data is shown in Figure 4.2. EEG data was firstly digitally filtered using band pass filter (pass band 1 to 40 Hz) to minimize drifts and line noises. Since EEG data involve plenty of eye movement artefacts, Independent Components Analysis (ICA; Makeig et al., 1996; Delorme & Makeig, 2004) was used for ocular artefacts removal. The EEG data was then segmented into 10-second epochs with 50% overlay (short-term variation of EEG parameters is also our concern though not included in this paper). An FFT analysis then estimated the spectrum for each epoch, while the relative power density spectrum (percentage to total power of frequency range 4-30 Hz) was calculated for each epoch. Frequency bands, theta and alpha, were extracted by accumulating the power of frequency bands 4-8Hz and 8-12Hz. This way, an average of 100 data points in theta and alpha power were extracted for each channel, task

condition and participant. To reduce the individual variation, the z-scores[27] of theta and alpha power were calculated over 900 data points (including all task conditions) for each participant. These data points were used to plot the topography to investigate spatial patterns of the changes in theta and alpha powers induced by task load. Later, the fro-theta powers were calculated by averaging the theta power over 5 frontal electrodes (Fz, FC1, FCz, FC2, Cz) and alpha power by averaging over 5 parietal electrodes (PCz, P3, Pz, P4, POz). The means of fro-theta and par-alpha were extracted by averaging over the data points within task condition and participant. Additionally, to investigate the variability of these short-term EEG parameters in different task load conditions, the standard deviation (SD) was calculated over the data points within each task conditions and participant before z-transformation.

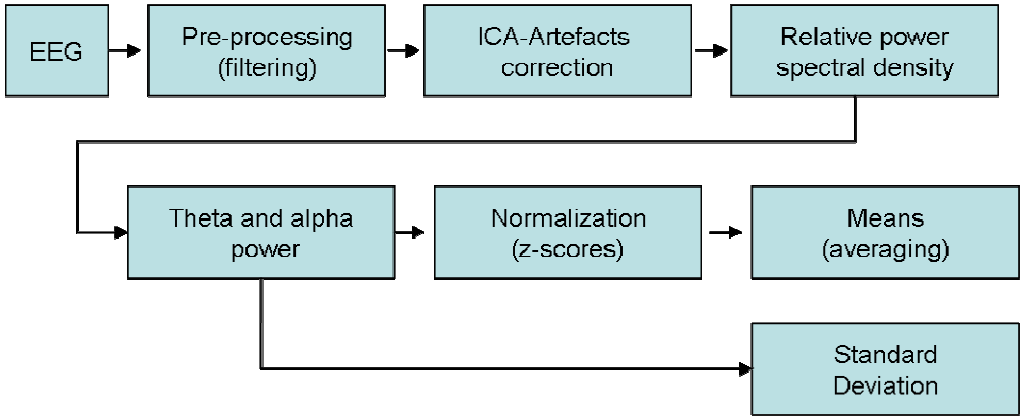


Figure 4. 2 The procedures for the EEG data analysis. The means and standard deviations of the theta and alpha power were extracted. The means were calculated by averaging over data points within each task condition and participant. The standard deviations were calculated over data points within each task condition and participant.

ECG analysis

A module supplied with EEGlab software used a QRS detection algorithm to

[27] Z-score is often used as a standard score in the statistics. It is calculated through Z-transform, in which the population is normalized by dividing the standard deviation after subtracting the mean of the population.

determine individual heart beats from the ECG signal. HR was calculated as the number of QRS divided by the duration for each driving track. The IBI of successive heart beats was then utilized to calculate the heart rate variability (HRV) statistics. The spectrum of HRV was estimated using a method developed by Malik et al. (1996). The ratio of energy around the low frequency (0.04-0.15Hz, LF) activity to the energy around the high frequency (0.15-0.4Hz, HF) activity was analyzed to offer an alternative variable for the workload evaluation.

Statistical Analysis

This was performed with PASW Statistics 18 (SPSS Inc., USA). A two-way ANOVA was used to test the significance of the differences in these parameters, and then multiple comparisons were made with post-hoc analysis (Bonferroni). An alpha of .05 determined the statistical significance.

4.3 Results

4.3.1 Subjective load (NASA-TLX)

As illustrated in Figure 4.3, the subjective load showed significant increases in both dimensions (for driving task load, $F(2, 24)=34.7$, $p<0.001$; for working memory load, $F(2,24)=95.2$, $p<.001$). The interaction of these two dimensions was also significant ($F(4, 24)=7.7$, $p<.001$). Post-hoc test indicated that within each n-back level, significant differences between each pair of driving task conditions were found, but with no differences between 'slow' and 'fast' in N1 and N2 conditions (in N0, 'base' & 'slow', $p<.05$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p<.01$; in N1, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p=.14$; in N2, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.001$; 'slow' &

'fast', p=.09). Within each driving task level, significant differences between each pair of n-back conditions were found (for all pairs, p<.001).

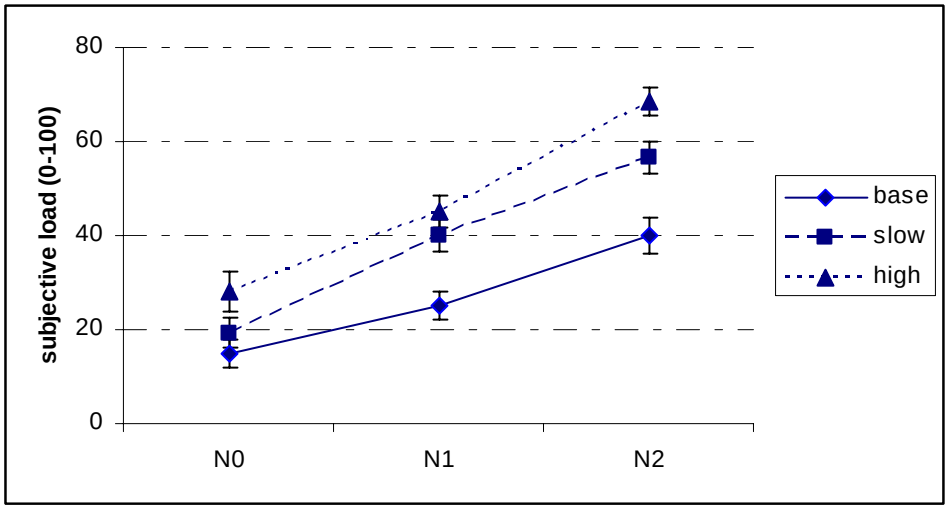


Figure 4. 3 The subjective loads in different task conditions averaged over participants (N=24). The error bar shows the standard error.

4.3.2 Task performance

As shown in Figure 4.4., the mean deviation in LCT showed significant increases with augmented driving task load ($F(1,24)=89.4$, $p<.001$) and working memory load ($F(2,24)=17.1$, $p<.001$). There was also a significant interaction effect ($F(2, 24)=5.9$, $p<.01$). Post-hoc test showed significant differences in mean deviation between 'slow' and 'fast' within each n-back level (in N0, 'slow' & 'fast', $p<.001$; in N1, 'slow' & 'fast', $p<.001$; in N2, 'slow' & 'fast', $p<.001$). Within each driving task level, there were significant differences in mean deviation between pairs N0 and N1, N0 and N2, but none between N1 and N2 (in 'slow', N0 & N1, $p<.05$; N0 & N2, $p<.05$; N1 & N2, $p=.64$; in 'fast', N0 & N1, $p<.001$; N0 & N2, $p<.001$; N1 & N2, $p=.88$).

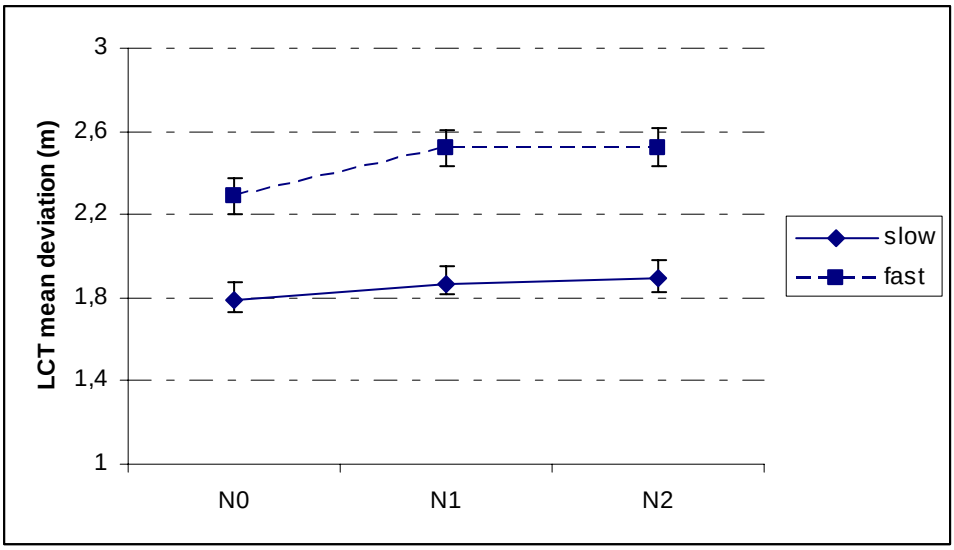


Figure 4. 4 The mean deviation for LCT in different driving conditions averaged over participants (N=24). The error bar shows the standard error.

The n-back error rate (Figure 4.5) showed significant increases with augmented driving task load ($F(2, 24)=12.6$, $p<.001$) and working memory load ($F(2,24)=26.2$, $p<.001$). There was also a significant interaction effect ($F(2, 24)=4.7$, $p<.05$). Post-hoc test indicated that within each n-back level, significant differences were found between each pair of driving task load conditions except for pairing between 'slow' & 'fast' in N2 condition (in N1, 'base' & 'slow', $p<.05$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p<.01$; in N2, 'base' & 'slow', $p<.01$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p=.06$). Significant differences between N1 and N2 conditions were found within each driving task level (in 'base', N1 & N2, $p<.01$; in 'slow', N1 & N2, $p<.001$; in 'fast', N1 & N2, $p<.001$).

Additionally, a significant RT delay for n-back task was found with increased driving task load ($F(2, 24)=16.1$, $p<.001$) and working memory load ($F(1,24)=69.7$, $p<.001$), shown in Figure 4.6. A significant interaction effect was also observed ($F(4, 24)=5.8$, $p<.01$). Post-hoc test indicated that for each n-back level, there were significant RT differences between 'base' and 'slow', 'base' and 'fast', but none between 'slow' and 'fast' (in N1, 'base' & 'slow', $p<.001$;

'base' & 'fast', p<.001; 'slow' & 'fast', p=.11; in N2, 'base' & 'slow', p<.01; 'base' & 'fast', p<.05; 'slow' & 'fast', p=.77). Significant differences between N1 and N2 conditions were found within each driving task level (in 'base', N1 & N2, p<.001; in 'slow', N1 & N2, p<.001; in 'fast', N1 & N2, p<.001).

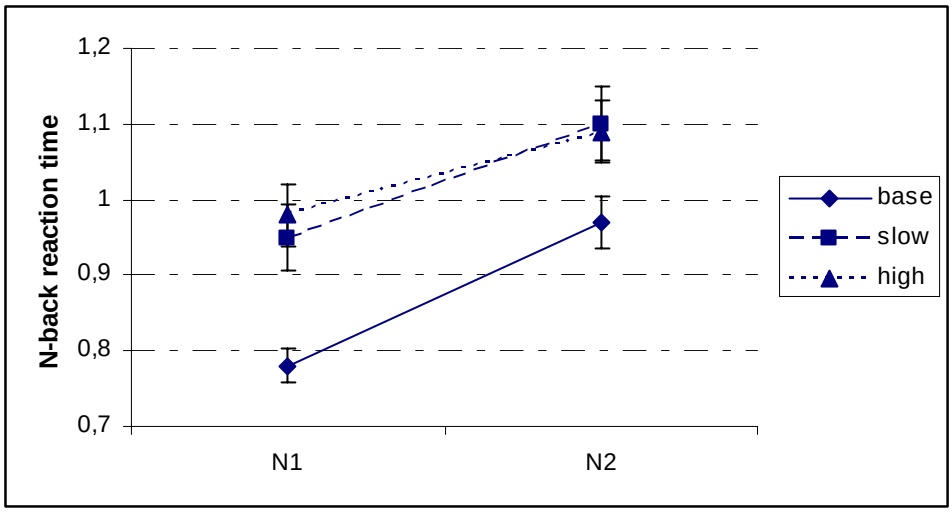


Figure 4. 5 The n-back reaction time in different task conditions averaged over participants (N=24). The error bar shows the standard error.

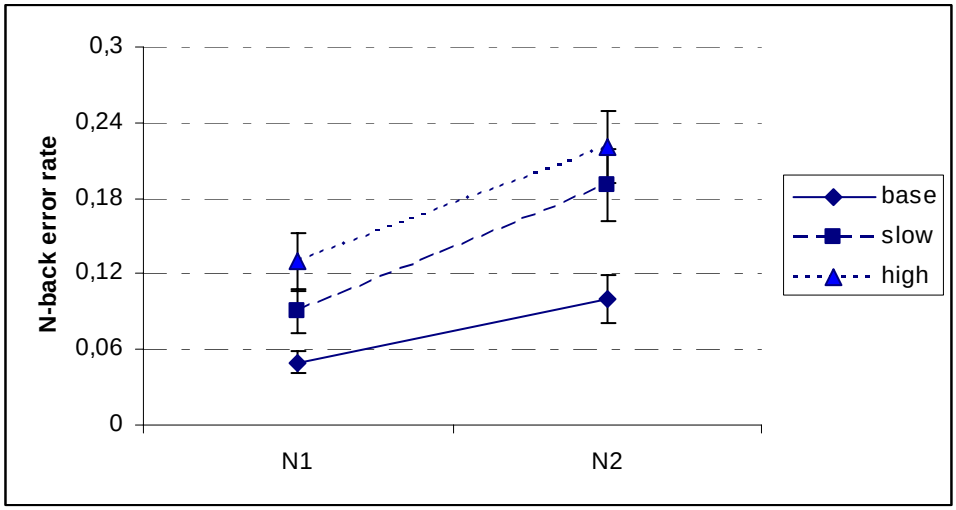


Figure 4. 6 The n-back error rate in different task conditions averaged over participants (N=24). The error bar shows the standard error.

4.3.3 Heart rate and heart rate variability

As shown in Figure 4.7, a significant rise in HR was found as the task load level increased in both dimensions (for driving task load, $F(2, 24)=12.2$, $p<.001$; for working memory load, $F(2,24)=31.9$, $p<.001$) with significant interaction effect ($F(4, 24)=8.1$, $p<.001$). Post-hoc test indicated that within each n-back level, there were significant differences in HR between 'base' and 'slow', 'base' and 'fast', but none between 'slow' and 'fast' (in N0, 'base' & 'slow', $p<.05$; 'base' & 'fast', $p<.05$; 'slow' & 'fast', $p=.43$; in N1, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.01$; 'slow' & 'fast', $p=.87$; in N2, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p=.36$). Within each driving task level, there were significant differences between each pair of n-back conditions except pair N0 and N1 in 'base' (in 'base', N0 & N1, $p=.31$; N0 & N2, $p<.01$; N1 & N2, $p<.01$; in 'slow', N0 & N1, $p<.001$; N0 & N2, $p<.001$; N1 & N2, $p<.05$; in 'fast', N0 & N1, $p<.001$; N0 & N2, $p<.001$; N1 & N2, $p<.001$).

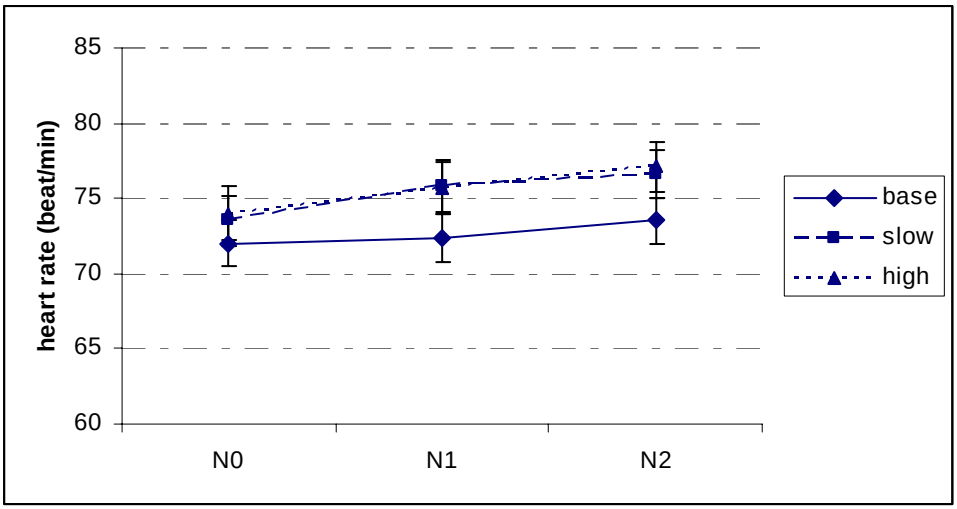


Figure 4. 7 The heart rate in different task conditions averaged over participants (N=24). The error bar shows the standard error.

The LF/HF ratio in HRV (Figure 4.8)showed a significant negative correlation with the workload in both dimensions (for driving task load, $F(2, 24)=3.9$, $p<.05$;

for working memory load, F(2,24)=8.1, p<.01). There was no significant interaction between driving task and n-back task on LF/HF (F (4, 24)=.5, p=.71). Post hoc test indicated that within each n-back level, significant differences were found only between 'base' and 'fast' in N0 condition (p<.05), 'slow' and 'fast' in N1 condition (p<.05). Within each driving task level, significant differences were found only between N0 and N3 in 'base' condition (p<.01), N1 and N2 in 'slow' condition (p<.05).

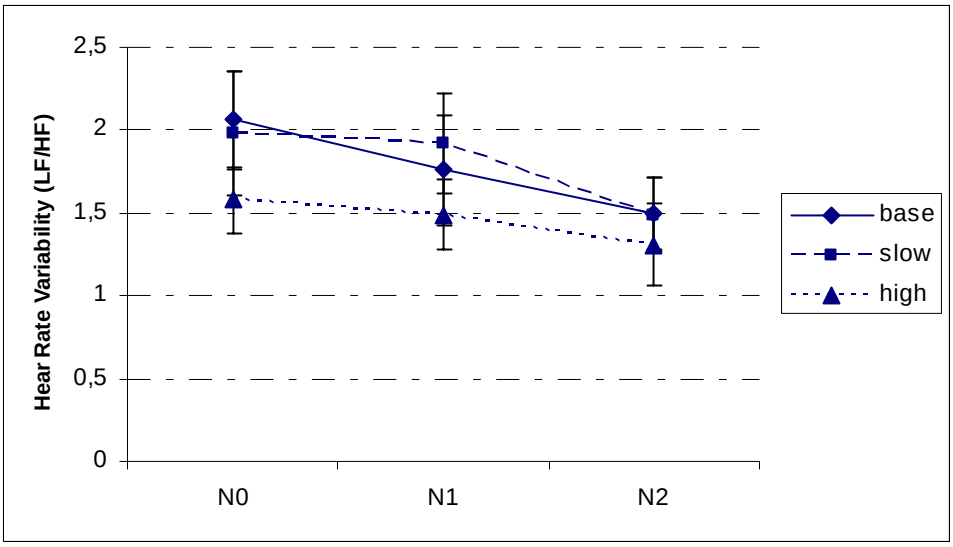


Figure 4. 8 The heart rate variability (LF/HF) in different task conditions averaged over participants (N=24). The error bar shows the standard error.

4.3.4 General modulation of the EEG parameters

The topographies of theta power in different task conditions are shown in Figure 4.9. As show in this figure, when the working memory load increased, the theta power increased in frontal area, especially at electrode sides Fz. Meanwhile, the red masked area became larger when the working memory load increased from N0 to N2, which demonstrated that the increased working memory load activated larger area in frontal brain. In contrast, the variation of driving task load did not elicit explicit changes in the frontal theta power. Figure 4.10 shows the

topographies of alpha power in different task conditions. As show in this figure, the alpha power decreased in parietal and occipital area as the task load increased in both working memory load and driving task load dimensions. However, it seemed that the variation of driving task load produced more differences in the parietal alpha power comparing with working memory load.

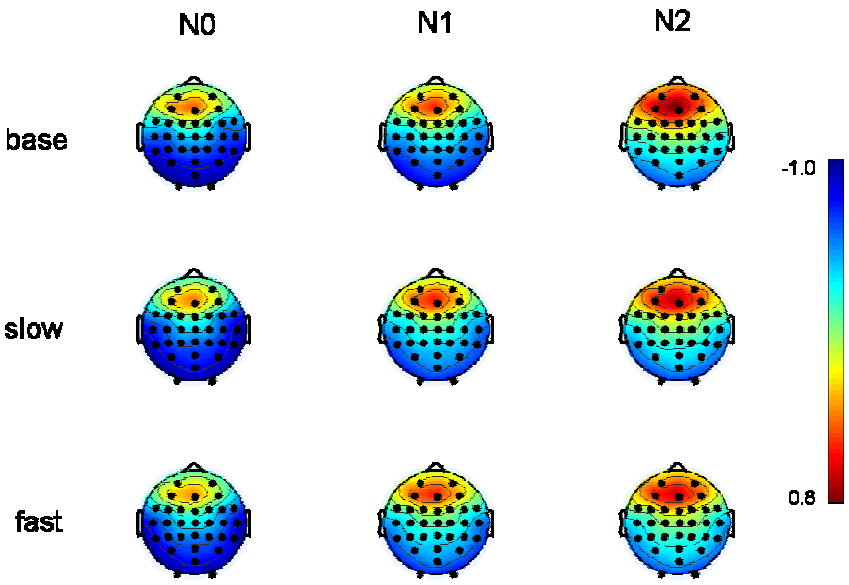


Figure 4. 9 The topographies of the theta power (z-scores) in different task conditions averaged over participants (N=24).

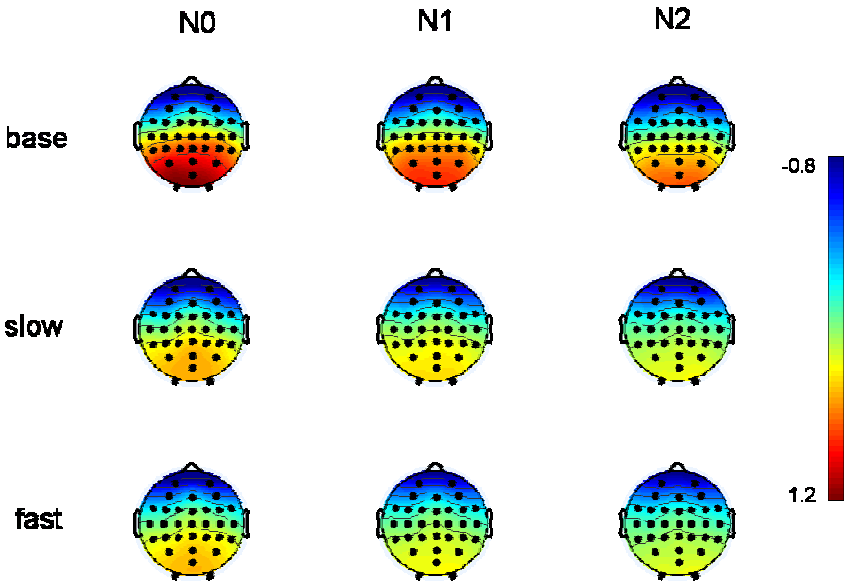
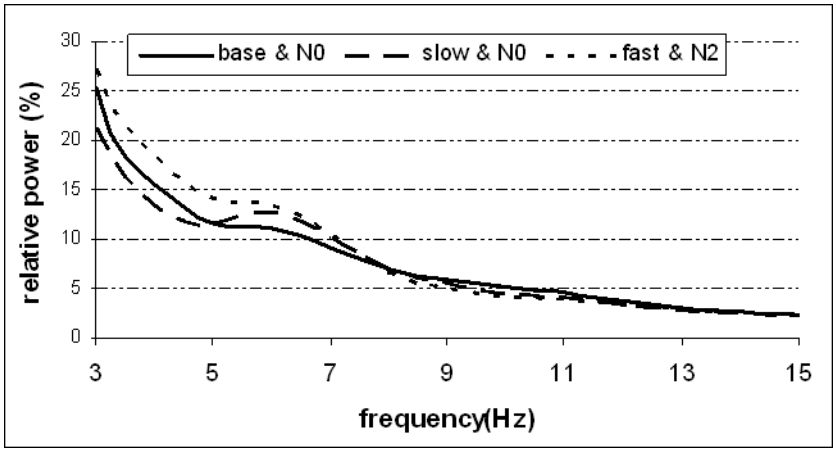
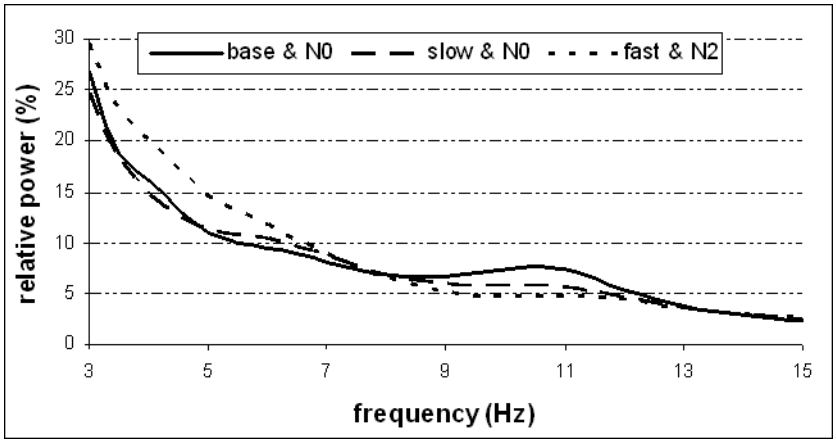


Figure 4. 10 The topographies of the alpha power (z-scores) in different task conditions averaged over participants (N=24).

Figures 4.11 (a) and (b) demonstrate the relative power spectrum density of 3 task conditions at electrode sites Fz and Pz, respectively. To illustrate a general pattern of EEG spectrum changes with task load, only 3 conditions (base & N0, slow & N0, and fast & N2, representing low, moderate and high task load levels respectively) were included in the figure. Generally, there was tendency in the power of the frontal theta frequency range to increase whenever the task load was increased (Figure 4.11(a)), whereas the parietal alpha activity strongly attenuated as the task load was increased (Figure 4.11(b)).



(a) Relative power spectrum at Fz



(b) Relative power spectrum at Pz

Figure 4. 11 Relative power spectrum density (the percentages to the total power of 4-30 Hz) (a) at frontal recording site (Fz); and (b) parietal recording site (Pz) averaged over participants (N=24).

4.3.5 Short-term modulation of the EEG parameters

Distributions the short-term parameters

Figure 4.12 and 4.13 demonstrate some examples of the statistical properties of the short-term theta and alpha power modulation respectively from subject 10. To illustrate explicit differences in the statistic properties of these two parameters, only the data from two task conditions, low task load ('base' & n0) and high task load ('high' & N2), were demonstrated here. As shown in Figure 4.12, the theta powers induced by high task load distributed in the higher value region compared to those induced by the low task load. However, the alpha powers showed an inverse tendency (Figure 4.13).

Moreover, it is interesting to note that the two curves of theta powers seemed parallel (Figure 4.12 (c)). However, the shape of alpha curve in high task load condition (Figure 4.13 (c)) became sharper and narrower compared to the low task load condition. This indicated that the alpha powers in the high task load conditions were more centre-distributed and had a smaller variance. The box plots of the theta and alpha powers (Figure 4.12 (c) and 4.13 (c)) also showed the same phenomenon. The box plot of alpha powers showed that the degree of dispersion in this parameter was lower in the high task load compared to the low task load. However, the dispersion of the theta powers demonstrated no significant difference.

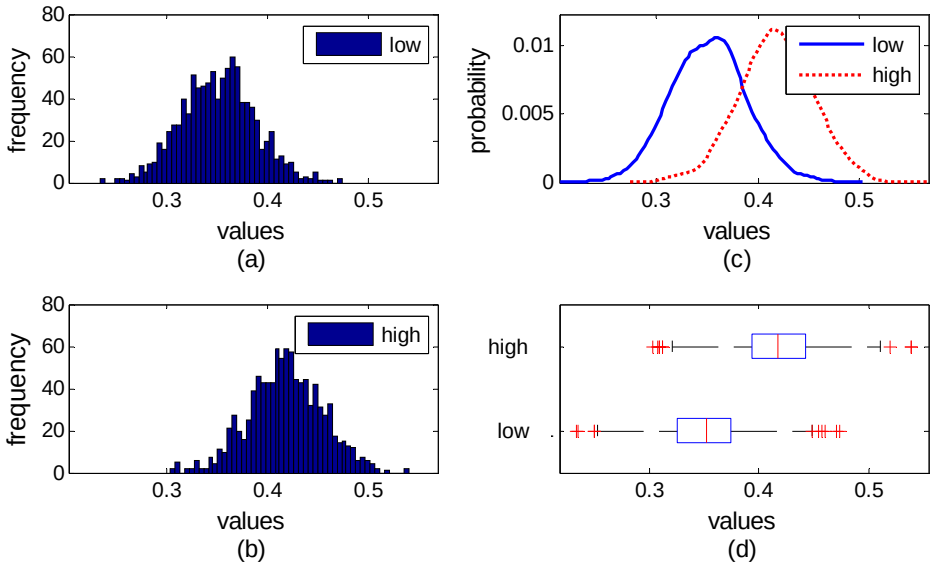


Figure 4. 12 Statistical properties of short-term theta powers in low and high task load conditions for subject 10. (a) Histogram of the theta powers in low task load condition (b) histogram of the theta power in high task load condition. (c) Comparison of the regressed probability distributions of theta powers in the low and high task load conditions (d) Boxplot of the theta powers in the low and high task load conditions.

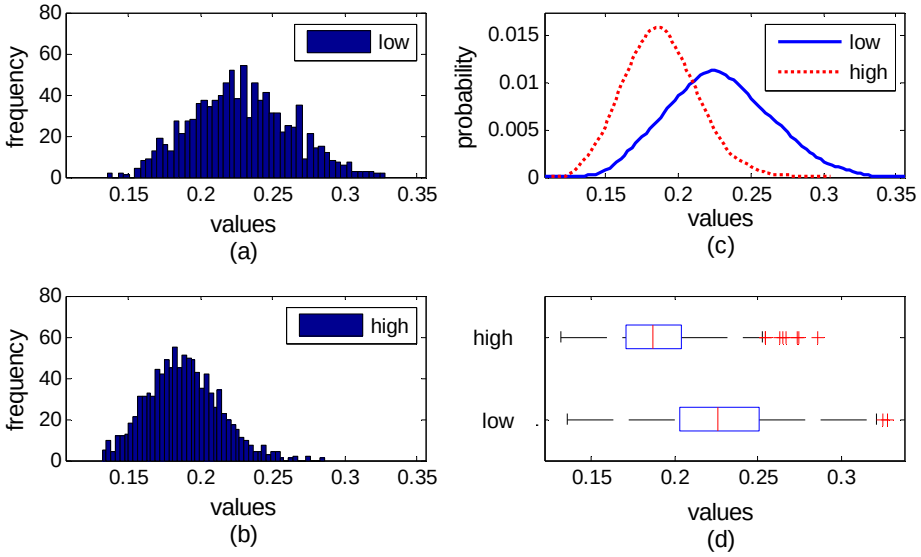


Figure 4. 13 Statistical properties of short-term alpha powers in low and high task load conditions for subject 10. (a) histogram of the alpha powers in low task load condition (b) histogram of alpha powers in high task load condition. (c) comparison of the regressed probability distributions of alpha powers in the low and high task load conditions (d) boxplot of the alpha powers in the low and high task load conditions.

Means of the short-term parameters

The results for fro-theta are shown in Figure 4.14. The variation of driving task load elicited no significant differences in the fro-theta ($F(2, 24)=2.4$, $p=.10$). However, the changes in working memory load produced significant increases in the fro-theta ($F(2,24)=23.6$, $p<.001$) with significant interaction between them ($F(4, 24)=4.2$, $p<.01$). Post-hoc test indicated that, for each driving task level, significant differences between each pair of n-back conditions were found in the 'base' and 'slow' condition but not in the 'fast' condition (in 'base', N0 & N1, $p<.001$; N0 & N2, $p<.001$; N1 & N2, $p=.45$; in 'slow', N0 & N1, $p<.01$; N0 & N2, $p<.001$; N1 & N2, $p<.05$; in 'fast', N0 & N1, $p=.08$; N0 & N2, $p=.06$; N1 & N2, $p=.15$).

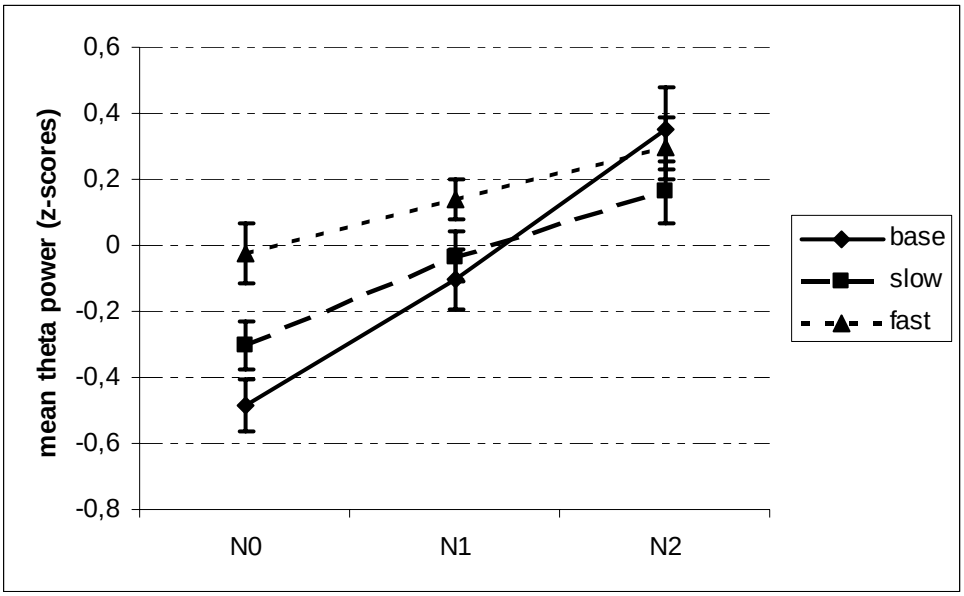


Figure 4. 14 Mean z-scores of frontal theta (4-8Hz) for each of 9 task conditions. Z-scores were converted for each participant and averaged over 24 participants (N=24). The error bars show standard errors of the z-scores.

As shown in Figure 4.15, there were significant decreases in z-scores of relative par-alpha whenever workload is increased with driving task load ($F(2, 24)=45.9$, $p<.001$) and working memory load ($F(2,24)=18.5$, $p<.001$) with significant

interaction effect (F(4, 24)=3.3, p<.05). Post-hoc test indicated that within each n-back level, there were significant differences between each pair of driving task load conditions except the pair between 'slow' & 'fast' in N1 condition (in N0, 'base' & 'slow', p<.001; 'base' & 'fast', p<.001; 'slow' & 'fast', p<.01; in N1, 'base' & 'slow', p<.001; 'base' & 'fast', p<.001; 'slow' & 'fast', p=.14; in N2, 'base' & 'slow', p<.001; 'base' & 'fast', p<.001; 'slow' & 'fast', p<.05). Within each driving task level, there were significant differences between each pair of n-back conditions except pairing between N1 & N2 in 'base' and 'slow' conditions (in 'base', N0 & N1, p<.001; N0 & N2, p<.001; N1 & N2, p=.45; in 'slow', N0 & N1, p<.01; N0 & N2, p<.01; N1 & N2, p=.32; in 'fast', N0 & N1, p<.05; N0 & N2, p<.01; N1 & N2, p<.05).

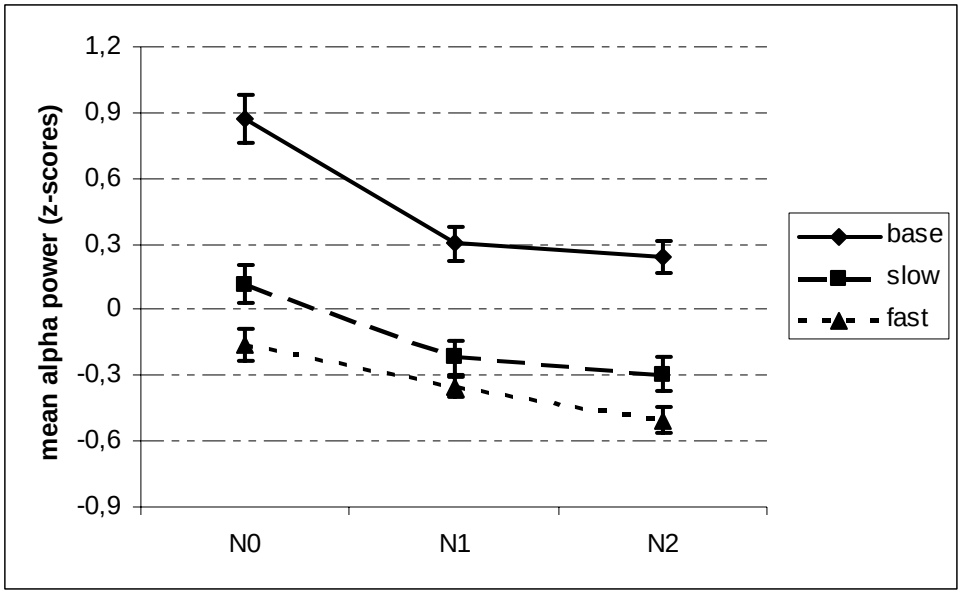


Figure 4. 15 Mean z-scores of parietal alpha (8-12Hz) for each of 9 task conditions. Z-scores were converted for each participant and averaged over 24 participants (N=24). The error bars show standard errors of the z-scores.

Standard deviations of the short-term parameters

The results of the mean SD of short-term theta powers are shown in Figure 4.16. Two-way ANOVA showed that there were no significant difference in SD of theta

powers with either driving task load (F(2, 24)=0.69, p=.50) or working memory load (F(2,24)=0.80, p=0.45). There were also no significant interaction effect (F(4, 24)=2.37, p=.06).

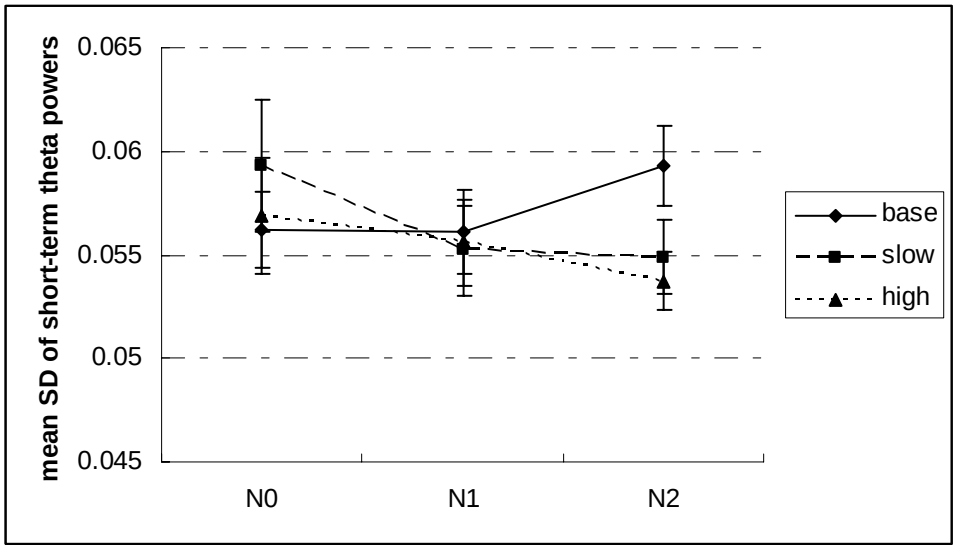


Figure 4. 16 Mean standard deviations of the short-term theta powers in different task conditions averaged over participants (N=24). The error bars show standard errors of this parameter.

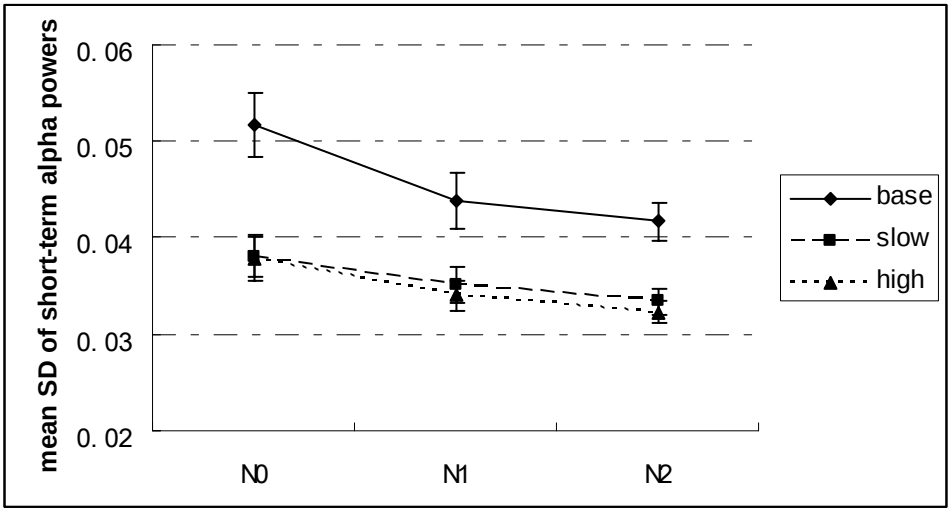


Figure 4. 17 Mean standard deviations of the short-term alpha powers in different task conditions averaged over participants (N=24). The error bars show standard errors of this parameter.

The results of the mean SD of short-term alpha powers are shown in Figure 4.17. Statistical analysis showed that there were significant decreases in SD of alpha powers with either driving task load ($F(2, 24)=23.5$, $p<.001$) or working memory load ($F(2,24)=16.3$, $p<0.001$). However, there were also no significant interaction effect ($F(4, 24)=1.79$, $p=.14$). Within the 'base' and 'fast' driving condition, there were significant differences between pairing N0 & N1, N0 & N2, but no significant differences between paring N1 & N2 (in 'base', N0 & N1, $p<.01$; N0 & N2, $p<.01$; N1 & N2, $p=.87$; in 'fast', N0 & N1, $p<.05$; N0 & N2, $p<.05$; N1 & N2, $p=.45$). However, no differences can be found between each n-back pairing in the 'slow' driving condition (N0 & N1, $p<.22$; N0 & N2, $p=.06$; N1 & N2, $p=.35$). Within each n-back level, there were significant differences between each pair of 'base' & 'slow', 'base' & 'fast', but no differences between 'slow' & 'fast' (in N0, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p=1.0$; in N1, 'base' & 'slow', $p<.05$; 'base' & 'fast', $p<.01$; 'slow' & 'fast', $p=1.0$; in N2, 'base' & 'slow', $p<.001$; 'base' & 'fast', $p<.001$; 'slow' & 'fast', $p=.88$).

4.3.6 Correlation of EEG parameters to other variables

The correlations of EEG parameters to other parameters were also investigated with the pearson correlation coefficient. As shown in Table 4.1, fro-theta was significantly correlated to the subjective load and HRV. However, par-alpha was significantly correlated to the subjective load, n-back error rate, HR, and HRV. Interestingly, par-alpha showed a higher correlation to HR than HRV, whereas for-theta showed a higher correlation to HRV than HR.

Table 4. 1 Correlation (Pearson r) of the EEG parameters to other variables

Variables	*Subjective Load*	*n-back error rate*	*n-back R.T*	*LCT M.D.*	*HR.*	*HRV (LF/HF)*
fro-theta	r=0.83	r=0.63	r=0.70	r =0.66	r =0.64	r=-0.91
	p<0.01	p=.18	p=.12	p=.15	p=.06	p<.001
par-alpha	r=-0.80	r=-0.81	r=-0.76	r=-0.69	r=-0.91	r=0.69
	p<.01	p<.05	p=.07	p=.12	p<.001	p<.05

Note: significant when p<0.05.

4.4 Discussion

In the present study, the task loads were manipulated by varying the demands imposed to participants who were asked to perform dual-task in a simulated driving environment. Both objective and subjective measurements indicated that the task manipulations increased the workload required for task performance. However, these measurements demonstrated separate sensitivities for the workload representation. In the following section, some of these results would be discussed.

4.4.1 Modulation of theta and alpha power with workload

The results of this study indicated that the increased working memory load lead to increased theta power. This finding is consistent with the results presented in Chapter 3 as well as prior literatures (Yamamoto & Matsuoka, 1990; Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998). The alpha power was observed to decline with increased task load in both dimensions, which is highly consistent with prior concept about the relation between alpha power and task load (Gevins et al., 1997, 1998; Smith et al., 2001; Wilson, 2002b; Hagemann, 2008). Specifically, the variability (standard deviation) of the short-term alpha power was observed to decrease as the task load increased. This point has been rarely concerned in prior literatures, though most of them reveal the suppression

of alpha in high task demand. This new finding would be beneficial for providing insight of the modulation of alpha rhythm induced by the variation of workload.

Interestingly, the combination of the two tasks led to a significant interaction effect on both frontal theta and parietal alpha activity. As shown in Figures 4.14 and 4.15, the degree of changes in fro-theta and par-alpha decreased with the combined task load. A reasonable assumption for this phenomenon is that both theta and alpha power have a non-linear relation with a decreased slope due to a workload increase. Such a concept is consistent with a few prior studies. Sterman & Mann (1995) reported graded suppressions in alpha power as the U.S. air force pilots flew progressively more difficult in-flight refuelling missions in a B2 aircraft simulator. However, as far as we know, the current research has no interpretation for this phenomenon.

Task-related differences were also found in this study. Statistical results show that the increase of working memory load elicited significant differences in theta power more than the driving task load. However, the increased driving task load produced more differences in the alpha power than the working memory load. Particularly, the increased driving task load did not induce significant changes in the frontal theta power. Such observation appeared also in prior studies. Wilson and Russel (2003a) investigated the changes of EEG activity in various flight conditions. Compared with the pre-flight baseline, the alpha band power significantly decreased over the parietal scalp during the flight tasks. However, only few experiment segments showed increased theta band activity at few scattered electrode sites and no consistent pattern was evident.

The task-related differences suggest that the theta and alpha activity may reflect separate neural resources utilization. Gevins et al. (1997) pointed out that task-related modulation of the ongoing EEG reflects changes in the state of the

functional networks underlying task performance. Different task may recruit separate neural resource linking with different brain regions. Metabolic studies suggest that working memory involves a functional network linking regions of the prefrontal cortex with posterior association cortices (Frisk and Milner, 1990; Gevins and Smith, 2000). On the other hand, brain imaging studies concerning the neural correlation of driving revealed the activation of cerebellar and occipital areas related to visuomotor integration (Walter et al., 2001; Calhoun et al., 2002). These findings can explain the task dependent modulation of the EEG spectrum parameters.

Another issue is the subject-related difference. Although, generally, a significant increase in theta power and a steady decrease of alpha power can be found, different subjects showed different sensitivities in these two variables. For instance, a small number of subjects showed systematic changes in frontal theta power but no systematic changes in parietal alpha power when the task load changes, and vice versa. The subject-related differences were reported also in prior studies (Gevins & Smith, 2000; Matthews et al., 2007). Gevins and Smith (2000) pointed out that the individual differences in the EEG activities, in term of either regionalization or amplitude, may reflect the individual differences in the cognitive ability and cognitive style. For instance, they stated that high-ability subjects developed strategies that made relatively greater use of parietal regions, whereas low-ability subjects relied more exclusively on frontal regions.

The individual- and task-related differences pose a big challenge to develop a general workload assessment metrics across tasks as well as individuals using the EEG parameters. These differences suggest that neither single theta power nor alpha could completely explain the variation of the human mental workload states. Therefore, a metrics, which can customize the individual differences and allow a rational combination of these EEG parameters, would be helpful for the workload

assessment.

4.4.2 Other variables and their correlations to EEG parameters

Subjective ratings increased with both workload dimensions and showed significant differences in each pairwise comparison. This indicates that the manipulation of task loads was successful. The detriments of performance, such as an increased LCT mean deviation and an n-back error rate and delayed n-back reaction time, are consistent with Wickens' multiple resource theory (Wickens, 2002). Wickens suggested that the attention resources were limited and that the resource structure can be described by four different dichotomies: two states of processing (perceptual-central and response), two modalities of perception (auditory and visual), two codes of processing (spatial and verbal), two aspects of visual processing (focal and ambient). Tasks involving the shared resources might lead to a decline in task performance when the resources are not sufficient to meet task demands. In the present study, both tasks involve the common visual perception modality. The dual task elicited competence in the visual resources, which produced impairment in the performance of both tasks.

The results of HR reinforce previous work indicating that the HR can be a sensitive workload in the driving environment (Lenneman & Backs, 2009; Mehler et al., 2009). The finding that HRV LF/HF decreased with task load is also consistent with prior studies (Kamada et al., 1992; Wilson, 2002b; Murai et al., 2004). Interestingly, the simple HR seems to demonstrate a more robust sensitivity than the complex HRV index. And this presumption is also supported by other studies that the HRV was not as sensitive to the varied cognitive demands of flight as other physiological variables such as HR and EDA (Veltman & Gaillard, 1996; Wilson, 2002b).

It is also interesting to compare the robustness of the ECG and EEG parameters. Generally, the sensitivity of EEG parameters was higher than HR and HRV to the index workload. Both fro-theta and par-alpha demonstrated significant differences for each pairwise comparison with the working memory load and driving task load respectively. However, neither HR nor HRV could completely distinguish three driving task conditions in pairwise comparisons. This robustness of the EEG observation is also supported by Brookings et al. (1996). However, the sensitivity of these physiological parameters seems task-dependent while each of the physiological parameters provides unique information concerning the cognitive load (Wilson, 2003). For example, in the present study, HR could distinguish all three n-back conditions, whereas fro-theta was not as sensitive in distinguishing driving conditions. Dolye et al. (2009) also suggested that EEG measures exhibit less sensitivity in distinguishing the cognitive load than HR during a satellite management decision-training task.

The finding that both theta and alpha are highly correlated to the subjective load emphasizes the fact that the EEG parameters can be used to represent the workload. What is interesting and surprising is that fro-theta demonstrates a higher correlation with HRV while alpha power demonstrates a higher correlation with HR. Unfortunately, as far as we know, the correlations of the EEG parameters to ECG parameters have not been systematically investigated.

4.5 Summary

This chapter aimed to investigate the changes of EEG band powers with task load in the case that the driver mental workload is simultaneously contributed by multiple factors. The experiment involved a two-factor 3*3 task design in which the LCT and a working memory task, the n-back task, were used. The task load

levels were manipulated in two dimensions, the driving task load and working memory load, with each containing three task load conditions. Except EEG signal, other variables including the subjective load, ECG, reaction time, and n-back accuracy were also recorded. Generally, the results indicated that frontal theta activity increased and parietal alpha activity decreased as the task load increased. Such results are consistent with previous studies. However, the different tasks contributed differently, such as driving task load contributed more to the changes in alpha power, whereas the working memory load contributed more to the changes in theta power, were also found. Such results suggest that different band powers may reflect separate neural resource utilization and the single EEG parameter, either theta or alpha could not completely represent the workload. Additionally, a new finding that the variation of the short-term alpha power was decreased with increased task load was also presented. Comparison of the EEG parameters with other variables indicated that, generally, the sensitivity of EEG parameters was higher than HR and HRV to index workload in the present study. However, it is suggested that the sensitivity of these physiological parameters is task-dependent with each of the physiological parameters providing unique information of the cognitive load.

Chapter 5. A computational model for online workload quantification

5.1 Motivation

A model allowing for workload assessment across tasks and individuals would be beneficial for integrating the EEG technology into dynamic human-machine systems. The results from the first and second experiments indicated that variation of task load leads to modulations of EEG theta and alpha rhythms. Particularly, Chapter 4 systematically investigated the short-term modulations of the theta and alpha powers with task load, which demonstrated the feasibility of online workload detection of workload using only the short pieces of EEG data (e.g 10s or even shorter). However, Chapter 4 addressed also several problems using the short-term parameters for the workload quantification. First of all, there are task as well as individual dependent differences in the theta and alpha powers. For instance, the working memory load contributes more to the changes of the theta power, whereas driving task load contributes more to the changes of the alpha power. Furthermore, some participants demonstrated significant changes in theta power but no changes in alpha power when task load changed, and vice versa. This suggests that either single theta power or alpha power could not completely represent the workload states. Secondly, even when the values of theta and alpha powers were customized using Z-scores, there were still huge individual variations, i.e. each participant may show different theta and alpha powers. This makes it difficult to define general thresholds or 'redlines' for adaptive aiding design.

In this chapter, we focus on the development of a workload quantification model aiming to solve the problems mentioned above. Firstly, a method using

p-quantiles (the points such that the cumulative probability that the random variable is less than these points is at most p) is promoted to generalize the individual differences. Secondly, for combining the theta power and alpha power, a logistic function model (LFM) was proposed for online workload assessment. In the following section, the details of these two methods are introduced.

5.2 P-quantile as a generalization method

We propose to use p-quantile as a method to generalize the individual differences. In statistics, q-quantiles (q is an integer) are often used to regulate intervals from the cumulative distribution function of a random variable. The motivation of q-quantiles is to divide ordered data into q essentially equal-sized data subsets and the quantiles are the data values marking the boundaries between consecutive subsets. For instance, 4-quantiles (also call quartiles) are those three points that divide a data set into four equal-sized groups with each representing a fourth of the distributed sampled population. Instead of using integers q, the p-quantiles are based on a real number p with $0 < p < 1$, illustrating the points such that the cumulative probability that the random variable is less than these points is at most p. They are useful measures because they are less susceptible to long-tailed distributions and outliers.

For illustrating the logic of using the p-quantile as a generalization method of EEG parameters, it is necessary to start with data from the second experiment. We firstly regressed the *probability distribution functions* of the short-term (10s with 50% overlay, see Chapter 4) theta and alpha powers in different task load conditions (here, for an explicit illustration, we used only three conditions 'base & N0', 'slow & N1', and 'fast & N2', representing the low, moderate, and high task load conditions). Meanwhile, we randomly selected the same size of the data points from each task condition for both theta and alpha powers. We used the all

selected data points to build a *general distribution function* and also an *empirical cumulative distribution function* for theta and alpha power respectively. This equal sized selection of data guaranteed that the general distribution model contains information from all task conditions without being dominated by certain task load condition.

Figure 5.1-5.4 show some examples of matching the task load specified theta or alpha powers distribution functions with the general distribution function and cumulative probability function. Subplot (a) shows the comparison of the task load induced short-term theta or alpha power probability distribution function; subplot (b) shows the general distribution function including all task conditions; subplot (c) shows the cumulative distribution function of the general distribution. As shown in subplot (a), when the task load increased from the lowest to highest level, the theta power (Figure 5.1 and 5.2) distribution functions moved from the left side (low value) to the right side (high value), whereas alpha power distributions (Figure 5.3 and 5.4) moved from the right side (high value) to the left side (low value). When matching 0.5-quantiles (medians) of the distribution function in task load condition with the general distribution model, the locations of these quantiles demonstrated a consistent pattern for individuals, even though the theta power and alpha power involved large inter-subject differences. As shown in Figure 5.1 and Figure 5.2, the medians of the theta power in the low, moderate, and high task load level for subject 4 were 0.356, 0.405, and 0.437 respectively. But for subject 9, these medians were 0.257, 0.298, and 0.327 respectively. Apparently, it is not appropriate to compare the workload between the subjects 4 and 9 using these data. However, the distributions of these medians in the general distribution function showed a quite similar pattern. The medians in low task load located in the left side; the medians of the moderate task load located in the middle; and the medians of high task load located in the right side.

The alpha power demonstrated the inverse trends with theta power which are shown in Figure 5.3 and 5.4. Although the medians of the alpha power in the low, moderate, and high task load level for subject 4 (0.264, 0.216, and 0.175 respectively) and for subject 9 (0.306, 0.269, and 0.237 respectively) were quite different, a quite similar pattern of the distribution of these medians of alpha power in the general distribution function can be found. The medians in low task load located in the left side; the medians of the moderate task load located in the middle; and the medians of high task load located in the right side.

These patterns could be described using p-quantiles in the cumulative distribution function. For instance, the medians of theta power in low, moderate and high task for subject 4 corresponded to 0.224-, 0.606-, and 0.766-quantiles respectively (see, Figure 5.1 (c)), whereas, for subject 9, they were correspondent with 0.166-, 0.482-, and 0.71-quantiles respectively (see, Figure 5.2(c)). Similarly, the medians of the alpha power in different task load conditions were responding to the p-quantiles in the cumulative distribution function, but showed an inverse relationship, e.g. medians in load task load matched high p values, whereas medians in high task load matched low p values (see, Figure 5.3 (c) and Figure 5.4 (c)).

Therefore, the analysis the p-quantile indicated that there were some general patterns of modulation of theta and alpha power induced by the task load, when matching the theta and alpha powers induced by specific task load levels with the general distribution function representing the full range task loads. This provides a solution for customizing the individual differences, which would be valuable for the development of a general model for workload quantification using the EEG parameters.

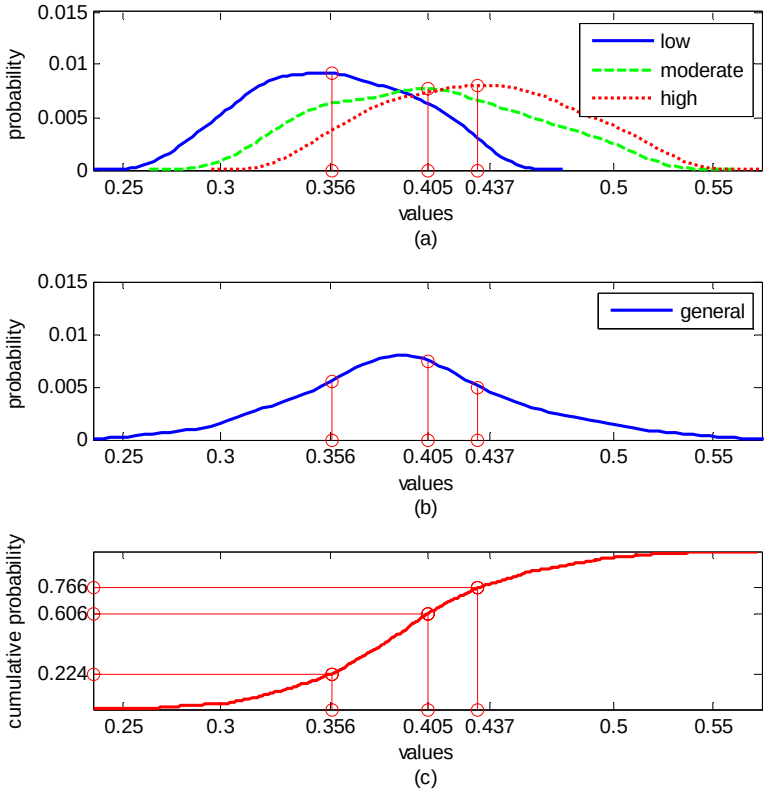


Figure 5. 1 Distribution functions of theta powers for subject 4

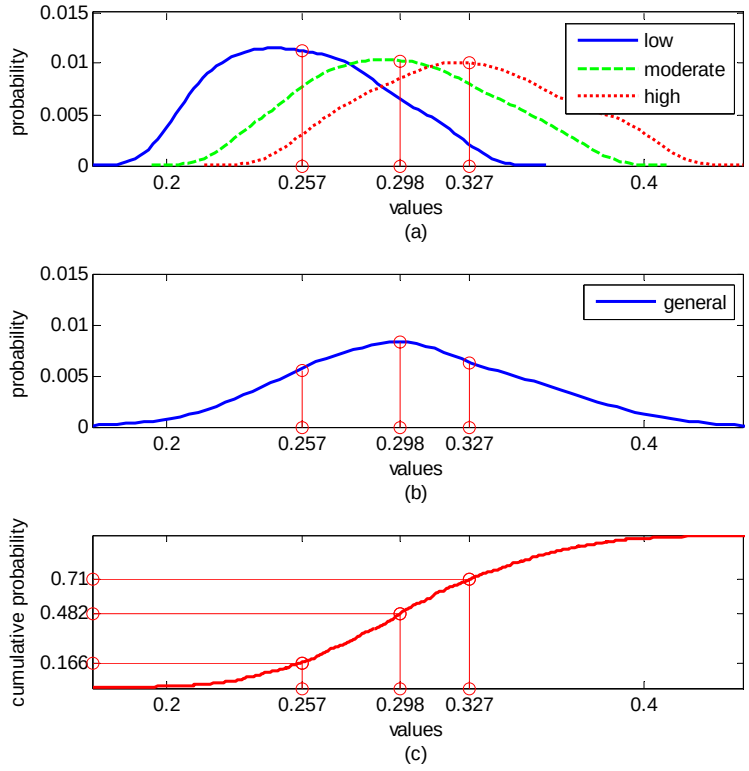


Figure 5. 2 Distribution functions of theta powers of subject 9

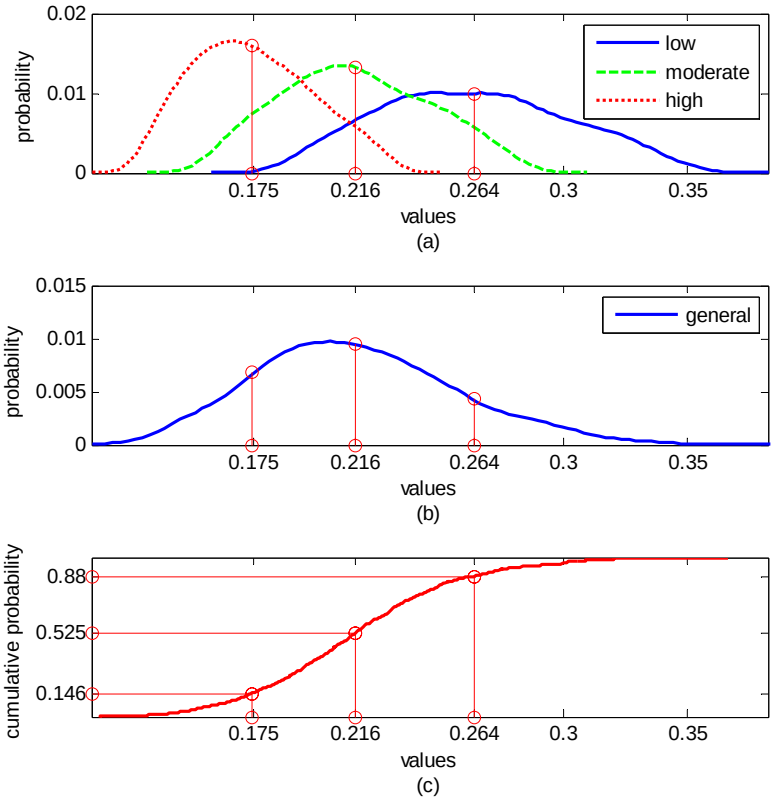


Figure 5. 3 Distribution functions of alpha powers for subject 4

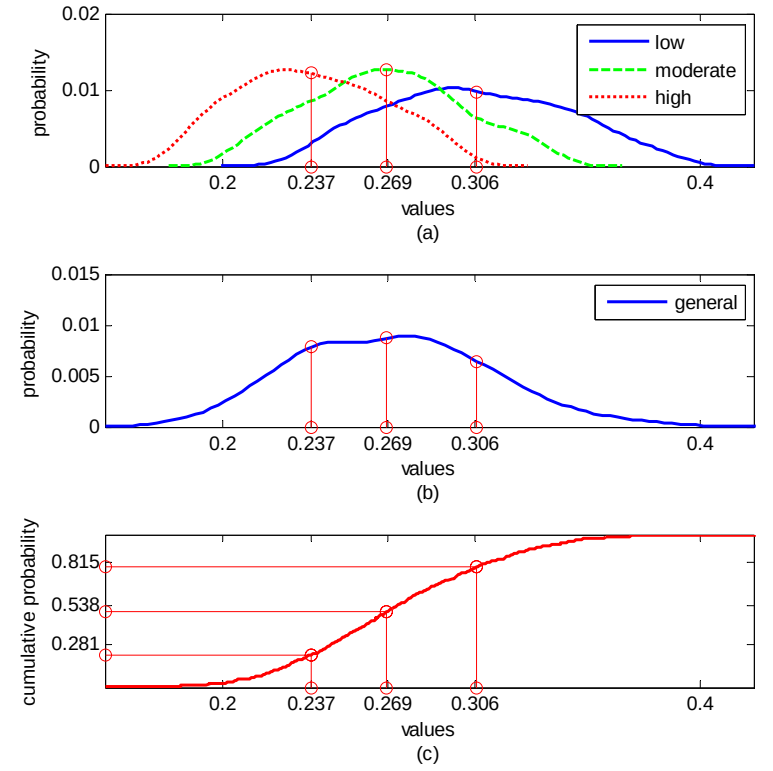


Figure 5. 4 Distribution functions of alpha powers of subject 22

The averaged p-values and standard deviations for each task condition across all subjects are shown in Table 5.1. The p values of the relative theta powers increased with task load, which is consistent with the previous finding that there is a proportional relation between theta activity and task load. For instance, the mean p values of theta band power in the lowest task load condition and highest task load condition are 0.355-quantile and 0.595-quantile respectively. As expected, the p-value of the normalized parietal alpha decreased from 0.768 to 0.373 when the task load increased from the lowest to the highest. These tendencies are consistent with the general pattern how theta and alpha change with the task load.

Table 5. 1 The averaged p values (standard deviation across subjects) for each task condition

Variables		Theta		Alpha	
		Mean	SD	Mean	SD
'base'	N0	0.355	0.112	0.768	0.101
	N1	0.467	0.132	0.642	0.117
	N2	0.599	0.160	0.608	0.121
slow	N0	0.422	0.113	0.556	0.121
	N1	0.505	0.116	0.468	0.120
	N2	0.565	0.139	0.428	0.110
fast	N0	0.501	0.139	0.476	0.130
	N1	0.550	0.090	0.424	0.065
	N2	0.595	0.123	0.373	0.086

In order to validate the normalization effect of the p-value, we extracted another two variables: the mean values and z-scores. The mean values of theta and alpha power were calculated by directly averaging the short-term theta and alpha over task conditions and participants. The z-scores were derived by subtracting the mean values of the population (over data from all conditions within each participant) and then dividing the standard deviation (over data from all conditions within each participant). The z-scores were also averaged over task conditions for each participant.

The normalization effect of the means, z-scores, and p values of the EEG parameters for workload presentation is shown in Figure 5.5. This Figure generally demonstrates the necessity of a normalization method. Figure 5.5 (a) shows the means of theta and alpha power in the low (base & N0) and high task load (fast & N2) conditions. As shown in this Figure, there are overlays in the data points of means of theta and alpha power and low and high task load condition could not be separated using these two parameters. However, after the normalization, both z-scores (Figure 5.5 b) and p values (Figure 5.5 c) could separate the low and high task conditions. This indicates that using the EEG parameters to represent the workload levels, it is necessary to normalize the data and both z-scores and p values have good normalization effect.

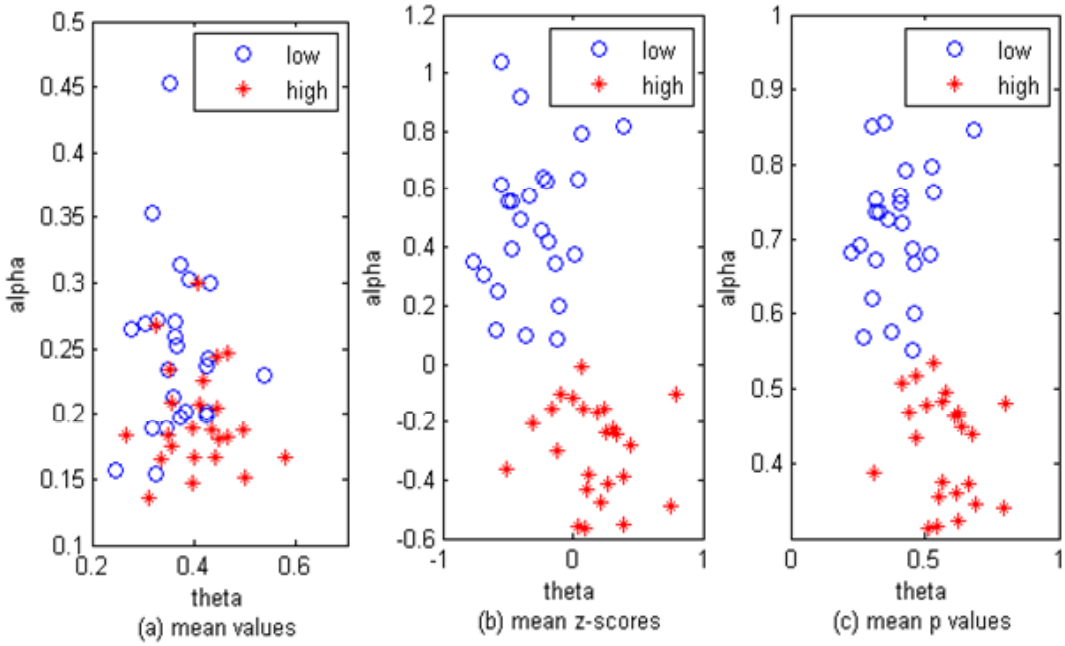


Figure 5. 5 Comparison of the individual variation of the mean values, mean z-scores, and mean p values with each containing 24 data points (24 participants) for theta and alpha.

For continuous comparison of normalization effect between the z-scores and p values, we used the absolute coefficient of variation (defined as the absolute value of the ratio of standard deviation to the mean value) to compare the dispersion of

z-scores and p-values. The coefficient of variation is a normalized measure of dispersion of a probability distribution. The higher coefficient indicates the population is more dispersed. Table 5.2 shows the coefficients of variation of z-scores and p-values in different task conditions for theta and alpha power. As shown in this table, the p-values of both theta and alpha power showed narrow coefficients of variation in all task conditions compared to z-scores. For instance, in condition 'base' & N0, the coefficient of variation across subjects for the theta p-value is 0.314 comparing with 0.809 for the z-scores; in the 'fast' & N2, this coefficient for p-value is 0.207 comparing with 0.584 for the z-scores. Similar results are also observed for the alpha power. The decreased coefficients of variation across subjects indicated a narrow individual variation in the p-values. Therefore, the p-value might be a good normalization method for the task load induced the EEG spectrum parameters. It could be helpful for establishing a generalized workload monitoring metrics using EEG spectrum parameters.

Table 5. 2 Comparison of the coefficients of variation between the z-scores and p-values

Variables		**theta**		**alpha**	
		z-scores	p-values	z-scores	p-values
'base'	N0	0.809	0.314	0.615	0.131
	N1	4.392	0.282	1.198	0.182
	N2	1.733	0.267	1.532	0.198
slow	N0	1.188	0.269	3.626	0.217
	N1	10.53	0.229	1.784	0.256
	N2	2.933	0.247	1.291	0.257
fast	N0	18.13	0.278	2.355	0.274
	N1	0.542	0.164	0.633	0.154
	N2	0.584	0.207	0.620	0.230

5.3 Logistic regression model for workload quantification

In the last section, we proposed a generalization model for workload

representation using EEG parameters using p-quantiles. It is possible to predict the workload using the p values, e.g, if the median of theta power matches the p-quantile in the general distribution function with a high p value, or the median of alpha power matches the p-quantile in the general distribution function with a low p value, it may be a high workload condition. However, as discussed in Chapter 4, either single theta or alpha power could not completely represent the workload level, and the generalization model does not solve the problem how to combine the theta power and alpha power into unified scales for workload representation, e.g. scales from 0-1, to index workload. In this section, we propose a logistic function as a transformation method to project the theta and alpha power into workload unified indexes.

5.3.1 The logistic function model (LFM) for workload quantification

Before introducing the LFM, it is necessary to explain the logistic function given by:

$$f(z) = \frac{1}{1+e^{-z}} \tag{1}$$

The equation (1) is the basic logistic function. The input is z and the output is $f(z)$. As shown in Figure 5.6, the logistic curve is a sigmoid curve. It is useful because it takes an input ranged from negative infinity to positive infinity to an output ranged between 0 and 1. In practice, it is sufficient to compute z over a small range of real numbers such as [−6, +6].

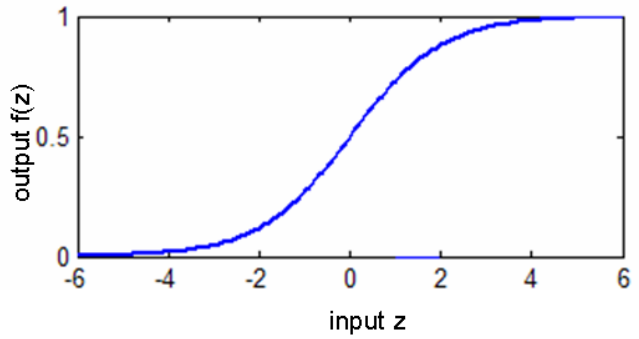


Figure 5. 6 Logistic function curve

The logistic function is itself a cumulative distribution function, which matches the theta power and alpha power cumulative distribution function. The output of the logistic function could be the p-values. The logistic function is useful in statistics for a number of reasons. Firstly, it is a cumulative distribution function, which makes it useful to simulate the cumulative distribution function of the random variables. Secondly, they are frequently used in logistic regression to model how the probability p of an event may be affected by one or more variables.

Therefore, we propose the following function for workload detection using EEG features:

$$W_{ind} = \frac{1}{1+e^{-f(x)}} \tag{2}$$

Where, W_{ind} is workload indices from 0 to 1 and $f(x)$ given by

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{3}$$

Where, $x_1, x_2, \cdots, x_n$ denote the normalized EEG features for workload indexing (e.g. theta, alpha, etc.) and β_0 is the intercept, $\beta_1, \beta_2, \cdots, \beta_n$ is the coefficient for $x_1, x_2, \cdots, x_n$ reflecting the size of the contribution of the corresponding factors to workload.

It is indeed quite reasonable to use the logistic function to index the workload. First of all, the logistic function, as shown in Figure 5.2, is itself a cumulative distribution function and the output of the logistic function could be the p values for the p-percentiles which matches perfectly the idea to use p value for workload representation. Secondly, the logistic function is often used in logistic regression to model how the probability p of an event may be affected by one or more factors, which provides a solution for combining the theta and alpha powers. Thirdly, the output of the logistic function is ranged from 0 to 1, which matches also our idea to use the scales from 0 to 1 to quantify workload levels.

5.3.2 Regression of the logistic function coefficients with p-quantiles

Defining the appropriate coefficients for the LFM is important. As presented in section 5.3, the average p-values were obtained from 9 different task conditions for both theta and alpha powers. The p-values describe how the medians of EEG parameters induced by certain task load level distribute in the general distribution functions containing information from a full range experiencing of task load.

We propose to use the subjective loads and p-quantiles to regress the coefficients of the logistic function. The subjective loads (presented in Chapter 4) were treated as the standard workload index W_{ind}, the p-quantiles of theta as the inputs of x_1, and the p-quantiles of the alpha power as input of variable x_2 in equation (3). In this way, the coefficients of the logistic function can be determined through regression of equation (2).

Here, we demonstrate an example using the data from subject 11. Table 5.3 shows the subjective loads and p-quantiles of the theta and alpha power using the p values provided in Table 5.1.

Table 5. 3 The subjective load and corresponding p-values of theta and alpha power for subject 11

Variables		Subjective load	theta	alpha
'base'	N0	0.148	0.367	0.214
	N1	0.251	0.384	0.199
	N2	0.401	0.402	0.196
slow	N0	0.193	0.376	0.192
	N1	0.398	0.390	0.185
	N2	0.567	0.398	0.180
fast	N0	0.282	0.389	0.186
	N1	0.453	0.396	0.179
	N2	0.685	0.396	0.176

The coefficients of the logistic function could be determined through regression of the data provided in Table 2

$$\beta_0 = \text{-}12.33\,,\ \beta_1 = 41.87\,,\ \beta_2 = \text{-}23.94$$

Thus, the model for workload prediction is given by

$$W_{ind} = \frac{1}{1+e^{12.33-41.87*theta+23.94*alpha}} \qquad (4)$$

The coefficients, β_1 and β_2, reflect the contribution of theta and alpha power to the workload indices. The positive β_1 denotes a proportional relation between theta power and workload, whereas β_2 is negative, representing a negative relation between alpha power and workload.

A general procedure for establishing the LFM based on continuous theta power and alpha power is described as follows:

Step 1: Establish a set of EEG parameters representing a full range of mental workload. This can be done with an experiencing of a full range of task performance scenarios.

Step 2: Estimate a general cumulative distribution function for theta and alpha power using the data from all task conditions; extract the 9 p-quantiles (shown in Table 5.1) for theta power and alpha power.

Step 3: Assume that sequences of theta percentiles and alpha percentiles represent the workload scale sequence (shown in Table 5.2); use these sequences to regress the coefficients, β_0, β_1, and β_2, in the LFM.

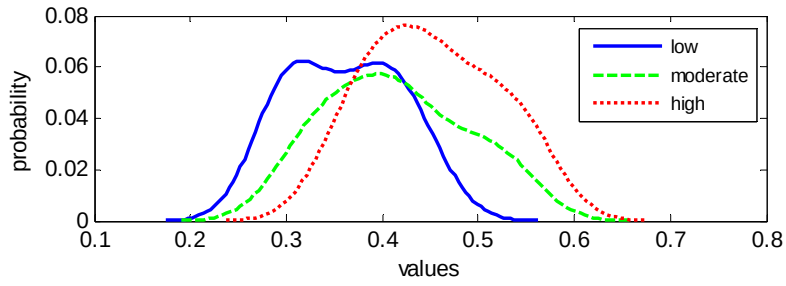
5.3.3 Results with the logistic regression model

Distribution Comparison

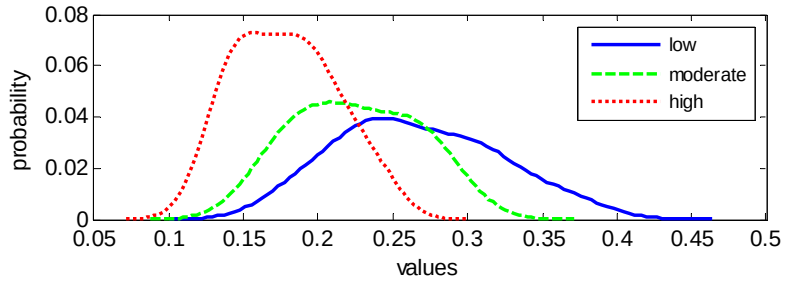
In section 5.2, we demonstrated the distribution functions of the theta and alpha powers induced by the different task load levels. Although the differences in these two parameters were statistically significant across the task loads, there were still large overlays in these distributions. Figure 5.7 shows the effect of separating workload scores induced by different task load conditions using LFM. After applying the LFM, the workload scores induced by low task load dominated in the low value area; those induced by moderate task load dominated in the moderate value area; and those induced by high task load dominated in the high value area. This matches the general expectation about a linear relation between workload and task load. Meanwhile, there were also overlays among the task conditions, which is also reasonable that there may be fluctuations in the workload states even when performing a task with constant demands. When compared to the original EEG parameter distribution functions (see, Figure 5.7 (a) and (b)), the overlay of the distributions of the workload scores (Figure 5.7 (c)) calculated using LFM was significantly decreased, which denoted that a better separation of the workload levels could be achieved using the LFM.

Mean values and the percentages for low, moderate and high workload level

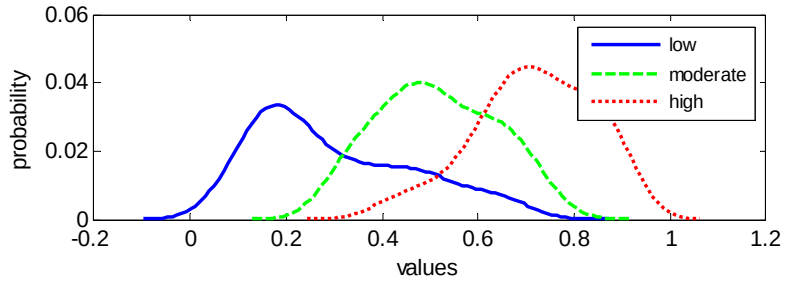
The mean values of the workload scores were calculated by the averaging across subjects and task conditions. Two thresholds 0.2 and 0.8 were used to separate the workload levels. Scores higher than 0.8, lower than 0.2 and between these two thresholds were classified as the high workload, low workload and moderate workload respectively. Table 5.4 shows the mean regressed scores and the percentages of low, moderate and high workload level which were directly classified using the regressed scores.



(a) distribution funcitons of the theta power



(b) distribution functions of the alpha power



(c) distribution functions of the workload scores using logistic function

Figure 5. 7 The comparison of the logistic function transferred workload indices with theta and alpha power in low (‘base’ & N0), moderate (’slow’ & N1) and high (‘fast’ & N2) for subject 11

As shown in Table 5.4, the mean workload scores increased from 0.26 to 0.58 when the task load levels increased from the lowest level to the highest level. This indicated that the LFM could be used for workload representation. Additionally, the percentages of the occurrences of low, moderate and high workload changed in different task load conditions. In the lowest task load condition, there were 60.6% epochs with low workload, and this percentage was reduced to 18.4% in the highest task load condition. In contrast, only 13.8% epochs with high workload in the lowest task load condition were observed comparing with an average of 47.5% in the highest task load condition. These results are consistent with the previous theory concerning the dissociation between the task demand and workload (Hancock et al., 1996). Hancock pointed out that dissociation occurs in some unusual circumstances that the experienced workload might be failed to linearly associate with external task demand.

Table 5. 4 Mean values (standard deviations) of the logistic function workload scores and the percentages of occurrences of the low, moderate and high workload in 9 task load conditions averaged over participants (N=24)

Variables		Mean workload scores	Percentages (%)		
			low	moderate	high
base	N0	0.268 (0.063)	60.6 (10.7)	25.6 (7.4)	13.8 (4.8)
	N1	0.385 (0.094)	44.8 (12.7)	32.4 (5.2)	22.8 (10.1)
	N2	0.482 (0.122)	31.6 (13.5)	32.9 (5.6)	35.5 (15.8)
slow	N0	0.384 (0.097)	40.4 (14.1)	33.4 (6.9)	26.2 (10.6)
	N1	0.481 (0.051)	29.7 (6.8)	35.9 (4.8)	34.4 (7.3)
	N2	0.552 (0.088)	21.9 (9.1)	34.1 (6.8)	44.0 (12.8)
fast	N0	0.483 (0.105)	29.4 (13.7)	35.3 (7.9)	35.3 (12.7)
	N1	0.552 (0.063)	21.9 (7.1)	38.3 (6.9)	39.8 (9.7)
	N2	0.583 (0.081)	18.4 (8.7)	34.1 (7.6)	47.5 (13.9)

Comparison with classic machine learning methods

For illustrating the merit of the LFM, we demonstrate a comparison of the traditional task load classification and LFM using data from subject 7 and subject

11, which represented two groups of task performance in this experiment. Table 5.5 shows the task performance (n-back error rate), theta, alpha power, and workload scores using logistic function for these two subjects. As expected, for subject 11 (also majority of the subjects), theta power increased and alpha power decreased as task load increased. Consequently, the workload score of subject 11 demonstrated a linear correlation to the task load, i.e. higher task load induced higher workload scores. However, the EEG spectrum components of subject 7 demonstrated an unusual case that the highest theta power and lowest alpha occurred in 'base' & N2 rather than 'fast' & N2, which indicated that this subject experienced the highest workload in 'base' & N2. The LFM indicated also that the highest workload score in 'base' & N2.

This can be interpreted with the performance data. As shown in Table 5.5, there was a significant decline in the task performance for subject 7 in the dual task condition compared with the single task condition, for instance, the error rate in DL & WH is 13.1% and this percentage sharply increased to 42.5% in 'slow' & N2, and 57.1% in 'fast' & N2, which denoted that the subject may give up task performance in 'slow' & N2 and 'fast' & N2, or at least in some periods. Thus, the workload scores in these two conditions were lower than that in 'slow' & N2. In contrast, subject 11 could maintain a certain level of the performance even when the task load was increased, for instance, 80% accuracy in 'fast' & N2. This denoted that subject 11 invested more effort for the task performance and therefore a higher workload score was observed.

For comparison, a classic machine leaning method, LDA, was used to classify the task load. We selected a 20 dimension of features consisting of theta and alpha power from 10 electrodes (5 at frontal-central area and 5 at parietal-occipital area) and used LDA to classify the task 'base' & N2 and 'fast' & N2. Our classification showed an accuracy of 75% and 83% for subject 7 and 11 respectively. However,

regarding to these classification results, a delicate discussion is definitely necessary. Although the classifier could distinguish these two task load levels with high accuracy, such classification was failed to provide important information that which condition induced higher workload. It may be straightforward assumed that the workload in 'fast' & N2 was higher than 'base' & N2. However, such an assumption was contradictory with the conclusion drawn from the quantitative analysis of the theta and alpha power that subject 7 experienced higher workload in the 'base' & N2 compared to 'fast' & N2. Another possibility is that such classification may not really distinguish the differential workload levels, but other components, e.g. there may be a spatial pattern induced by eye movement. Therefore, the machine learning methods for workload classification have 'black box' effect. The high classification accuracy may mislead to a straightforward assumption based on the predefined label. The ignoring of the internal changes of the features may result in a wrong conclusion. From this point, the LFM with a scoring scheme seems more transparent and also more convincing.

Table 5. 5 Comparison of the n-back performance, theta power, alpha power, and workload scores between subject 7 and 11

Subs.	variables	base			slow			fast		
		N0	N1	N2	N0	N1	N2	N0	N1	N2
Ss 7	n-back error rate(%)	-	6.3	13.1	-	20.6	42.5	-	51.2	57.1
	theta	0.321	0.332	0.353	0.330	0.350	0.341	0.332	0.343	0.351
	alpha	0.190	0.185	0.178	0.183	0.178	0.186	0.187	0.185	0.183
	Workload scores	0.341	0.409	0.506	0.478	0.504	0.448	0.482	0.419	0.481
Ss11	n-back error rate(%)	-	2.5	9.9	-	4.1	12.5	-	16.5	20.0
	theta	0.361	0.362	0.385	0.358	0.369	0.385	0.354	0.387	0.442
	alpha	0.211	0.204	0.189	0.201	0.192	0.185	0.193	0.179	0.166
	Workload scores	0.323	0.346	0.476	0.360	0.438	0.539	0.392	0.564	0.788

Correlation of the workload scores using LFM with other variables

Table 5.6 shows the correlation coefficients of the EEG parameters and regressed scores to other variables (SL, HR, HRV, RT and n-back error rate), which have been also reported to be sensitive to the workload variation in Chapter 4. As shown in Table 5.6, only the theta power significantly correlated to SL, HRV, and the alpha power significantly correlated to SL, HR, HRV, error rate of n-back. However, the workload score significantly correlated to all other 5 parameters. Moreover, the correlation coefficients of the workload score with other variables were higher than that of the fro-theta and par-alpha with other variables except that with HRV (the coefficient of HRV with fro-theta is 0.90, while that with regressed indices is -0.85). The correlation analysis shows again workload score using logistic function is rational for workload assessment. Furthermore, the enhanced correlation coefficients show that the LFM provides a promising solution for establishing a general workload quantification metrics.

Table 5. 6 Comparison of correlation coefficients of the EEG parameters and regressed scores to other variables (SL, HR, HRV, RT, n-back error rate). The regressed scores showed better correlations with other workload indexes comparing with the fro-theta and par-alpha.

variables	SL	HR	HRV	RT	n-back Error
theta	0.83 (p<.01)	0.63 (p=.06)	-0.90 (p<.001)	0.69 (p=.12)	0.62 (p=.12)
alpha	-0.79 (p<.01)	0.90 (p<.001)	0.69 (p<.05)	-0.76 (p=.07)	-0.80 (p<.05)
Workload scores	0.90 (p<.001)	0.90 (p<.001)	-0.85 (p<.01)	0.96 (p<.01)	0.94 (p<.01)

5.4 Discussion

The quantification of human mental workload is crucial for task analysis as well as for designing adaptive aiding strategies for human operator to improve

human-machine interaction. In this chapter, we propose to use p-quantiles for the individual difference generalization and to use a logistic function for mental workload quantification. This model projects EEG parameters (theta and alpha) into unified indexes ranged from 0 to 1 for mental workload representation. Particularly, the p-quantiles are used to determine the coefficients of the LFM, which could be customized for individuals. The results showed that such LFM owns numerous merits for the workload quantification. In the following section, some issues concerning this model are discussed.

5.4.1 P-quantiles and z-scores

In the section 5.2, we compared the dispersion of the p-values with another normalization method, Z-score, which are commonly used in the prior EEG studies (Smith et al., 2001; Thatcher et al., 2005; Fukami et al., 2010). The comparison of the dispersion of the z-scores and p-values indicated that the p-value owns a better normalization effect. Here, it is interesting to discuss the dissimilarities and similarities between p-values with Z-scores. Z-score is calculated with a standard transform, a method by dividing the standard deviation after subtracting the mean of the population. The use of Z-score often entails a predefining condition that the population is normally distributed, though this condition is often ignored in many applications. In fact, when the population is a strict normal distribution, these two methods, p-values and z-scores, are equivalent for the normalization. Unfortunately, most of the cases (shown in the Figure 5.1- 5.4 as well as in Figure 5.7), the distributions of the theta and alpha power actually are not necessarily normally distributed. For instance, the standard deviation of the short term alpha power was found to become narrow when the task load increased. This makes the distribution functions of the alpha power are more like log distribution rather than norm distribution. Therefore, use the Z-score as the normalization method would be not appropriate. In contrast, the use of

p-quantiles can be applied in any distribution function. In this perspective, the p-quantiles might be more suitable to handle the EEG parameters, which are likely to be not normal distributions.

5.4.2 Merits and demerits of the logistic function model

The results in section 5.3.3 showed that the LFM is effective for workload assessment. Such effectiveness focus on the following aspects:

First of all, the properties of the logistic function make it suitable to quantify workload. The output of the logistic function is in a range of [0 1] and it matches the idea to use the scales from 0 to 1 to quantify workload. Such unification allows an easy definition of the workload thresholds in the practice for adaptive aiding design. It enables also an easy adjustment of the thresholds in different application cases. Additionally, the logistic function enlarges the centre values and suppresses the extreme value, which offers an effective solution to maximize the differences of the overlaid distributions. Such property allows a better separation of the workload scores induced by different task loads. The comparison of the distributions in Figure 5.7 proves this effect. The overlay of the distributions of workload scores in different task conditions significantly decreased.

Secondly, the LFM provides a solution to combine the EEG parameters and enables to establish a general workload detector using EEG signals. In the present study, we showed an example only using the theta power and alpha power. It is also possible to involve other parameters. In a relative simple task, one or two parameters that are sensitive to the variation of workload could be used. In a complex situation, some other parameters such as beta power or gamma power could be also introduced in this model if these parameters are modulated by the variation of workload. Thus, this model could even assess the workload across

tasks. For instance, the subject load demonstrated equivalent scores between 'base' & N2 and 'slow' & N1 (0.402 and 0.398 respectively). The workload score using logistic function also showed this equivalence (0.482 and 0.481 respectively). Additionally, results shown in 5.3.3 demonstrated the improved correlations of the workload scores to other variables using the LFM (Table 5.6). It is well known that the different workload assessment candidates may be sensitive to separate aspects of the workload (Kramer, 1991). For instance, subjective load is sensitive to the number of the tasks even if these tasks are not difficult (Yeh and Wickens, 1988); it was suggested that HRV may be sensitivity to computational effort but not to compensatory effort (Brookhuis and De Waard, 2001). Therefore, the improvement in the correlations to these workload indicates the LFM shows a general sensitivity to workload which may be a results of different factors.

Finally, it has been confirmed that the theta power increases and alpha power decreases with the task load (Gevins et al., 1997, 1998; Smith et al., 2001). In the proposed model, a positive and negative coefficient for theta and alpha can be derived through the regression. In other words, the proposed model reflects the general tendencies in the modulation of the EEG parameters with the workload variation. This makes the LFM to produce more convincing outputs for the workload representation compared to the classic machine learning methods. As demonstrated in 5.3.3, the classification of workload states using classic machine learning methods is prone to be contaminated by other artefacts. Furthermore, such methods simply separate the features into groups but are failed to provide important information about the workload states, for example, which is higher or lower. Therefore, in this perspective, the LFM would be more reliable to reflect the mental workload states.

Nevertheless, some limitations of this model should be addressed and need further scientific investigations. First of all, we used the subjective loads to regress the theta and alpha coefficients in the logistic function model. However, as pointed in the previous discussion, the subjective load may sometimes dissociate with the operator experienced workload in the dual task or multi-task conditions (Yeh and Wickens, 1988). Therefore, using the subjective load as the workload standards would be problematic. During the regression of the logistic function model, we tried to slightly adjust the subjective loads and found even a small adjustment may produce significant differences in the regressed coefficients. Therefore, a standard workload scheme and the correspondent theta and alpha p-values would be beneficial for obtaining an accurate model. Secondly, the estimation of the general distribution function and accumulative distribution function requires an experiencing of full-range task loads. However, in practice, it's unrealistic to manipulate too many workload levels in an experiment. Therefore, a standardized procedure, e.g. the development of a standard task, for the estimation of subject specific general distribution of the EEG parameters is definitely essential. Additionally, other issues, e.g. the criteria for the definition of the workload thresholds, should be also addressed in the future research.

5.5 Summary

In this chapter, a new computational model for driver mental workload detection is presented. Firstly, the analysis of short-term theta and alpha powers from the second experiment indicated that p-quantiles could be used as an effective generalization method to customize the individual differences. Secondly, the logistic function model (LFM) is proposed to quantify mental workload through combining the theta and alpha power into scales ranged from 0 to 1. The p-quantiles were used to determine the individual-specific coefficients of the LFM. The results indicated that LFM could reduce the overlays of the distribution

functions of workload scores induced by different task loads, which suggests a better separation of these scores for task load levels. Additionally, the workload scores using the LFM increased with the task load and showed improved correlations with other workload parameters compared to the single theta or alpha powers, which indicates that the LFM has a benefit for establishing a general workload metrics. In the discussion section, the comparison of the p-quantile with the Z-score and the merits of LFM are given. The biggest advantage of LFM is that it can reliably reflect the workload states and enables an easy definition of the workload thresholds, which are adjustable as well for various application cases. However, the disadvantages for LFM, e.g. an uneasy predefinition of standard workload scores for coefficients regression, and an unclear definition of the workload thresholds, still need to be addressed in the future research.

Chapter 6. Driver adaptive task allocation in driving simulator

6.1 Motivation

Adaptive task allocation is a new concept for the human-centered system design. As introduced in Chapter 1 and Chapter 2, studies attempting to build a closed-loop in the human-machine systems based on psychophyisological markers have already shown some promising results, e.g. both Prinzel et al. (2000) and Wilson & Russell (2007) revealed that psychophysiologically driven adaptive systems are helpful to improve the operator's task performance. In driving context, Kohlmorgen et al. (2007) reported that an EEG-based system could improve the driver's overall task performance under real traffic situation. However, one study of course is not enough to support the integration of such a concept. Adaptive task allocation in the driving context needs more convincing evidence from both laboratory and applied studies. In this chapter, we focus on building an EEG-driven closed-loop for driver adaptive task allocation in the driving simulator (see, Figure 6.1).

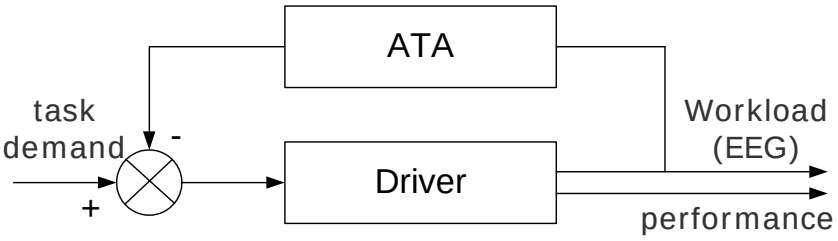


Figure 6. 1 EEG-based neural adaptive interface for driver adaptive task allocation

Chapter 3 and Chapter 4 present two experiments investigating the modulation of the EEG parameters with the mental workload. And in Chapter 5, a computational model (i.e. LFM) using theta and alpha power was promoted. These works have set the fundamental basis for the integration the EEG interface into the driver-vehicle-system to provide adaptive aiding. Particularly, in Chapter 2, the

theoretical DWM-ATA model is proposed and provides a protocol for the adaptive aiding design. The objective of this chapter is to build a closed-loop for driver adaptive task allocation and take the DWM-ATA model into the practice.

In this experiment, the LCT, and the n-back task were used again. Task loads were adjusted in two dimensions, the driving task load (adjusted by driving speed) and working memory load (adjusted by n-back task). Two experiment modes, non-control (without EEG adaptive aiding) and control (with EEG adaptive aiding), were used in this study. In the control mode, EEG signals were collected to detect the driver's mental workload state in real-time and the driving speed was dynamically adjusted according to workload states. The comparison between the control mode and non-control mode in various variables, including EEG, ECG, performance, subjective loads, enabled a comprehensive understanding of the effect of such adaptive system. In the following section, the details of methods, experimental results and discussion of the results are presented.

6.2. Methods

6.2.1 Participants

Overall, 23 participants (12 men and 11 women) between the age of 22-34 (mean age is 27.3 with a standard deviation of 3.65) participated in this experiment. All individuals reported to be free of illness and medication. They were paid for their participation in the study. Results from three participants were excluded because of sickness or extremely bad n-back performance (only 30% and 40% accuracy in 2-back task).

6.2.2 Tasks and task load manipulation

The descriptions of the tasks, the LCT and the n-back, are given in Chapter 3 and Chapter 4 respectively. Here, we do not rereiterate them. But some details about the task condition design should be addressed. In the present study, task load was modulated in two dimensions: the primary driving task demand (by driving speed) and the secondary working memory load (by n-back load factor n). The whole experiment contained two task modes: non-control and control, with each including three n-back task conditions. In non-control task mode, three sub-conditions were involved: a slow driving (75km/h) without n-back task (slow driving & no n-back), a combination of slow speed driving (75km/h) and 1-back task (slow driving & 1-back), and a combination of high speed driving (110 km/h) and 2-back task (fast driving & 2-back), representing low, moderate and high task load respectively. In control mode, there were also three n-back task conditions, no n-back, 1-back and 2-back. However, the driving speed was not constant but adjusted between 40 and 120 km/h through a closed-loop, in which the EEG was used to estimate the workload states and driving task load was modulated dynamically according to the workload states by adjusting the speed settings. Working memory load was used to produce different secondary task load levels. However, we did not change the working memory load according the EEG estimated workload states.

6.2.3 Experiment setup

Like the second experiment, the basic experimental setup included a driving box (see, Chapter 3 and Chapter 4), EEG systems, PCs and two projectors which were used to project the LCT driving scene and n-back digits respectively. The details for the EEG recording, ECG electrode placing could be found in Chapter 3. For establishing a closed-loop, the whole system consisted of 4 PCs (Figure 6.2). PC1

and PC 2 ran the LCT and n-back software respectively. Two projectors were used to project the LCT driving scene and n-back digits respectively. The n-back digits projection was overlaid upon the LCT driving scene and randomly located in the driving scene. The Brain Vision Recorder (Brain Products, Germany) was installed on PC4 and recorded the EEG data. A Matlabserver (developed by PhyPa Group at Chair of Human-Machine-System in Technische Universitaet Berlin) is a Microsoft Visual C++ based toolbox to perform online processing of EEG data and was installed on PC4 to detect the workload with an output ranging from 0-1. These results were sent to the other PCs through TCP/IP. A microcontroller was used to replace the gas pedal for LCT speed control, i.e. the speed could be automatically set without any input from the participants.

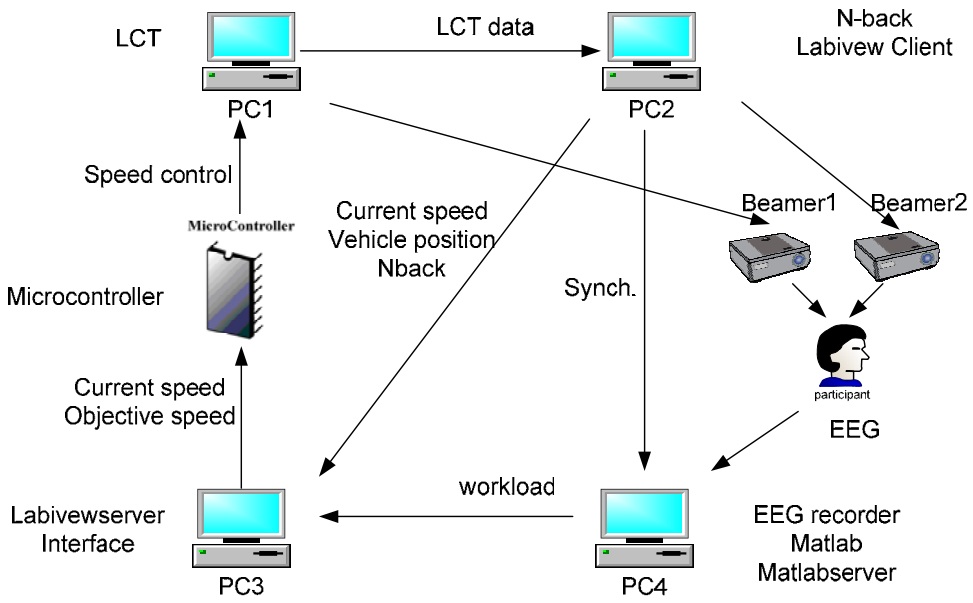


Figure 6. 2 System overview in the third experiment. The whole system consisted of four computer and two projectors. A microcontroller was used as well to control the speed settings.

The data communication among LCT, n-back and EEG was solved using two self-developed Labview (National Instruments, USA) tools, a Labview Client (installed on PC2) and a Labview Server (installed on PC3). The Labview Client

received LCT log data through TCP/IP and automatically triggered the n-back program when it detected the start of each track of LCT and shut it down when it came to the end of each track. Meanwhile it transferred the information, such as the current driving speed, n-back condition and vehicle position, to the Labview Server. The Labview Server received data from the Labview Client as well as the workload result from the Matlabserver. It calculated an objective speed which will be set according to current driving speed, n-back condition and workload scores. Most important is that the Labview Server sent a command containing the current speed and objective speed to the Microcontroller, which adjusted the driving speed to the objective speed with a digital potentiometer. Additionally, two buttons embedded in the steering wheel allowed the driver to respond to the n-back task and the response information was sent to the EEG recorder, which enabled the evaluation of the n-back performance and reaction time.

The Labview Server includes also a data visualization interface (Figure 6.3). This interface allows a real-time monitoring of different variables. The charts ‘driving speed’, ‘task demand’, and ‘mental workload’ (left side of Figure 6.3) show the historical fluctuations of driving speed, task demand and EEG estimated mental workload respectively. The current values of the driving speed, task demand and workload are shown with a sector display and two bar charts (right side of Figure 6.3) respectively. The right bottom is placed with three lights ‘underload’, ‘normal’, and ‘overload’ to indicate the current workload states. Such an interface makes it easy for the experimenter to dynamically assess how driver’s mental states modulate with the driving task load. It is also valuable for the experimenter to evaluate how reliable such EEG-based system is.

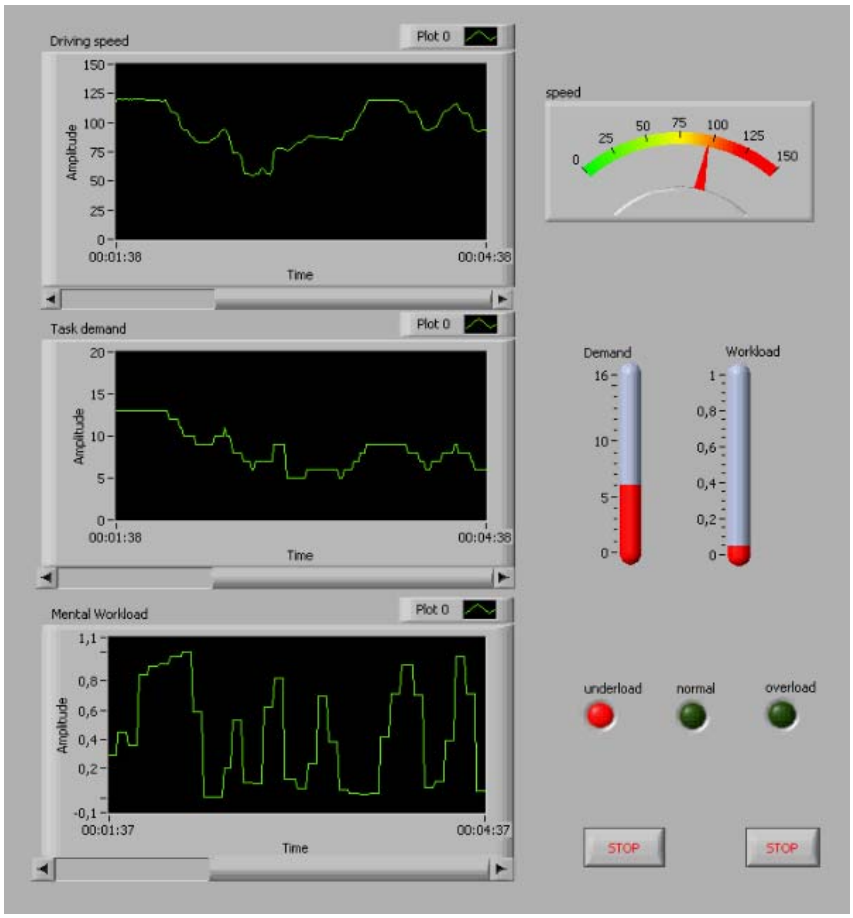


Figure 6. 3 Labview interface for EEG-based closed-loop system in the third experiment. Speed, task demands, mental workload, and the mental workload states can be dynamically visualized driving the experiment.

6.2.4 Quantification of workload with EEG signal

The data processing was based on the online framework of PhyPha group, which includes a Matlab toolbox and a Matlabserver (based on C++). The Matlab toolbox contains various models for EEG processing. We integrated our model for workload quantification (workload_detection_toolbox_1.3, see the Appendix 1) into this framework. The Matlabserver could access EEG data from the Brain Vision Recorder, execute the Matlab scripts, and exchange parameters with the Matlab workspace. We extended the original Matlabserver with another communication channel with the Labviewserver, in which the workload detection results can be sent for the online speed adjustment.

The workload quantification using EEG data was done in the following procdures: the EEG data were firstly filtered with a bandpass filter (1-40Hz) and a laplacian filter (Hjorth, 1975), which helped to reduce the influence of the ocular artefacts. Then data were segmented into 10-second epochs with 50% overlay. FFT was used to estimate the power spectrum density and relative theta power and alpha power (the ratios to the total power of 4-30 Hz) were extracted from the electrode site Fz and Pz respectively. The data from the training session was used to extract the p-quantiles using the p-values shown in Table 5.1 (Chapter 5). These p-quantiles combining with the predefined workload scores (shown in Table 5.3) were used to determine the subject specified coefficients $\beta_0, \beta_1, \beta_2$ in the LFM (see section 5.3.1). In the control experiment mode, the workload scores were calculated with the trained LFM and two thresholds (0.2 and 0.8) were directly used to classify the workload into three states, underload (<0.2), normal (0.2-0.8), and overload (>0.8).

6.2.5 Task demand adjustment using DWM-ATA

The task demand in control modes was modulated with the DWM-ATA model. The DWM-ATA model requires firstly a quantification of the task demand. In this study, we quantified the task demand into 16 scales. We scored 1 for the driving condition with fixed speed 40km/h and added 1 scale for task demand when the speed increased with 10 km/h. For quantifying the n-back task, we used the results showed in Table 5.4 in section 5.3.3. As shown in this table, the mean workload scores for 1-back single task was 0.385 and for 2-back single task was 0.482, which is approximately equal with the workload for the single slow driving (0.384) and fast driving (0.482) respectively. Therefore, the n-back task demand was quantified as following routine: no n-back, 1-back and 2-back were scored as 0, 4 (equal to driving at speed 75 km/h), and 7 (equal to driving at speed 100 km/h)

respectively. The general task demand is calculated by adding the speed scores with n-back task scores. Since the maximum speed was restricted to 120 km/h (equivalent to task demand 9), the highest task demand (highest driving speed with 2-back) is 16.

We modelled the task demand into three levels with two thresholds 4 (equivalent to 70 km/h driving or 1-back only in which almost all participants reported it was easy to cope with) and 12 (equivalent to 80 km/h driving combining 2-back task in which most participants reported it was difficult): low (task demand < 4), moderate (4 <task demand < 12) and high (> 12). The EEG workload score was also divided into three levels using a low and high threshold: low level (EEG workload score <0.2), moderate level (0.2< EEG workload score <0.8) and high level (EEG workload score> 0.8). The demand-workload plane was then divided into 5 areas A, B, C, D and E (see, Figure 6.4). The speed adjusting rules were designed as following:

- In area A, task demand was low but workload was high, speed was maintained.
- In area B, task demand was moderate or high and workload was high, speed was reduced.
- In area C, the workload was moderate, speed was maintained
- In area D, the workload was low and task demand was low or moderate, speed was increased.
- In area E, the workload was low and task demand was high, speed was reduced.
- The speed adjustment was made in step of 10 km/h.

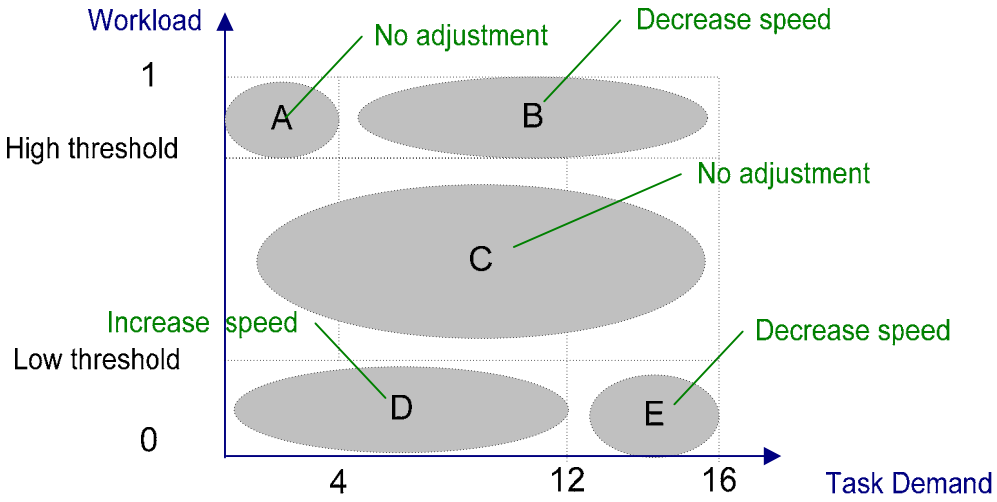


Figure 6. 4 Task demand control strategy according to the DWM-ATA model

6.2.6 Experiment procedure

Participants were firstly asked to fill out a form regarding personal information such as age, driving experience, disease situation, etc., and were also given an experiment introduction. After the EEG electrode preparation, they had a 10-minute practice session. NASA-TLX weights were collected after the practice session. Then, they were required to perform a 25-minute training session containing 9 randomly organized combinations (no driving & no n-back, no driving & 1-back, no driving & no 2-back, slow driving & no n-back, slow driving & 1-back, slow driving & 2-back, fast driving & no n-back, fast driving & 1-back, fast driving & 2-back), which gave the participant a full-experience of all task load conditions. This session allowed obtaining the statistical properties of EEG data and training the logistic regression workload detection model. After the training session, 12 blocks containing twice of each of no-speed-control conditions (slow driving & no n-back, slow driving & 1-back, fast-driving & 2-back) and speed-control conditions(speed-control driving with no n-back, 1-back and 2-back), were randomly performed to avoid a learning effect. These 12 blocks allowed a comprehensive pairewise comparison between no-speed-control

and speed-control under low task load, moderate task load, and high task load levels to investigate the effect of the closed-loop system. A 3-minute break followed each block and during this break, participant was required to evaluate the NASA-TLX rating scores. The whole experiment lasted around 150 minutes.

6.2.7 Data analysis

Totally, 8 variables, including the driving speed, n-back task performance and reaction time, NASA-TLX, heart rate (HR) and heart rate variability (HRV), EEG workload scores (EEG W.L. score) and EEG percentages of the workload levels, were analyzed from all 20 participants afterwards. The preprocess of EEG data was similar to that used in Chapter 4 and the calculation of EEG W.L. scores followed the LFM which presented in Chapter 5. Two workload threholds (0.2 and 0.8) were used to calculate the percentages of low, moderate, and high workload levels. The ECG data analysis was the same with the procedure used in Chapter 4. The ratio of the energy around the low frequency (0.04-0.15Hz, LF) to the energy around the high frequency (0.15-0.4Hz, HF) was analyzed to offer an alternative variable for the mental workload evaluation (Kamada et al., 1992; Murai et al., 2004).

The general concept for the data analysis was: (1) a general comparison between the non-control and control conditions for each parameter. This step might help to generally show the effect of adaptive task allocation. (2) A detailed comparison of the non-control and control conditions in different task demand levels. The original idea is to compare the non-control and control conditions in different task demand levels which were defined by the DWM-ATA model. However, since the speed was dynamically changing, the quantification of the task demand in control conditions was difficult. Therefore, we simple define the workload levels according to the n-back conditions. This comparison may show the effect of

adaptive task allocation in different task demand conditions.

Additionally, the n-back task performance in different areas (defined by DWM-ATA model) was extracted. Again, since the task demand varied from time to time in control conditions, such separation was only done in the non-control conditions. Furthermore, since in the low task demand area we had no n-back task, there was no nback performance in areas 1, 4, and 7 (DWM-ATA model).

The statistical comparision was performed with one way Analysis of Variance (ANOVA) at a 0.05 level.

6.3 Results

6.3.1 Task demand

General comparison

The general comparison of the task demand between non-control and control mode is shown in Figure 6.5. Statistical analysis indicated the task demand significantly decreased in the control mode compared to the non-control mode ($F(1,20)=5.2$, $p<0.05$).

Comparison in different n-back conditions

Figure 6.6 shows the task demand comparison between the non-control and control mode in different n-back conditions. Compared to non-control mode, the task demand in control mode was significantly increased under no n-back ($F(1, 20)=34.4$, $p<0.001$) and 1-back conditions ($F(1, 20)=24.3$, $p<0.001$). In contrast,

the task demand was significantly reduced in control mode under 2-back condition compared to non-control mode ($F(1, 20)=199.2$, $p<0.001$). Additionally, the mean task demand levels in the control modes were all adjusted into a predefined moderate task demand level (from 4-12).

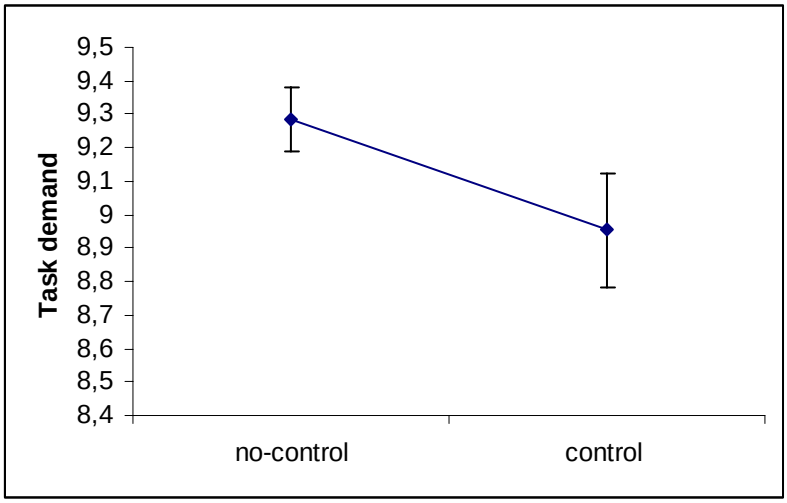


Figure 6. 5 General comparison of task demand between the non-control and control task mode (N=20). The error bar shows the standard error.

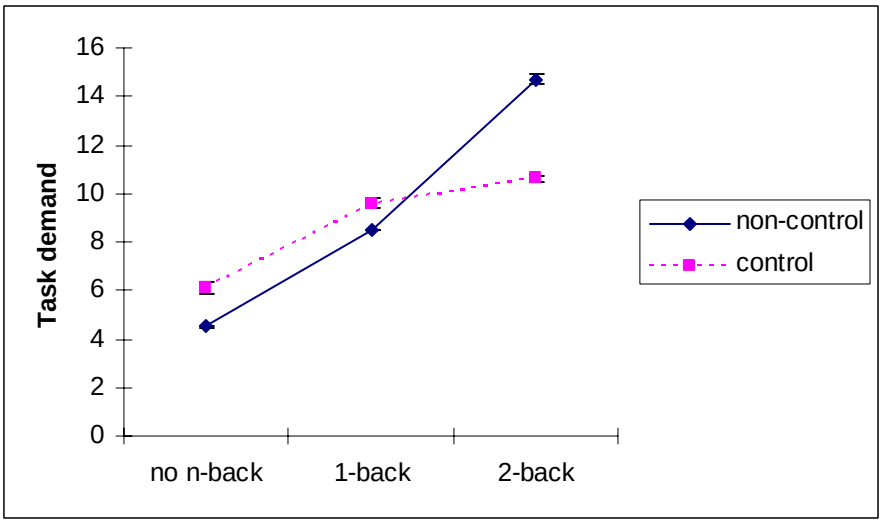


Figure 6. 6 Task demand comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.

6.3.2 Subjective load (NASA-TLX)

General comparison

Figure 6.7 shows the general comparison of the subjective load between non-control and control mode. Statistical analysis indicated that there was significant decrease in the subjective load in the control mode compared to non-control mode ($F(1,20)=4.54$, $p<0.05$).

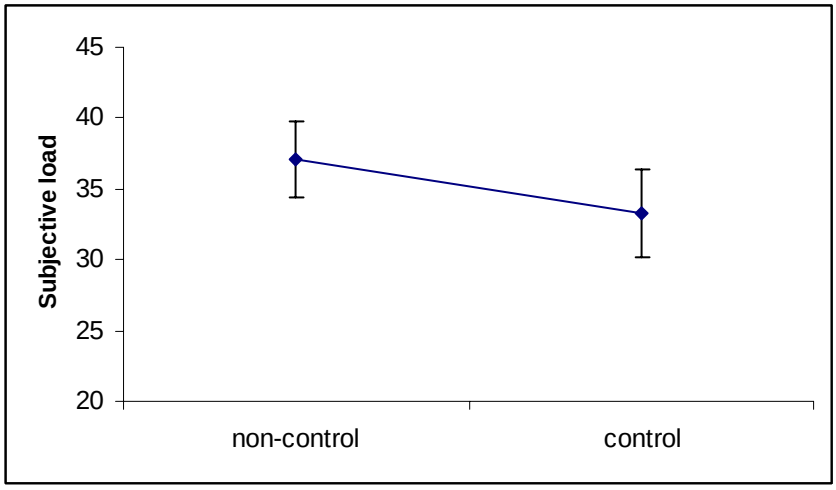


Figure 6. 7 General comparison of subjective loads comparison between the non-control and control mode (N=20). The error bar shows the standard error.

Comparison in different n-back conditions

The comparison of subjective load in different n-back task conditions is shown in Figure 6.8. In no n-back and 1-back conditions, the subjective load showed no significant differences between the non-control and control mode (for non-control VS control in no n-back, $F(1,20)=0.01$, $p=0.93$; for non-control VS control in 1-back, $F(1,20)=0.22$, $p=0.64$). However, a significant decrease in NASA-TLX score was found in the control & 2-back compared to the non-control & 2-back ($F(1,20)=4.93$, $p<0.05$).

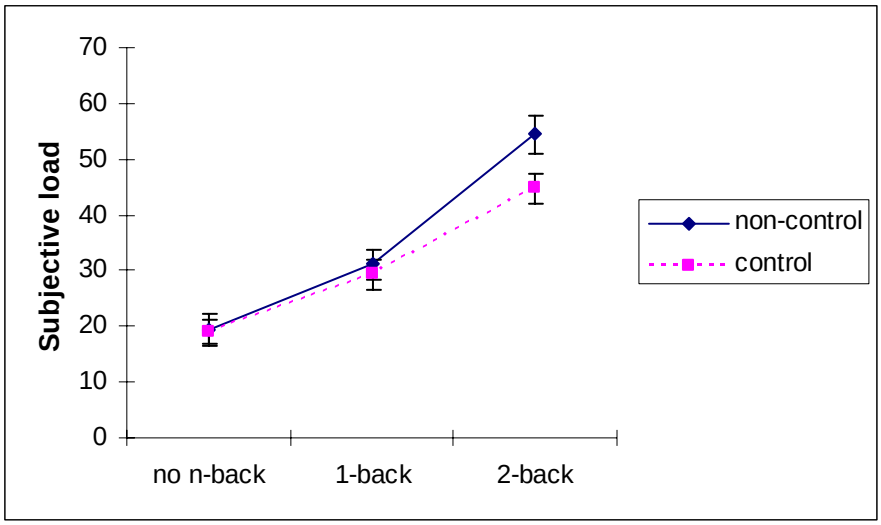


Figure 6. 8 Subjective load comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.

6.3.3 Task Performance

General comparison

The general comparisons of n-back error rate and reaction time between non-control and control mode are shown in Figure 6.9 and Figure 6.10 respectively. Statistical analysis indicated that there was a significant decrease in n-back error rate in the control mode compared to non-control mode ($F(1,20)=4.13$, $p<0.05$). However, there was no significant difference in the reaction time between the control mode and non-control mode ($F(1,20)=0.37$, $p=0.54$).

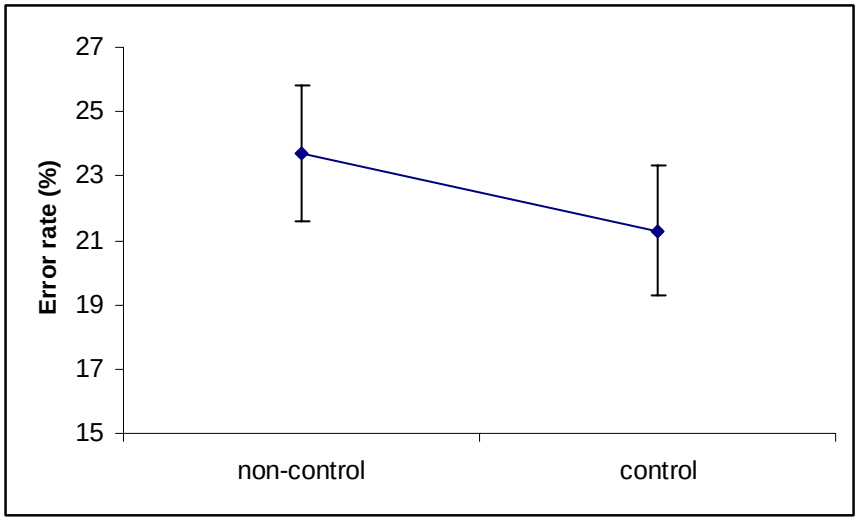


Figure 6. 9 General comparison of n-back error rate between the non-control and control task mode (N=20). The error bar shows the standard error.

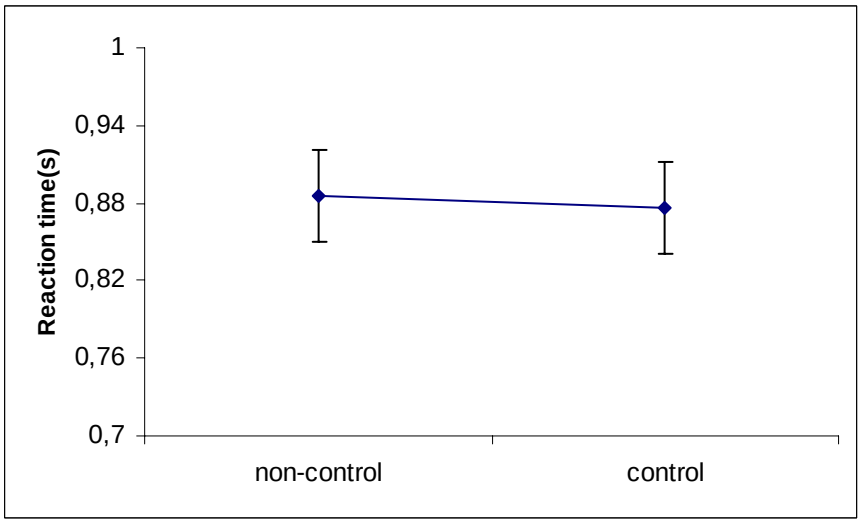


Figure 6. 10 General comparison of n-back reaction time between the non-control and control task mode (N=20). The error bar shows the standard error.

Comparison in different n-back conditions

The comparison of n-back error rate and reaction in different n-back conditions is shown in Figure 6.11. The error rate showed less than 20% for both non-control and control modes in 1-back condition, and no significant differences between these two modes ($F(1,20)=1.68$, $P=0.20$) were found. In fast driving & 2-back 33.2% error rate was found and this percentage decreased significantly to 23.5%

in control model ($F(1,20)=4.69$, $p<0.05$), which showed that the speed adjustment significantly enhanced n-back task performance in 2-back task. Reaction time (shown in Figure 6.12) demonstrated no significant differences between non-control and control mode in both 1-back and 2-back conditions (for, non-control VS control in 1-back, $F(1,20)=0.03$, $p=0.85$; for non-control VS control in 2-back, $F(1,20)=0.29$, $p=0.59$).

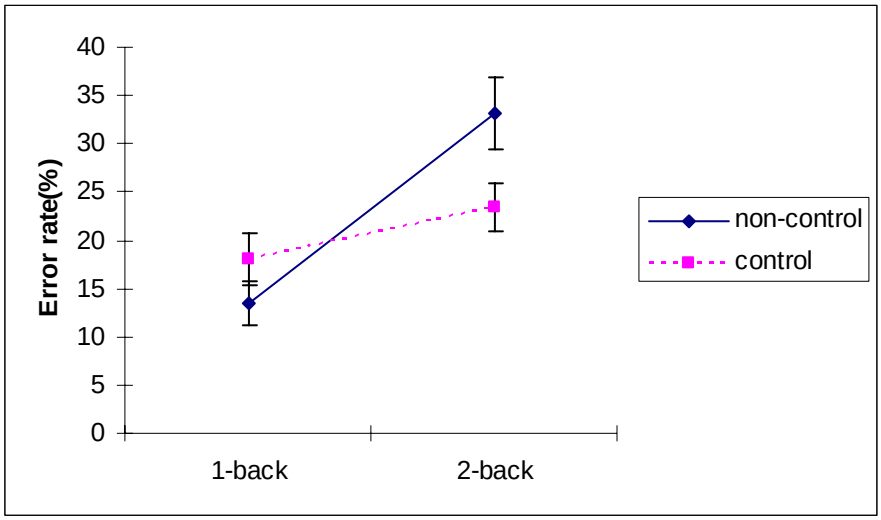


Figure 6. 11 n-back error rate comparison between the non-control and control task mode (N=20). The error bar shows the standard error.

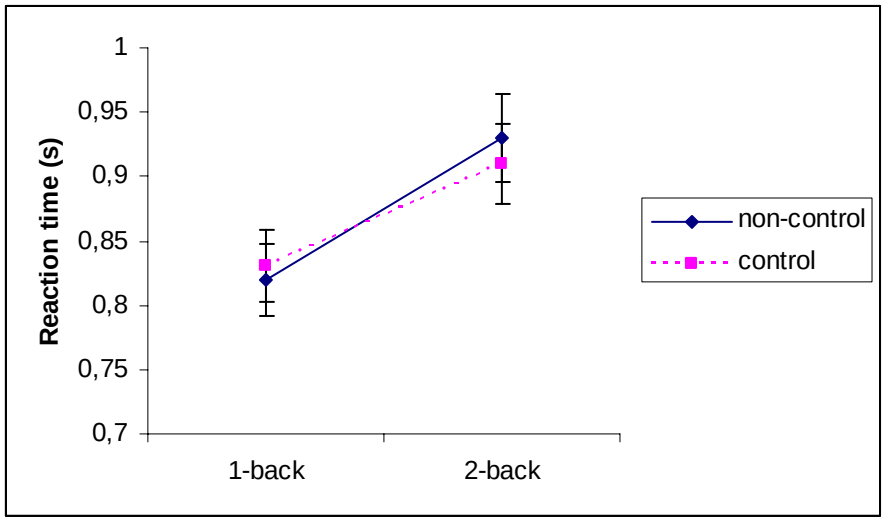


Figure 6. 12 n-back reaction time comparison between the non-control and control task mode (N=20). The error bar shows the standard error.

6.3.4 HR and HRV

General comparison

The general comparison of HR and HRV between non-control and control mode is shown in Figure 6.13 and Figure 6.14 respectively. There was no significant difference in HR between control mode and non-control mode ($F(1,20)=0.49$, $p=0.48$). There was also no significant difference in HRV between the control mode and non-control mode ($F(1,20)=0.05$, $p=0.82$).

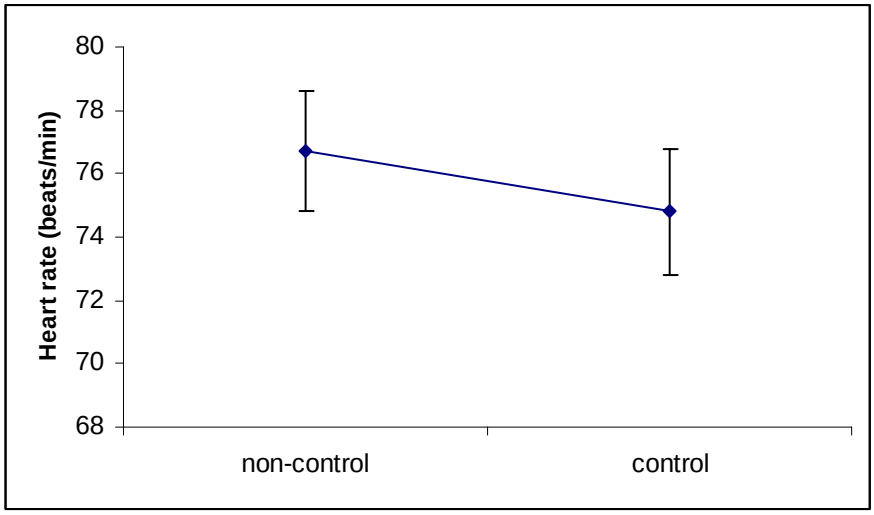


Figure 6. 13 General comparison of HR between non-control and control model (N=20). The error bar shows the standard error

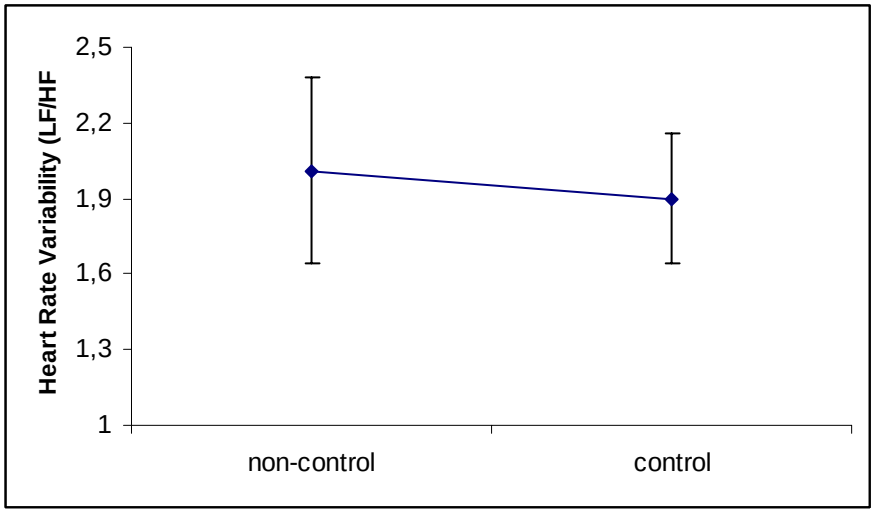


Figure 6. 14 General comparison of HRV(LF/HF) between non-control and control model (N=20). The error bar shows the standard error.

Comparison in different n-back conditions

The results of HR and HRV in different n-back conditions are shown in Figure 6.15 and 6.16 respectively. There were no obvious differences in HR between each pair of the non-control and control mode (for non-control VS control in no n-back, $F(1,20)=0.11$, $p=0.74$; for non-control VS control in 1-back, $F(1,20)=0.05$, $p=0.82$; for non-control VS control in 2-back, $F(1,20)=0.12$, $p=0.73$). And no significant differences were observed in HRV (LF/HF ratio) as well (for, non-control VS control in no n-back, $F(1,20)=0.05$, $p=0.82$; for non-control VS control in 1-back, $F(1,20)=0.16$, $p=0.69$; for non-control VS control in 2-back, $F(1,20)=0.0$, $p=0.97$).

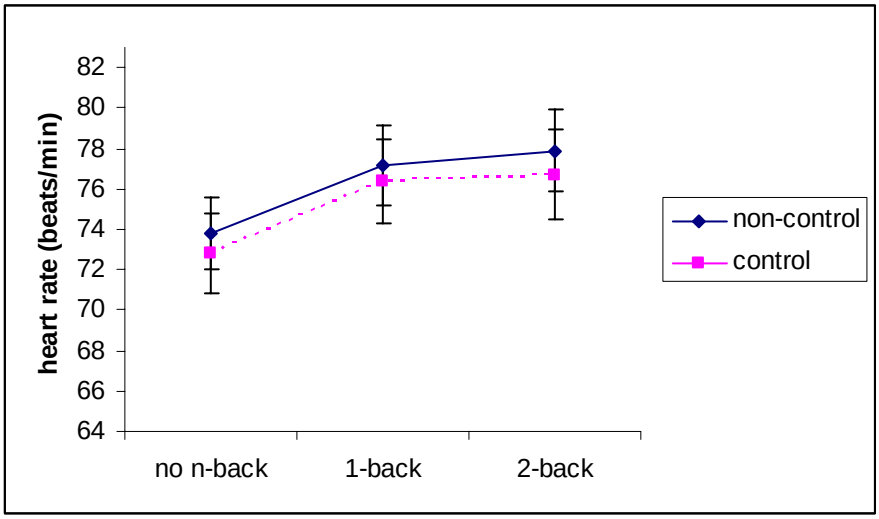


Figure 6. 15 Heart Rate comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.

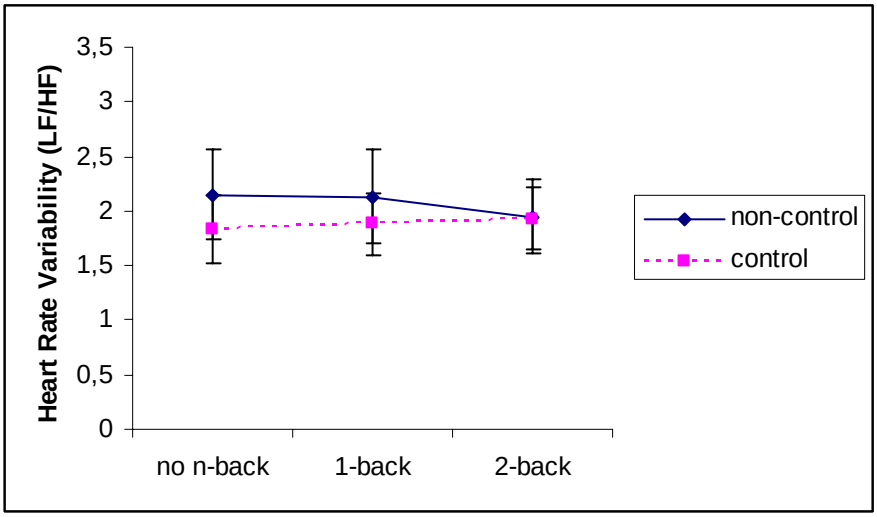


Figure 6. 16 HRV (LF/HF) comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.

6.3.5 EEG Parameters

General comparison

The general comparison of EEG W.L. score between non-control and control mode is shown in Figure 6.17. Statistical analysis indicated that there was no significant difference in this parameter between the non-control and control mode ($F(1,20)=0.94$, $p=0.37$). The general comparison of the percentages of low, moderate and high workload between non-control and control mode is shown in Figure 6.18. Statistical analysis demonstrated that there was no significant differences in these three parameters between the control mode and non-control mode (for low workload percentage, $F(1,20)=1.91$, $p=0.18$; for moderate workload percentage, $F(1,20)=2.28$, $p=0.14$; for high workload percentage, $F(1,20)=0.46$, $p=0.53$).

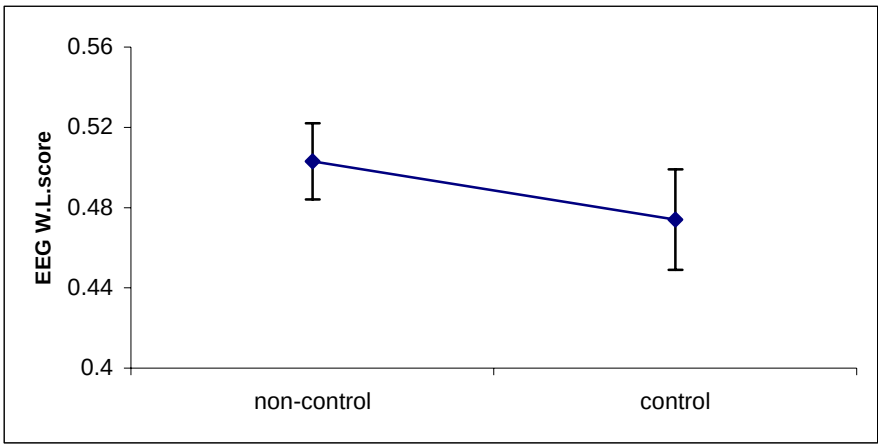


Figure 6. 17 General comparison of EEG W.L. score between non-control and control model (N=20). The error bar shows the standard error.

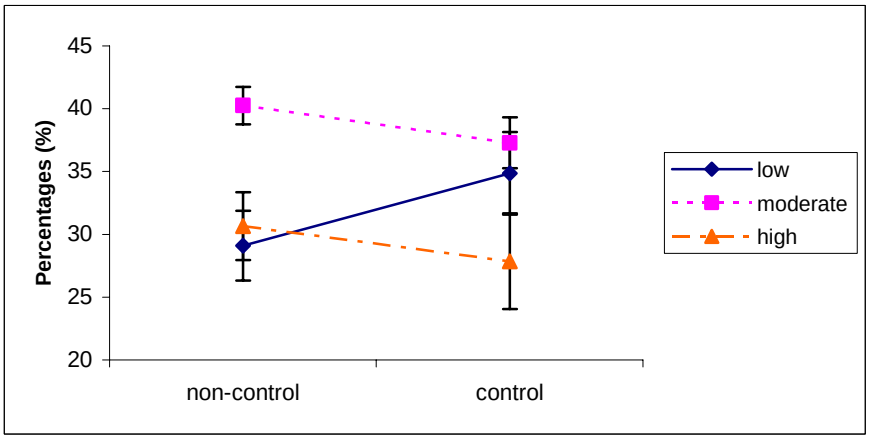


Figure 6. 18 General comparison of percentages of low, moderate and high workload between non-control and control model (N=20). The error bar shows the standard error.

Comparison in different n-back conditions

Comparison of the EEG workload scores (EEG W.L. socre) in different n-back conditions is shown in Figure 6.19. No significant differences in the EEG workload scores were observed between non-control and control mode in no n-back condition ($F(1,20)=0.04$, $P=0.84$) as well as in 1-back condition ($F(1,20)=0$, $P=0.95$). However, the EEG workload score decreased significantly from 0.60 in the non-control & 2-back to 0.52 in control & 2-back ($F(1,20)=4.44$, $P<0.05$). It indicated that the adaptive speed adjustment in the 2-back condition helped to

reduce driver's mental workload.

The percentages of low, moderate and high workload classified using EEG workload scores showed significant differences only between the two sub-conditions of 2-back (see, Table 6.1). The low workload percentage in control & 2-back was significantly increased compared to non-control & 2-back ($F(1,20)=4.41$, $P<0.05$), whereas the high workload percentage showed an opposite tendency ($F(1,20)=4.15$, $P<0.05$). The moderate percentage showed no significant difference between non-control & 2-back and control & 2-back.

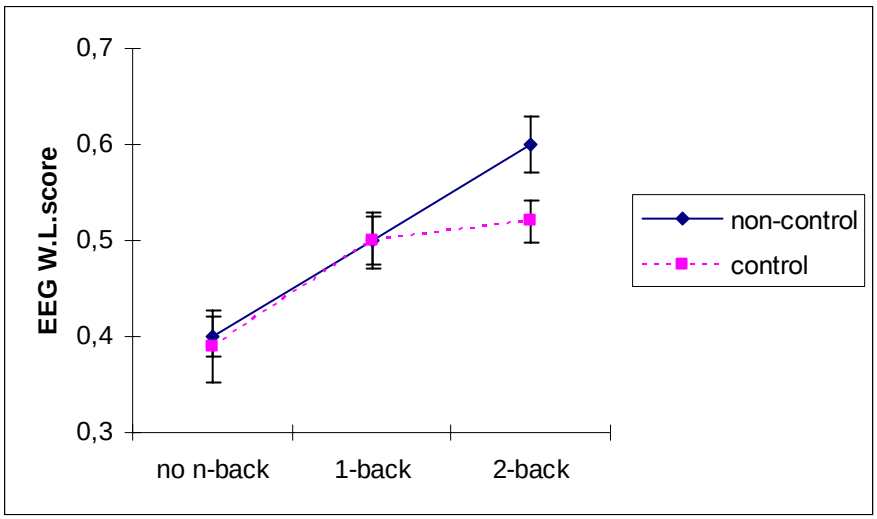


Figure 6. 19 EEG workload scores comparison between the non-control and control task mode (N=20). The error bar shows the standard error.

6.3.6 DWM-ATA model and performance

The results linking the performance with the DWM-ATA model were shown in Figure 6.20. In general, the performance in high task demand demonstrated much higher error rate compared to the medium task demand ($F(1,20)=66.8$, $p<0.001$). The lowest error rate occurred in area 5 (medium-medium area). By contrast, the highest error rate (37.2%) was observed in the area 3, which followed by area 9 (35.6%). Statistical analysis showed that the difference in the error rate between

area 3 and area 6 was significant ($p<0.05$). The difference in the error rate between area 6 and area 9 was also significant ($p<0.05$). However, there was no significant difference between area 3 and area 9 ($p=1$).

Table 6. 1 The mean percentages (standard deviation) of low, moderate and high workload averaged over participant (N=20).

Conditions		Low (%)	Moderate (%)	High (%)
no n-back	non-control	42.4(16.2)	41.6(12.5)	16.0(11.1)
	control	47.0 (24.5)	33.0 (14.7)	20.0 (23.0)
1-back	non-control	28.3(20.4)	41.8(13.2)	29.8(15.9)
	control	30.8 (17.3)	40.3 (10.6)	28.8 (22.7)
2-back	non-control	16.2(0.161)	37.3 (10.2)	46.2 (20.1)
	control	26.7 (145)	38.5 (9.7)	34.7 (15.0)

Note: Workload was classified by EEG workload scores (smaller than 0.2 is low; higher than 0.8 is high; between these two thresholds is moderate).

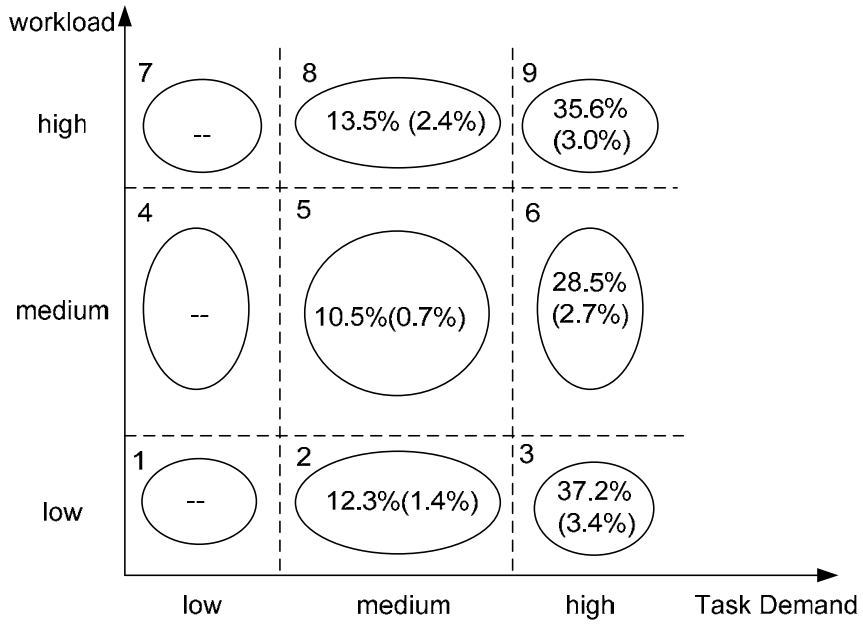


Figure 6. 20 The n-back error rate (standard error across participant) according to DWM-ATA model (There was no n-back task in low task demand region. Thus, there was no performance data in this region)

6.4 Discussion

In this chapter, the theoretical DWM-ATA model was applied in an EEG-based closed-loop for driver mental workload detection and adaptive task allocation. A laboratory study was conducted to investigate the effect of such adaptive system. In the following section, some of these results are discussed.

6.4.1 Psychophysiology based adaptive task allocation

The general comparison in different parameters between the non-control mode and control mode showed that the adaptive system could significantly reduce the overall task demand, though in the low demanding cases it increased. The subjective load and n-back error rate both were observed to be reduced in the control mode compared to the non-control mode. This indicates such psychophysiology-based adaptive system and the DWM-ATA model exhibit potential benefits for improving the human-machine-interaction. Similar supporting can be also found in previous studies (Prinzel et al., 2000; Wilson & Russel). Prinzel et al. (2000) reported a psychophysiology based closed-loop system could moderate the level of engagement based on engagement index (beta/ (alpha plus theta)). They used a modified version of the MAT (Multiple Attribute Task) battery involving EEG-based negative and positive feedback. In negative feedback conditions, the level of automation in the tasks was lowered when the EEG index reflecting decreasing engagement. On the other hand, when the EEG reflected increases in task demands, the automation level was increased. This negative feedback process matched the concept of Parasuraman's model for adaptive allocation. The results from Prinzel's study showed that negative feedback condition elicited better tracking performance, which is consistent with our finding.

Particularly, the detailed comparison between the non-control and control mode in different n-back conditions showed that improvement of task performance was only found in the high workload condition (2-back) in the control mode, but no differences in the performance under the low and moderate task demand conditions were found. Specifically, results from the EEG workload scores, NASA-TLX ratings both supported that such system could reduce the operator's mental workload when the task is too demanding. Such results are also consistent with Prinzel's study (Prinzel et al., 2000). Prinzel and his colleagues noticed that more task allocations were made under the multiple task condition and concluded that such closed-loop appeared to be sensitive to increases in task load. The present finding that control mode only elicited significant differences in the performance, EEG workload scores and NASA-TLX ratings in 2-back condition was in line with Prinzel's conclusion.

One objective of this experiment is to maintain the participant's mental workload in the moderate level through the adjustment of task demand. The system was designed to increase percentage (between 0.2 and 0.8) of moderate workload and decrease the percentages of high and low workload. As expected, the high workload percentage was significantly reduced in the 2-back condition due to the speed adjustment. However, interestingly, there were no obvious increases in the percentages of the moderate workload in the feedback control condition. Furthermore, the percentage of low workload level in the 2-back condition was significantly increased in the speed control condition compared to fast driving condition. It seems that the change of EEG workload score involves a 'binary' effect. The task load changes lead to a switching of two functional states, 'rest' and 'busy', rather than a continuous changing from the 'rest' to 'busy'. One possible reason for this effect is due to the definition of the thresholds for the low, moderate and high workload. However, concerning the appropriate thresholds, we need further investigations.

6.4.2 The needs of DWM-ATA

In the present study, the percentages of the low, moderate and high workload in different task load conditions show that the workload, generally, linear associated with the task load. For instance, the low workload periods dominated in the low task load condition; whereas high percentages of high workload periods in the high task load condition was observed. However, the results also improved the existing of the area 3 and 7 in DWM-ATA model, in which the occasional dissociation of the workload and task demand may happen, e.g. the EEG-based workload estimation revealed there were 16% high workload periods in low workload condition as well as 16% of low workload periods in high workload condition.

These findings are consistent with the previous theory (Yeh and Wickens, 1988; Hancock, 1996; Parasuraman and Hancock, 2001). Parasuraman and Hancock (2001) pointed out that dissociation occurs in some unusual circumstances that the experienced workload might be failed to linearly associate with external task demand. More theoretically, Yeh and Wickens (1988) pointed out that such dissociation may happen when increased resources are invested to improve performance of a resource-limited task; when task demands are increased by time-sharing. In the present study, attentional resource competition between the primary LCT and secondary n-back become severer with the increasing working memory load and driving speed. Some participants gave up the performance of 2-back in non-control mode. But they still reported high subjective load scores.

The evidence of the dissociation between task demand and workload provide empirical support for extending the Parasuraman's workload-match ATA model to the proposed DWM-ATA model. In DWM-ATA model, regions 4, 7 and 8 represents a state of 'overpaced', in which the participant's workload is higher

than required by the task. These situations could be a result of inappropriate emotional states (Averty et al., 2004) or a result of improper strategy the participant adopted (Parasuraman and Hancock, 2001). Some methods which help to release the participant's mental stress (e.g. relaxing music, Knight and Rickard, 2001) or improve the task strategy would be preferred in this area rather than decrease the task demand. In contrast, regions 2, 3, and 6 in DWM-ATA model represent a state of 'underpaced' in which the participant pays less effort than required by the task. Particularly, in region 3, the participant might give up the task since the task is too demanding for them. The n-back performance according to DWM-ATA (Figure 6.20) proved that the participant had a worse performance when s/he experienced low workload state in a high-demanding task. This indicates that when the participant is in area 3 (low workload and high task demand), to continuously increase the task demand (according to Parasuraman's model) may pose a risk of performance breakdown. Many factors, e.g. motivation and the risk associating the task, may contribute to these regions (Hilburn and Jorna, 2001). It is dangerous when the participant is in these regions, since participant devote inadequate engagement in the task. Such situations might be more frequently observed in the laboratory studies with artificial environment, in which participant has less sense about the safety.

The occurance of the dissociation between the task demand and workload and also the potential breakdown in the performance in area 3 emphasize the necessity of the DWM-ATA model, particularly, for the psychophysiology-based the adaptive task allocation, in which the workload is assessed instantaneously and the dissociation between task demand and workload may more frequently occur. In general, in DWM-ATA model, task load assignment in most of the regions (1, 2, 4, 5, 6, 8, 9) is consistent with the Parasuraman's (1999) model for adaptive task allocation, increasing task demand when the workload of the participant is low and decreasing task demand when the workload is high. The differences lie only in

the region 3 and 7. In region 3, it is also necessary to lower the task load level to avoid ‘give up’ because of high task load. In region 7, the task load level is maintained, since the high workload is probably unrelated to the task load in this area. We suggest that other solutions (e.g. trigger light music) may be more efficient to help the drivers to release themselves.

6.4.3 DWM-ATA model in driving context

DWM-ATA model portrays the relation between the workload and task demand including both association and dissociation conditions. This model requires firstly an estimation of the task demand placed on the human operator from the external task. In driving context, the task demand may be a result of the various factors including the road curvature, weather conditions, traffic density (De Waard, 1996; Baldwin et al., 2004) as well as the in-vehicle tasks induced by the in-vehicle technologies (Horrey & Wickens, 2004; Caird et al., 2005). Therefore, the DWM-ATA interprets not only the interaction of the vehicle and driver, but also the interaction of three elements in the traffic loop, driver, vehicle, and environment. Figure 6.21 shows the general concept of applying the DWM-ATA model in the driving context. The task demand posed by the in-vehicle tasks and environmental factors can be quantified as the one input of the DWM-ATA model. Another input comes from the psychophysiological assessment of the driver’s mental workload. These inputs can be mapped with the demand-workload plane to determine the optimal task allocation solution.

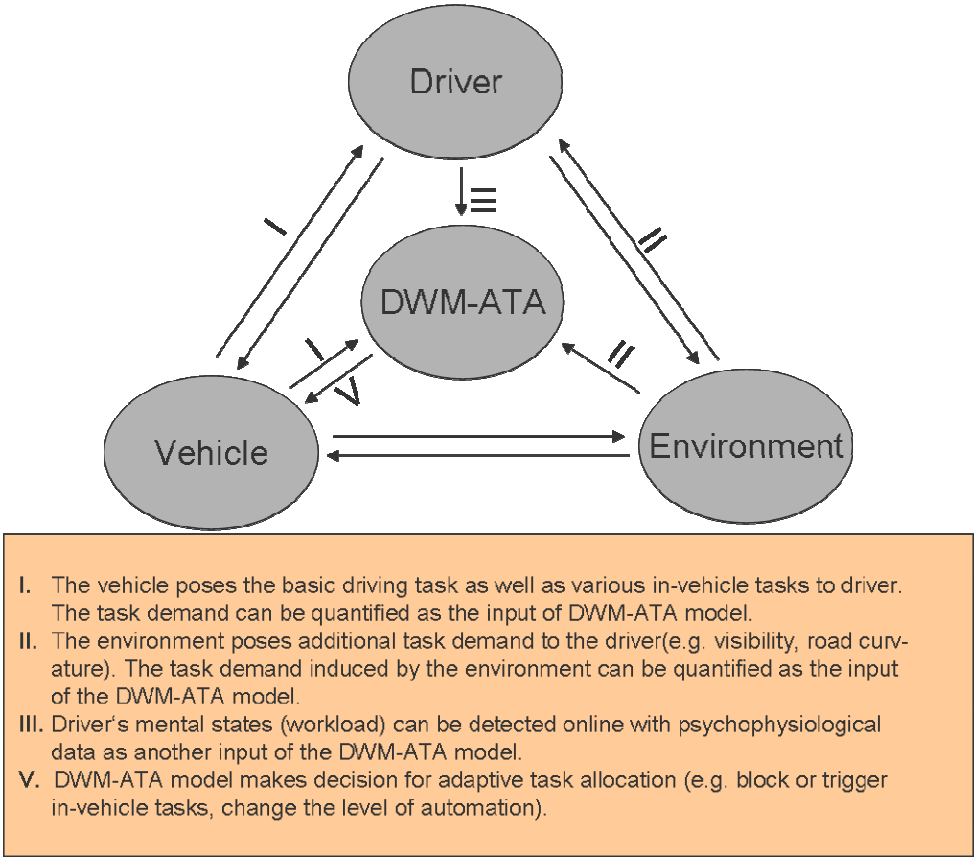


Figure 6. 21 DWM-ATA model for the driver-vehicle-environment loop

The prior studies of the driver workload can be categorized into two groups according to this concept. On the one hand, psychophyiological assessment has been established a valid and effective method for the driver workload monitoring (Brookhuis and De Waard, 1993, 2010; Kohlmorgen et al., 2007; Hagemann, 2008; Lei et al., 2009 a, b,c). On the other hand, other researchers attempted to quantify the task demand from road or environment situation and in-vehicle tasks. Environment parameters, such as road curvature, traffic density, weather, etc., might contribute to the external demand for the driver (Musa and Navin, 2005; Horrey and Lesch, 2009). Musa and Navin (2005) proposed a rate of information processing demand (RID) combining the road complexity and driving speed in term of time constrain. They proposed to regulate the driving speed corresponding to the task demand from the driving environment, which is similar with the idea in the present study to dynamically regulate the driving speed by the driver mental

workload. More recently, Horrey and Lesch (2009) quantified the demand of the different road section with scales, for example, the demand of the narrow road scored 1 and a turn scored 0.85, whereas a straight road scored 0.37. They also presented some results, the demand score for different types of in-vehicle tasks, e.g. a long phone conversation scored 0.73. These studies surely provided some solutions for quantifying the objective task demand exposed to the driver.

These studies deriving the driver workload either from driver's psychophysiology or from external task demand driven by the environment or in-vehicle tasks set a fundamental basis for the application of the DWM-ATA model in the driving context. The matching of the task demand and driver states could be dynamically associated to different demand-workload areas in the DWM-ATA model and plausible regulation of task allocation can made accordingly. However, quantifying the task demand is a challenge for bringing the DWM-ATA model into practice. Although many studies suggested a context-based solution for this, there is still a lack of systematic and well-accepted methodology to quantify task demand across various driving contexts.

6.5 Summary

In this chapter, the methods and results from the third experiment are presented. This experiment is designed to take the theoretical DWM-ATA model and the proposed LFM into practice to investigate the feasibility of the EEG-driven adaptive task allocation in a simulated driving environment. Again, the LCT and n-back task were used in this experiment. The whole experiment consisted of n-back conditions with each containing two modes: the non-control mode (without adaptive aiding) and control mode (with adaptive aiding). In the control mode, the results of EEG estimated workload states were immediately used to regulate the driving task load induced by driving speed. The results from study demonstrated

that the LFM can effectively quantify driver mental workload using the EEG parameters. The comparison of the non-control and control mode shows that psychophysiologically driven adaptive task allocation can improve operator's performance and reduce operator workload in the high task load condition. In the discussion section, the needs of the theoretical DWM-ATA model and its application in driving context are particularly addressed.

Chapter 7. Overall discussion, conclusion, and outlook

7.1 Overall discussion and conclusion

As discussed in Chapter 1, the driving task undergoes intensive changes with the development of technology. As more and more in-vehicle devices are continually included in the vehicle, the driver engages often in other attention-demanding tasks which are not related to the primary driving task. This may lead to driver mental overload (De Waard, 1996; Sayer et al., 2007; Lenneman and Backs, 2009). On the other hand, in some scenarios, e.g. in the highway, the use of automation technology, e.g. ACC, may lead to a state of underload, resulting in driver fatigue and low vigilance (Brookhuis et al., 2001; May & Baldwin, 2009). Such workload states are supposed to be dangerous for traffic safety and a potential solution for both tendencies is to dynamically regulate the information flow from in-vehicle devices or change the level of automation according to the driver's mental states, namely adaptive task allocation, which has been proposed for decades in the area of human-machine interaction. However, for such a concept, a computational model for workload assessment as well as theoretical and empirical evidence supporting the implementation of ATA in the driving context would definitely be essential. The whole dissertation intensively concentrated on these two points and accordingly a series of studies have been conducted.

In the first experiment, the changes in EEG parameters with the task load including ERP components (P300) and frequency components (theta, alpha, and beta) were investigated with both single and dual task paradigms in the LCT, during which the task loads were manipulated in terms of speed settings and combination of a two-paced PASAT conditions. Results indicated that the amplitude of the P300 attenuated with the task load in both single and dual task

paradigms. Such findings were consistent with previous results from various studies (Wickens et al., 1997; Schultheis and Jameson, 2004; Brookhuis & De Waard, 2010). They are also consistent with an earlier theory that the amplitude of the P300 reflects the amount of attentional resources allocated to the task performance (Johnson, 1986; Polich, 2004). Alternatively, the changes in EEG frequency activities were investigated as well. The results revealed that as the speed increased, there were significant increases in the theta power but with a significant decrease in the alpha power. As the PASAT was added to the driving task[28], a significant decrease in alpha power was found while there was no difference in the theta power. When comparing the single PASAT conditions with the dual task conditions[29], there was a significant increase in theta power and a significant decrease in alpha power in dual task conditions compared to single PASAT conditions. The findings that an increased task load may lead to increased frontal theta activity and suppressed parietal alpha activity were consistent with previous studies (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Wilson, 2001, 2002; Hagemann, 2008). The comparison of ERP components with frequency components indicated that EEG frequency parameters demonstrated a higher classification accuracy in an Adaboost-based workload classification. These results address the demerits of using ERPs for the workload detection. Since ERPs are often obtained by averaging a number of trials for improving the signal-noise-ratio, it has been questioned whether EPRs based on a few trials let alone a single trial can reliably discriminate between different levels of workload (De Waard, 1996; Baldwin et al., 2004). Given the advantages and disadvantages of ERPs and EEG frequency components, in this dissertation the latter is concluded to be preferable for driver mental workload detection.

In the second experiment, we continuously investigated the changes in the EEG

[28] Comparison of single driving condition (80 km/h) with dual task conditions, driving+p5, driving+p3.
[29] Comparison of the single PASAT conditions p5, p3 with dual task conditions, driving+p5, driving+p3.

band powers with task load in LCT combined with the n-back task. The task load levels were manipulated in two dimensions, the driving task load and working memory load, with each containing three task load conditions. Unlike the first experiment, this experiment focused on the EEG spectrum modulation induced by task combination, in which the driver's mental workload was attributed simultaneously to different factors. The results demonstrated that frontal theta activity showed significant increases in the working memory load dimension, but no differences in theta power were found with the driving task load dimension. However, significant decreases in parietal alpha activity were found when the task load was increased in both dimensions. The increased theta activity and decreased alpha activity when the working memory load was increased were consistent with previous findings (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Smith, 2001). The decreased alpha power with the driving task load was also consistent with previous studies on the driving context (Hagemann, 2008) or on other applied settings (Wilson, 2001, 2002b). Additionally, task-related differences were also found. The driving task load contributed more to changes in alpha power, whereas the working memory load contributed more to changes in theta power. These findings suggested that different frequency bands may represent different neural resources. Such assumption can explain the inconsistency in previous results related to the EEG spectrum modulation induced by different types of tasks. Moreover, these findings suggested that any single EEG spectrum parameter could not completely represent the operator's workload states. Such a suggestion raises the need for the development of a combined metrics for workload representation using various EEG parameters.

Following the second experiment, a computational model for workload quantification was proposed based on the analysis of the short-term modulations of EEG parameters. Firstly, the p-quantiles were used to determine the individual-specific coefficients. Such an approach can customize the individual

differences. The positive regressed theta power coefficient and negative regressed alpha power coefficient are in line with previous contention about changes in these two parameters with the task load. Secondly, the LFM was used to transfer the theta and alpha power into scales ranging from 0 to 1, which could represent the workload levels. The results indicated that the workload scores using the LFM increased with the task load and showed improved correlations with other workload parameters more than the single theta or alpha powers. Additionally, the classification of driver mental workload using the workload thresholds [30] demonstrated that there are rare presences of dissociation between the task load and workload, e.g. a small percentage of the high (low) workload was observed in the low (high) task load condition. Such findings were consistent with previous theory about the occasional occurrence of dissociation between workload and task load (Hancock, 1996; Parasuraman & Hancock, 2001). The biggest merit of the proposed model is that it can reliably reflect the workload states and allows for an easy definition of the workload thresholds, which are adjustable as well as applicable to various cases.

In the third experiment, the EEG-driven adaptive interface was integrated into the driver-vehicle-system in a simulated driving environment. The EEG estimated workload states were used to immediately regulate the driving speed, which is thought to be an important contributor to the driver task load (Senders et al., 1967; Lei et al., 2009a; Fitzpatrick et al., 2010). This study was actually an application of the proposed theoretic DWM-ATA model and the computational workload quantification model. The results demonstrated that the computational model for workload quantification proposed in this dissertation can effectively regulate the task load into a moderate level. Most importantly, it was also found that the psychophysiologically driven adaptive task allocation can improve operator

[30] Two workload thresholds 0.2 and 0.8 were used to classify the whole range of mental workload scores into low, moderate, and high ranges.

performance and reduce operator workload in a high task load condition. These findings are in line with previous studies (Prinzel et al., 2000; Wilson & Russell, 2007; Kohlmorgen et al., 2007). Still, a key contribution of this study is that it provides new evidence supporting the applicability of the ATA in the driving context, what was rarely a concern in previous studies (Kohlmorgen et al., 2007).

The whole dissertation focused on workload issues for improving the interaction of driver-vehicle-environment. The results from the experimental studies indicated that the EEG signal provides useful information for inferring driver workload states. The proposed LFM, based on the EEG spectrum components, was proven to be effective for representing workload and can be used to regulate the task flow to the driver. Most importantly, this dissertation provides both theoretic theory and empirical evidence in support of driver adaptive task allocation.

7.2 Originality, innovations, and new findings

A discussion of some of the originalities, innovations, and novel findings of this dissertation is now in order.

Firstly, the DWM-ATA is innovative in and of itself. Parasuraman (1999) had proposed a workload-matched model for adaptive task allocation. However, this model only interprets the case in which workload is linearly associated with task demand and fails to provide a solution for the situation where the workload is dissociated from the task load. Many studies reported the existence of an occasional dissociation between the workload and task demand (Hancock, 1996; Parasuraman & Hancock, 2001), e.g. a high task load does not necessarily induce a high workload level. Our findings from the second and third experiment, that there was a low percentage of high workload periods in the low task condition as well as a low percentage of low workload periods in the high task condition,

provide conclusive evidence on this point. In such situations, the workload-matched model may be not applicable. Therefore, there is actually a need for a general model that can handle both association and dissociation conditions. From such a perspective, the DWM-ATA model is an effective extension of Parasuraman's model (Parasuraman, 1999). The DWM-ATA model provides a solution for both association and dissociation of mental workload and task demand. This model is particularly practical when integrating the psychophysiological interfaces into a closed-loop for adaptive aiding designs, since temporal changes in psychophysiologically estimated workload (or instantaneous workload[31]) may frequently be dissociated from task demand. In the driving context, the DWM-ATA depicts the interactions of three components in the transportation loop, namely, driver-vehicle-environment. As discussed in Chapter 5, the driving task demand imposed by the environment or in-vehicle task can be evaluated using various methods (Musa and Navin, 2005; Horrey and Lesch, 2009). Meanwhile, the driver's psychophysiology can be used for mental workload estimation. By mapping the task demand and workload with the DWM-ATA model, a corresponding adaptive strategy can be determined.

Secondly, our finding, that changes in working memory load contributed more to frontal theta power changes but less to parietal alpha power changes when compared to the driving task load, was new. Although the EEG spectrum modulations with the working memory tasks have been well-studied and there are also a number of studies investigating the EEG band power changes in complex tasks, e.g. Multiple Attributes Task Battery (Smith et al., 2001), there are only few studies that provide a clear comparison of the similarities and differences in EEG spectrum parameters induced by different workload dimensions as well as their combinations. The findings concerning task-dependent or factor-dependent differences in EEG parameters suggest that changes in different EEG parameters

[31] Xie & Salvendy, 2000b.

may represent separate neural resources utilities associated with task performance. Such an assumption further suggests the inefficiency of using a single parameter to index workload while conjuring up the necessity of having a computational model that can combine various EEG parameters and thus enable indexing workload even across tasks and individuals. Additionally, our discovery that a reduced variation of the short-term alpha is associated with an increased task load was also new. There are indeed quite a number of studies purporting to show that the attenuation of the alpha power can be observed when the workload is increased (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Smith, 2001). But few of these studies concern themselves with the short-term variation in the alpha power. According to our finding, the variation (e.g. standard deviation) in the alpha power would also be a reliable parameter for workload assessment.

Thirdly, the LFM for workload quantification is original. As discussed in Chapter 4, the direct transformation of the EEG parameters, e.g. theta/(alpha+beta) (Pope et al., 1995; Prinzel et al., 2000), and machine learning methods, e.g. ANN (Wilson, 2003), both have limitations for the workload quantification. For instance, the index theta/(alpha+beta) is strongly individual-dependent and the definition of workload thresholds for adaptive aiding design is difficult using such methods. On the other hand, the use of machine learning methods could indeed classify the task load with high accuracy. However, such classification has a 'black box' effect and is even doubtful, since it is prone to being involved with other cognition components rather than workload. The LFM is proposed to customize the individual differences (with the p-quantile) and to transfer the EEG parameters towards indices ranging from 0 to 1 for workload representation. Such an approach allows for an easy definition of the workload thresholds and even comparable workload quantification across tasks. Most importantly, this model is in line with the general tendencies of changes in EEG parameters with the

workload, e.g. the positive theta coefficient and negative alpha coefficient respectively reflect the proportional and inversely proportional relationship of these two parameters with the workload, which is consistent with the majority of previous findings (Gundel & Wilson, 1992; Sterman et al., 1994; Gevins et al., 1997, 1998; Smith, 2001). On this point, results with LFM are more reliable and convincing for indexing mental workload.

The three points discussed above constitute the main contributions of this dissertation to the topic at hand. The DWM-ATA model as well as the LFM would be valuable for transferring the concept of adaptive task allocation into application. These have a great potential for application in both laboratory and real-world settings. For example, the psychophysiology-driven interfaces contribute to the development of a vehicle workload manager which would allow for regulating the information flow to the driver, e.g. changing the availability of in-vehicle cellphones, triggering the adaptive automation, or changes in the level of automation. The driving context is but one application case. These works are applicable to many other domains.

7.3 Outlook for future research

The outcomes of this dissertation as well as the large body of related studies have set a fundamental basis for driver adaptive task allocation. However, to take these into real traffic application, there is still a long way to go. For future work(s) on this topic, particular attention should perhaps be paid to the following aspects:

First of all, a general task demand metrics for driving task is needed. The DWM-ATA model clearly presumes a clear definition of the task demand. That is, the task demands from the road situation as well as from the in-vehicle tasks have to be quantified independently of the individual experience. Indeed, a number of

previous studies have focused on this and various models and methods have been proposed (Baldwin & Coyne, 2003; Musa and Navin, 2005; Horrey and Lesch, 2009). For example, a rate of information processing demand (RID) combining road complexity and driving speed in terms of time constraint was proposed (Musa and Navin, 2005). Horrey and Lesch (2009) also reported scales quantifying the demands of different road sections and in-vehicle tasks. However, as far as we know, these studies only focused on particular driving contexts (e.g. the degree of curve, traffic density, or in-vehicle tasks) and there is no agreed-upon metrics to allow for modelling the general driving task demands in a dynamic pattern across various driving contexts. Therefore, further work towards building a general task demand metrics would be beneficial for the application of the DWM-ATA model.

Secondly, the improvement of the LFM is needed. The LFM projects the theta power and alpha power into unified workload indices that ranged from 0 to 1. This allows for an easy definition and adjustment of workload thresholds in various application cases. However, the criteria for the thresholds definition would be beneficial in such cases to avoid over- or less-fitting of the adaptive task allocation. Furthermore, in the present study, we focused only on frontal theta and parietal alpha. However, theta and alpha from other electrode sites also definitely provide valuable information on neural resource utilities. Additionally, the beta power is reported to be sensitive to task load changes, even though changes in this band from previous studies are inconsistent or even contrary (Wilson and O'Donnell, 1998; Dussault et al., 2005; Brookings et al., 1996; Doyle et al., 2009). The selection of the EEG parameters has to be linked with the underlying neural mechanisms behind mental workload. This may raise some interesting questions, such as: 'could the LFM be used when more features are involved?', or 'how could it be improved?' These need further investigations.

Thirdly, validation of the DWM-ATA in field tests is also needed. Three experiments in this dissertation were conducted in the simulated driving environment. Such settings are of course far-off from the real traffic situations. Hence, integrating adaptive task allocation in the vehicle definitely needs further empirical evidence from the field driving test or real traffic situations. Unfortunately, so far, only very few studies have dealt with the topic of the field driving test (Kohlmorgen et al., 2007). This is clearly insufficient.

Finally, the psychophysiology-driven adaptive automation in the driving context also needs further investigation. Although adaptive automation has been proposed for decades, the investigation of adaptive automation in the driving context is rare. The majority of related studies actually focus on the human factor issues in the interaction between driver and a highly automated vehicle (Rauch et al., 2009, 2010; Schieben et al., 2011). Kollmorgen's study (2007) focused actually on the adaptive modulation of the information flow to the driver according to the driver's mental states. In our third experiment, the modulation of the driving speed was actually some kind of a modulation of information flow, since speed in a sense reflects the rate of information processing. Therefore, as far as we know, there is actually no extant studies that truly integrate the psychophysiology-driven interfaces for adaptive automation (e.g. modulating the LOA) in either laboratory or real-traffic situation.

References

Aasman, J., Mulder, G. & Mulder, L.J.M. (1987). Operator effort and the measurement of heart-rate variability. *Human Factors*, 29, 161-170.

Akshoomoff NA, Courchesne E, Townsend J (1997): Attention coordination and anticipatory control. *International Review of Neurobiology*, 41, 575-598.

Averty, P., Collet, C., Dittmar, A., Athènes, S., Vernet-Maury, E. (2004). Mental workload in air traffic control: an index constructed from field tests. *Aviation, Space, and Environmental Medicine*. 75(4), 333-41.

Backs, R. W., & Seljos, K. A. (1994). Metabolic and Cardiorespiratory Measures of Mental Effort: The Effects of Level of Difficulty in a Working Memory Task. *International Journal of Psychophysiology*, 16 (1), 57-68.

Baddeley, A.D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 302, 311-324.

Bailey, B.P., Iaqbal, T. (2008). Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transaction on Computer-Human Interaction*, 14(4), 1-28.

Baldauf D, Burgard E, Wittmann M (2009). Time perception as a workload measure in simulated car driving. *Applied Ergonomics*, 40, 929-935.

Baldwin, C. L., and Coyne, J. T. (2003). Mental workload as a function of traffic density in an urban environment: Convergence of physiological, behavioral, and subjective Indices. In *Proceedings of the Human Factors in Driving Assessment 2003 Symposium* (July 2003. Park City, Utah).

Baldwin, C. L., Freeman, F. G., & Coyne, J.T. (2004). Mental Workload as a Function of Road Type and Visibility: Comparison of Neurophysiological, Behavioral, and Subjective Indices. In *proceedings of Human Factors and Ergonomics Society Annual Meeting* (pp:2309-2313), Santa Monica.

Baldwin, C. L., & Coyne, J. T. (2005). Dissociable Aspects of Mental Workload: Examinations of the P300 ERP Component and Performance Assessments. *Psychologia*, 48, 102-119.

Banquet, J.P., & Grossberg, S. (1987). Probing cognitive processes through the structure of event-related potentials during learning: an experimental and theoretical analysis. *Applied Optics*, 26 (23), 4931-4946.

Beatty, J. (1982). Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. *Psychological Bulletin*, 91(2), 276-292.

Beeli, G., Koeneke, S., Gasser K., and Jancke L.(2008). Brain stimulation modulates driving behaviour. *Behavioral and Brain Functions*, 4, 34.

Berka, C., Levendowski, D.J., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P.D., & Craven, P.L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat Space Environmental Medicine*, 78(5, Suppl.), B231-B244.

Birbaumer, N., & Schmidt, R.F. (1996). *Biologische Psychologie* (3rd ed.). Heidelberg: Springer-Verlag.

Bornemann, E. (1942). Untersuchungen über den Grad der geistigen Beanspruchung.II.Teil: Praktische Ergebnisse. *Arbeitsphysiologie*, 12, 173-191.

Bower J.M. (1997): Control of sensory data acquisition. *International Review of Neurobiology*, 41, 489-513.

Brookhuis, K.A., De Vries, G. & De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention*, 23, 309-316.

Brookhuis, K.A., & de Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics*, 36 (9), 1099-1110.

Brookhuis, K.A., & De Waard, D. (2001) Assessment of Drivers' Workload: Performance and Subjective and Physiological Indexes. In P. A. Hancock

and P. A. Desmond (eds.). *Stress, Workload, and Fatigue* (pp. 321-333). Erlbaum Associates, Mahwah, N.J.

Brookhuis, K.A. De Waard, D., & Janssen. W. H. (2001). Behavioural impacts of Advanced Driver Assistance Systems - an overview. *European Journal on Transportation and Infrastructure Research*, 1(3), 245-253.

Brookhuis K.A., van Driel C.J., Hof T., van Arem B., & Hoedemaeker M. (2009) Driving with a Congestion Assistant; mental workload and acceptance. *Applied Ergonomics*, 40(6), 1019-1025.

Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis and Prevention*, 42, 893-903.

Brookings, J.B., Wilson G.F., & Swain, C.R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42(3), 361-377.

Buzsáki, G (2002). Theta oscillations in the hippocampus. *Neuron*, 33 (3): 325-40.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, 42(3), 249-268.

Cain, B. and Defence research and development Toronto (Canada). (2007). *A Review of the Mental Workload Literature*.
Retrieved from:
http://ftp.rta.nato.int/public//PubFullText/RTO/TR/RTO-TR-HFM-121-PART-II///TR-HFM-121-Part-II-04.pdf

Caird, J. K., Scialfa, C. T., Ho, G., & Smiley, A. (2005). A meta-analysis of driving performance and crash risk associated with the use of cellular telephones while driving. In *Proceedings of the third international driving symposium on human factors in driver assessment, training and vehicle design* (pp. 478-485). University of Iowa Public Policy Center.

Caleefato, C., Montanari R., & Tango, F. (2007). Advanced Drivers Assistant Systems in Automation. *Lecture Notes in Computer Science*, 4551, 768-777.

Calhoun, V. D., Pekar, J. J. McGinty, V. B., Adali, T., Watson, T. D., & Godfrey, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Human Brain Mapping*, m16, 158-167.

Calhoun, V. D. (2007). Functional Magnetic Resonance Imaging (fMRI): Advanced Methods and Applications to Driving. In R. Parasuraman, & M. Rizzo (Eds.). *Neuroergonomics: The Brain at Work* (pp. 51-64). New York: Oxford University Press.

Cantero J.L., Atienza M., Stickgold R., Kahana M.J., Madsen J.R., Kocsis B. (2003). Sleep-dependent theta oscillations in the human hippocampus and neocortex. *Neuroscience*, 23 (34), 10897-10903.

Carsten O. (2007) From driver models to modelling the driver: what do we really need to know about the driver? In P.C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments*, Springer, London, pp. 105-120.

Castor M, Hanson E, Svensson E, Nählinder S, LeBlaye P, MacLeod I, Wright N, Alfredson J, Ågren L, Berggren P, Juppet V, Hilburn B, & Ohlsson K. (2003). *GARTEUR Handbook of mental workload measurement*, GARTEUR TP 145.

Caton, R. (1875). The electric currents of the brain. *British Medical Journal*, 2, 278

Cnossen, F., Rothengather, T., & Meijman, T. (2000). Strategic Changes in Task Performance in Simulated Car Driving as and Adaptive Response to Task Demands. *Transportation Research Part F*, 3, 123-140.

Colle, H.A., & Reid, G.B. (1999). A framework for mental workload research and applications using formal measurement theory. *International Journal of Cognitive Ergonomics*, 1(4), 303-313.

Collet C, Guillot A, & Petit C. (2010). Phoning while driving I: a review of epidemiological, psychological, behavioural and physiological studies. *Ergonomics*, 53(5), 589 - 601.

Collins, M., Schapire, R., & Singer Y. (2002). Logistic regression, Adaboost, and Bregman distances. *Machine Learning*, 48, 1-3.

Comstock, J. L., & Arnegard, R. J. (1992). *The Multi-attribute Task Battery for human operator workload and strategic behavior research*. Technical Report 104174. Hampton, VA:NASA Langley Research Center.

Cooper, N.R., Burgess, A.P., Croft, R.J., & Gruzelier, J.H. (2006). Investigating evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task. *Neuroreport*, 17, 205-208.

Corwin, W.H. Sandry-Garza, D.L., Biferno, M.A., & Boucek, G.P. (1989). Assessment of crew workload measurement methods, techniques and procedures. Volume II - *Guidelines for the use of workload assessment techniques in aircraft certification*. Dayton, OH, Wright-Patterson Air Force Base, Wright Research and Development Centre, Cockpit Integration Directorate: 51.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA.

Delorme, A. & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics. *Journal of Neuroscience Methods*, 134, 9-21.

De Waard, D.(1996). *The measurement of drivers' mental workload*. Ph.D. Thesis. University of Groningen, Traffic Research Centre, Haren, Netherlands.

Donchin, E., Ritter, W., & McCallum, C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, &

S.Koslow (Eds.), *Brain Event-Related Potentials in Man* (pp. 349-411). New York: Academic Press.

Donges, E. (1999). A Conceptual Framework for Active Safety in Road Traffic, *Vehicle System Dynamics*, 32(2), 113-128.

Dooley, C. (2009). The Impact of Meditative Practices on Physiology and Neurology: A Review of the Literature. *Scientia discipulorum*, 35-59.

Doyle, M. J., Gould, J. E., Kass, S. J., Raj, A. K., Andrasik, F., & Higgins, J. P. (2009). Psychophysiological Correlates of Cognitive Workload during a Satellite Management Decision -Training Task. In: *proceedings of Human Systems Integration Symposium*, Annapolis, MD, USA.

Duncan-Johnson C.C., Donchin E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology*. 14(5), 456-467.

Dussault C, Jouanin JC, Philippe M, Guezennec CY. (2005). EEG and ECG changes during simulator operation reflect mental workload and vigilance. *Aviation, Space, and Environmental Medicine*, 76(4), 344-351.

Eggemeier, F.T., Wilson, G.F., Kramer, A.F. & Damos, D.L. (1991). Workload assessment in multi-task environments. In D.L. Damos (Ed.). *Multiple-task performance* (pp. 207-216). London: Taylor & Francis.

Eulitz C., Maess B., Pantev C., Friederici A.D., Feige B., & Elbert T. (1996). Oscillatory neuromagnetic activity induced by language and non-language stimuli. *Cognitive Brain Research*, 4, 121-132.

Fitzpatrick, K., Chrysler, Susan., Park, E. S., Nelson, A.A., Robertson, J.A., and Iragavarapu, V. (2010). *Driver workload at higher speeds*. Report Number: 0-5911-1. Texas Transportation Institute, College Station, TX. February 2010.

Fowler B. (1994). P300 as a measure of workload during a simulated aircraft landing task. *Hum Factors.* 36(4), 670-683.

Freeman F.G., Mikulka P.J., Prinzel L.J., Scerbo M.W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, 50(1), 61-76.

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28(2).

Frisk, V., & Milner B. (1990). The relationship of working memory to the immediate recall of stories following unilateral temporal or frontal lobectomy. *Neuropsychologia*, 28(2), 121-35.

Freund Y. & Schapire R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139.

Funk, K., Lyall, B., Wilson, J., Vint, R., Miemcyzyk, M., Suroteguh, C., et al. (1999). Flight deck automation issues. *International Journal of Aviation Psychology*, 9, 125-138.

Fukami, T., Shimada, T., Ishikawa, F., Ishikawa, B., & Saito, Y. (2010). Evaluation of brain aging by EEG analysis of photic driving response. In *proceedings of 2010 IEEE/ICME International Conference on Complex Medical Engineering (CME),* 161-164, Gold Coast, QLD,

Furedy, John J. (1987). Beyond Heart Rate in the Cardiac Psychophysiological Assessment of Mental Effort: The T-Wave Amplitude Component of the Electrocardiogram. *Human Factors*, 29(2), 183-194.

Fu, S., & Parasuraman, R. (2007). Event-related brain potentials in neuroergonomics. In R. Parasuraman, & M. Rizzo (Eds.) *Neuroergonomics: The Brain at Work* (pp. 32-50). New York: Oxford University Press.

Fu, S., Fedota, J., Greenwood, P. M., & Parasuraman, R. (2010a). Early interaction between perceptual load and involuntary attention: an event-related potential study. *Neuroscience Letters*, 468, 68-71.

Fu, S., Fedota, J., Greenwood, P. M., & Parasuraman, R. (2010b). Dissociation of visual C1 and P1 components as a function of attentional load: An event-related potential study. *Biological Psychology*, 85(1), 171-178.

Gevins, A., Smith, M.E., McEvoy, L., Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of difficulty, type of processing, and practice. *Cerebral Cortex*, 7, 374-385.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S. and Du, R. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, 40, 79-91.

Gevins, A., and Smith, M. E. (1999). Detecting transient cognitive impairment with EEG pattern recognition methods. *Aviation, Space, and Environmental Medicine*, 70, 1018-1024.

Gevins A., & Smith M. E. (2000). Neurophysiological Measures of Working Memory and Individual Differences in Cognitive Ability and Cognitive Style. *Cerebral Cortex,* 10(9), 829-839.

Giger-Mateeva V.I., Riemslag F.C., Reits D., Schellart N.A., & Spekreijse H. (1999). Isolation of late eventrelated components to checkerboard stimulation. *Electroencephalogr Clin Neurophysiol Suppl*, 50, 133-149.

Gopher, D. and Donchin, E. (1986). Workload: An examination of the concept. In K. Boff, and L. Kaufman, (Eds.): *Handbook of Human Perception and Performance*, NewYork: John Wiley, 2(41), 41-49.

Gray, C. M., König, P., Engel, A. K. and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338, 334-337.

Graydon, F X, Young, R, Benton, M D, II, R J Genik, Posse, S, Hsieh, L. and Green, C F (2004). Visual event detection during simulated driving: identifying the neural correlates with functional neuroimaging, *Transportation Research Part F: Psychology and Behavior*, 7(4-5), 271-286.

Greef, T.E.d., Dongen, C.J.G.v., Grootjen, M., Lindenberg, J.: (2007). Augmenting Cognition: Reviewing the Symbiotic Relation Between Man and Machine. In: Schmorrow, D.D., Reeves, L.M. (eds.). *Foundations of Augmented Cognition* (pp. 439-448), Springer, Heidelberg.

Green. P.(2004). Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions. In *proceedings of Convergence International Congress & Exposition on Transportation Electronics*, October 2004, Detroit, MI, USA

Grimes, D., Tan, D. S., Hudson, S. E., Shenoy, P., and Rao, R. P. (2008). Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI 08. ACM, New York, 835-844.

Groeger, J.A. (2002). Trafficking in cognition: applying cognitive psychology to driving. *Transportation Research Part F*, 5, 235-248.

Grossman, P. (1992). Respiratory and cardiac rhythms as windows to central and autonomic biobehavioral regulation: selection of window frames, keeping the panes clean and viewing the neural topography. *Biological Psychology*, 34, 131-161.

Gruzelier, J., Liddiard, D., Davies, L., & Wilson, L. (1990). Topographical EEG differences between schizophrenic patients and controls during neuropsychological functional activation. *International Journal of Psychophysiology*, 8, 275-282.

Gonzalez Gonzalez, J., Mendez Llorens, A., Mendez Novoa, A., Cordero Valeriano, J. J. (1992). Effect of acute alcohol ingestion on short-term heart rate fluctuations. *Journal of Studies on Alcohol*, 53, 86-90.

Gronwall, D. (1977). Paced Auditory Serial Addition Task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367-373.

Gundel A., & Wilson G.F. (1992). Topographical changes in the ongoing EEG related to the difficulty of mental tasks. *Brain Topography*, 5, 17-25.

Hagemann, K. (2008). *The alpha band as an electrophysiological indicator for internalized attention and high mental workload in real traffic driving*. Dissertation, Universitaet Duesseldorf, Germany.

Halgren E., Baudena P., Clarke J.M.; Heit G., Marinkovic K., Devaux B., Vignal J.-P., Biraben A. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. II. Medial, lateral and posterior lobe. *Electroencephalography and Clinical Neurophysiology*, 94, 229-250.

Hancock, P. A., Chignell, M. H., & Lowenthal, A. (1985). An adaptive human-machine system. In: *Proceedings of the 1985 IEEE International Conference on Systems, Man and Cybernetics* (pp. 627-630). New York: Institute of Electrical and Electronics Engineers.

Hancock, P. A., & Warm, J. S. (1989). A Dynamic Model of Stress and Sustained Attention. *Human Factors and Ergonomics Society*, 31(5), 519-537.

Hancock P. A. (1996). Effects of control order, augmented feedback, input device and practice on tracking performance and perceived workload. *Ergonomics*. 39(9), 1146-62.

Hart, S. G. and Staveland, L. E. (1988). Development of the NASA TLX: results of empirical and theoretical research. In: P.A. Hancock and N. Meshkati, *Human Mental Workload* (pp.139-183). North Holland, Amsterdam.

Healey J., & Picard R. (2005). Detecting stress during real world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6 (2), 156-166.

Hebert, R., & Lehmann, D. (1977). Theta bursts: An EEG pattern in normal subjects practising the transcendental meditation technique. *Electroencephalography and Clinical Neurophysiology*, 42, 397-405.

Heger D., Putze F., Schultz, T. (2010). Online Workload Recognition from EEG data during Cognitive Tests and Human-Computer Interaction. In *proceedings of 33rd Annual German Conference on Artificial Intelligence 2010*, Karlsruhe, Germany.

Heino, A., Van der Molen, H.H. & Wilde, G.J.S. (1996). Differences in risk experience between sensation avoiders and sensation seekers. *Personality and Individual Differences*, 20, 71-79.

Herrmann, C. S. & Knight, R. T., (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews*. 25, 465-476.

Horikawa, E., Okamura, N., Tashiro, M., Sakurada, Y., Maruyama, M., Arai, H. (2005). The neural correlates of driving performance identified using positron emission tomography. *Brain and cognition*, 58, 166-171.

Horrey, W.J. & Wickens, C.D. (2004). Cell phones and driving performance: A meta-analysis. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 2304-2308). Santa Monica, CA: Human Factors and Ergonomics Society.

Horrey, W.J., Lesch, M.F., & Garabet, A. (2009). Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task situations. *Journal of Safety Research*, 40(1), 7-12.

Hilburn, B., & Jorna, P.G.A. M. (2001). Workload and Air Traffic Control. In P. A. Hancock & P. E. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 384-394). Mahwah, NJ: Erlbaum.

Hillyard S.A., Hink R.F., Schwent V.L., Picton T.W. (1973). Electrical signs of selective attention in the human brain. *Science*, 182(108), 177-180.

Hillyard, S. A., Picton, T. W. and Regan, D.(1978). Sensation, perception and attention: Analysis using ERPs. In Callaway, E., Tueting, P. and Koslow, S.

H. (eds.). *Event-related potentials in man* (pp.223-347). Academic Press, New York.

Hillyard S.A., & Münte T.F. (1984). Selective attention to color and location: an analysis with event-related brain potentials. *Percept Psychophys*, 36(2), 185-98.

Hink, R. F., Van Voorhis, S. T., Hillyard, S. A. and Smith T. S. (1977). The division of attention and the human auditory evoked potential. *Neuropsychologia*, 15(4-5), 597-605.

Hjorth, B. (1975). An on-line transformation of EEG scalp potentials into orthogonal source derivations. *Electroencephalography and Clinical Neurophysiology*, 39, 526-530..

Hjortskov N, Rissen D, Blangsted AK, Fallentin N, Lundberg U, Sogaard K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology*, 92(1-2), 84-89.

Ho, C., & Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. Aldershot, England: Ashgate.

Honal M., & Schultz T. (2008). Determine task demand from brain activity. In *Proceedings of the 3rd International Conference on Bio-inspired Systems and Signal Processing*, Funchal, Madeira-Portugal.

Humphrey D. G. & Kramer A. F. (1994). Toward a psychophysiological assessment of dynamic changes in mental workload. *Human Factors*, 36(1), 3-26.

Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y., Toi, S., and Ishihara, T. (1994). Potential flow of frontal midline theta activity during a mental task in the human electroencephalogram. *Neuroscience Letters*, 169, 145-148.

Insko, B. E. (2003). Measuring presence: Subjective, behavioral and physiological methods. In G. Riva, F. Davide, & W. A. IJsselsteijn (Eds.), *Being There:*

Concepts, effects and measurement of user presence in synthetic environments. Amsterdam: Ios Press.

Ishii R, Shinosaki K, Ukai S, Inouye T, Ishihara T, Yoshimine T, Hirabuki N, Asada H, Kihara T, Robinson SE, Takeda M. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport*, 10(4), 675-9.

ISO. (2010). Road vehicles - *Ergonomic aspects of transport information and control systems -Simulated lane change test to assess driver distraction*. ISO/TC 22/SC 13 N WG8 N579, DIS 26022.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 371-375.

Jennings, J.R., Stringfellow, J.C. & Graham, M. (1974). A comparison of the statistical distributions of beat-by-beat heart rate and heart period. *Psychophysiology*, 11, 207-210.

Jensen, O., Kaiser, J., & Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends in Neurosciences*, 30, 317-324.

Jex, H.R. (1988). Measuring mental workload: problems, progress, and promises. In Hancock, P.A., & Meshkati, N. (Eds.), *Human mental workload* (pp. 5 - 40). Amsterdam: North-Holland.

Johnson R Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23(4), 367-84.

Jorna, P.G.A.M. (1992). Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. *Biological Psychology*, 34, 237-257.

Kaber, D. B., & Riley, J. (1999). Adaptive automation of a dynamic control task based on secondary task workload measurement. *International Journal of Cognitive Ergonomics*, 3, 169-187.

Kaber, D.B., Wright, M.C., Prinzel, L.J., Clamann, M.P. (2005). Adaptive Automation of Human-Machine System Information-Processing Functions. *Human Factors*, 47(4), 730-741.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Kalsbeek, J.W.H. & Ettema, J.H. (1963). Scored regularity of the heart rate and the measurement of perceptual load. *Ergonomics*, 6, 306.

Kamada, T., Miyake, S., Kumashiro, M., Monou, H., & Inoue, K. (1992). Power spectral analysis of heart rate variability in Type As and Type Bs during mental workload. *Psychosomatic Medicine,* 54, 462-470.

Karavidas MK, Lehrer PM, Lu SE, Vaschillo E, Vaschillo B, Cheng A.(2010). The effects of workload on respiratory variables in simulated flight: A preliminary study. *Biological Psychology*, 84(1), 157-60.

Kirchner, W. K. (1958), Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352-358

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). 'Paradoxical' alpha synchronization in a memory task. *Cognitive Brain Research*, 7, 493-501.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Research Reviews*, 53, 63-88.

Knight W.E. J. & Rickard, N.S. (2001). Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure, and heart rate in healthy males and females. *Journal of Music Therapy*, 38, 254-272.

Kohlmorgen, J., Dornhege, G., Braun, M. L., Blankertz, B., Müller, K. R., Curio, G., Hagemann, K., Bruns, A., Schrauf, M., & Kincses, W. E. Improving Human Performance in a Real Operating Environment through Real-Time Mental Workload Detection. In: G. Dornhege; J. del R. Millán; T. Hinterberger; D. McFarland; K.-R. Müller (Eds.). *Towards*

Brain-Computer Interfacing (pp.409-422). Cambridge, Mass.: MIT Press, Cambridge.

Kornhuber, H.H., Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Arch*, 284, 1-17.

Kotchoubey, B. (2006). Event-related potentials, cognition, and behavior: A biological approach. *Neuroscience and Biobehavioral Reviews*, *30*(1), 42-65.

Kramer, A.F. (1991). Physiological metrics of mental workload: a review of recent progress. In D.L. Damos (Ed.), *Multiple-task performance*. (pp. 279-328). London: Taylor & Francis.

Kramer, A., & Parasuraman, R. (2007). Neuroergonomics-application of neuroscience to human factors. In J. Caccioppo, L. Tassinary, L., & G. Berntson, (Eds.) *Handbook of Psychophysiology* (2nd ed). New York: Cambridge University Press.

Kubota, Y., Sato, W., Toichi, M., Murai, T., Okada, T., Hayashi, A. et al. (2001). Frontal midline theta rhythm is correlated with cardiac autonomic activities during the performance of an attention demanding meditation procedure. *Cognitive Brain Research*, 11, 281-287.

Kubová Z., Kremlábek J., Szany J., Chlubnová J., Kuba M. (2002). Visual event-related potentials to moving stimuli: normative data. *Physiological Research,* 51, 199-204.

Kutas M, McCarthy G, Donchin E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, 197(4305), 792-795.

Lee, D.H. & Park, K.S. (1990). Multivariate analysis of mental and physical load components in sinus arrhythmia scores. *Ergonomics*, 33, 35-47.

Lee , J. H., & Tan, D.S. (2006). Using a low-cost electroencephalograph for task classification in HCI research, In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, Montreux, Switzerland.

Lee J. D. & Seppelt. B. D. (2009). Human Factors in Automation Design. Shimon Y. Nof (Ed.): *Springer Handbook of Automation* (pp. 417-436). Springer 2009.

Lehrer, P.M., Karavidas, M., Lu, S., Vaschillo, E., Vaschillo, B., Karavidas, P. (2008). Contribution of Psychophysiological Measures to Evaluating Flight Task Workload Under Simulated Conditions.
Retrieved from:
http://www.tc.faa.gov/logistics/grants/pdf/2004/04-G-004.pdf

Lei, S., Welke, S., Roetting, M. (2009a). Representation of Driver's Mental Workload in EEG Data. In: D. de Waard, J. Godthelp, F.L. Kooi, and K.A. Brookhuis (Eds.). *Human Factors, Security and Safety* (pp. 285-294) Maastricht, the Netherlands: Shaker Publishing.

Lei, S. Welke, S., Roetting, M.(2009b). Driver mental workload assessment using EEG data in a dual task paradigm. In *Proceedings of Enhanced Safety of Vehicle 2009*. Stuttgart, Germany.

Lei, S. Welke, S., Roetting, M. (2009c). EEG Band Power for Driver Mental Workload Monitoring. In *Proceedings of 17th Congress of International Ergonomics Association* (IEA2009). Beijing, China.

Lei, S. Welke, S., Roetting, M. (2009d). A comparison of Classification for Driver Mental Workload Using ERP and Band Power Parameters. In *proceedings of 8. Berliner Workshop of Human-Machine-System*, Berlin, Germany.

Lei, S, & Roetting, M. (2011). Influence of Task Combination on EEG Spectrum Modulation for Driver Workload Estimation. *Human Factors*, 53(2), 168-179.

Lei, S., Zhang, J., & Roetting, M. (2011). A Closed-loop for Driver Mental States Monitoring Based on EEG Signals. In D. de Waard, N. Gérard, L. Onnasch, R. Wiczorek, and D. Manzey (Eds.). *Human Centred Automation* (pp.351-365). Maastricht, the Netherlands: Shaker Publishing.

Lenneman, J. K., Shelley, J. R., & Backs, R. W. (2005). Deciphering psychological-physiological mappings while driving and performing a secondary memory task. *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Rockport*, ME, 493-498.

Lenneman, J. K., & Backs, R. W. (2009). Cardiac autonomic control during simulated driving with a concurrent verbal working memory task. *Human Factors*, 53(3), 404-418.

Leuthold, H., & Sommer, W. (1998) Postperceptual effects and P300 latency. *Psychophysiology*, 35 (1). 34-46.

Lin, Y., & Cai, H. (2009). A method for building a real-time cluster-based continuous mental workload scale. *Theoretical Issues in Ergonomics Science*, 10(6), 531 - 543.

Lopes da Silva, F. H., Vos, J. E., Mooibreck, J., and Van Rotterdam, A. (1980). Relative contributions of intracortical and thalamocortical processes in the generation of alpha rhythms, revealed by partial coherence analysis. *Electroencephalography and Clinical Neurophysiology*, 50, 449-456.

Lorenz, B., & Parasuraman, R. (2003). Human operator functional states in automated systems: The role of compensatory control strategies. In G. R. J. Hockey, A. W. Gaillard, & O. Burov (Eds.). *Operator Functional State: The Assessment and Prediction of Human Performance Degradation in Complex Tasks* (pp. 224-237). Amsterdam: IOS Press.

Luck, S. J., Hillyard, S.A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. (1994). Effect of spatial cueing on luminance detectability: Psychophysical and electrophysiological evidence for early selection.

Journal of Experimental Psychology: Human Perception and Performance, 20(4), 887-904.

Magliero A, Bashore TR, Coles MG, Donchin E. (1984). On the dependence of P300 latency on stimulus evaluation processes. *Psychophysiology*, 21(2), 171-186.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural In-formation Processing Systems*, 8, 145-151.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J. and Schwartz, P. J. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*, 93, 1043-1065.

Manly, T., Robertson, I.H., Galloway, M., and Hawkins, K.(1999). The Absent Mind: Further Investigations of Sustained Attention to Response. *Neruopsychologia*, 37(6), 661-670.

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Roesler (Ed.), Ergebnisse und Anwendungen der Psychophysiologie. Enzyklopädie der Psychologie: Vol. C/I/5. *Biologische Psychologie* (pp. 799 - 864). Göttingen: Hogrefe.

Mascord, D.J., & Heath, R.A. (1992). Behavioral and physiological indices of fatigue in a visual tracking task. *Journal of Safety Research*, 23, 19-25.

Mascord, D.J., Walls, J. & Starmer, G.A. (1995). Fatigue and alcohol: interactive effects on human performance in driving-related tasks. In L. Hartley (Ed.), *Fatigue and Driving: Driver Impairment, Driver Fatigue and Driving Simulation* (pp.189-205). London:Taylor & Francis.

Marshall, S.P. (2000). *Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity*. US Patent No. 6,090,051.

Marshall, S.P. (2002). The Index of Cognitive Activity: Measuring Cognitive Workload. In *Proceedings of the 7th Conference on Human Factors and Power Plants*. IEEE Computer Society, pp. 75-79.

Martens, M. H., & van Winsum, W. (2000). *Measuring distraction: the peripheral detection task*. Soesterberg, Netherlands: TNO Human Factors.

Mattes, S. (2003). The Lane Change Task as a Tool for driver Distraction Evaluation. *IHRA-ITS Workshop on Driving Simulator Scenarios*, Dearborn, Michigan.

Matthews R., McDonald N. J., Anumula H., Woodward J., Turner P. J., Steindorf M.A., Chang K., Pendleton J.M. (2007). Novel Hybrid Bioelectrodes for Ambulatory Zero-Prep EEG Measurements Using Multi-channel Wireless EEG System. *Human-Computer-Interaction*, 16, 137-146

May, J.F. & Baldwin, C.L. (2009). Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F*, 12, 218-224.

Maclver B. (2003). The Hippocampus. In: In: Antognini JF, Carstens EE, Raines DE (Eds). *Neural Mechanisms of Anesthesia* (pp. 149-167). New Jersey, Humana Press.

McCarthy G, Donchin E. (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science*, 211(4477), 77-80.

McDonald, B., & Ellis, N. C. (1975). Driver workload for various turn radii and speed. *Transportation Research Record*, 530, 18-29.

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). The impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record*, 2138 (12), 6-12.

Michon J. A. (1985). Acritical view of driver behaviour modelsWhat do we know, what should we do? In L. Evans & R. Schwing (Eds.). *Human behaviour and traffic safety* (pp. 485-525). New York: Plenum press.

Muckler, F.A. & Seven, S.A. (1992). Selecting performance measures: 'objective' versus 'subjective' measurement. *Human Factors*, 34, 441-455.

Mulder, G. & Mulder, L.J.M. (1980). Coping with mental workload. In S. Levine & H. Ursin (Eds.). *Coping and Health* (pp. 233-258). New York: Plenum Press.

Mulder, L.J.M. (1988). *Assessment of cardiovascular reactivity by means of spectral analysis*. PhD Thesis. Groningen: University of Groningen.

Mulder, L.J.M., De Waard, D., Brookhuis, K.A.(2004). Estimating mental effort using heart rate and heart rate variability. In: Stanton, N., Hedge, A., Hendrick, H.W., Brookhuis, K.A., Salas, E. (Eds.), *Handbook of Ergonomics and Human Factors Methods*. Taylor & Francis, London, pp. 201-208.

Müller, A., Schandry, R., Montoya, P. & Gsellhofer, B. (1992). Differential effects of two stressors on heart rate, respiratory sinus arrhythmia, and T-wave amplitude. *Journal of Psychophysiology*, 6, 252-259.

Murai, K., Hayashi, Y., Nagata, N., & Inokuchi, S. (2004). The mental workload of a ship's navigator using heart rate variability. *Interactive Technology & Smart Education*, 2, 127-133.

Musa, P. T. & Navin, F. P. T. (2005). Driver Attention Demand and Adaptive Information Management. In *proceedings of 2005 Annual Conference of the Transportation Association of Canada*, Calgary, Canada.

Nachreiner, F., P. Nickel, and C. von Ossietzky. (2003). Sensitivity and Diagnosticity of the 0.1-Hz Component of Heart Rate Variability As an Indicator of Mental Workload. *Human Factors*, 45, 575-590.

Nash AJ, Williams CS. (1982). Effects of preparatory set and task demands on auditory event-related potentials. *Biological Psychology*, 15(1-2), 15-31.

Nikolaev, A. R., Ivanitskii, G. A., & Ivanitskii, A. M. (1998). Reproducible EEG alpha-patterns in psychological task solving. *Human Physiology*, 24, 261-268.

Nunez P.L. (1981). *Electric Fields of the Brain: The Neurophysics of EEG* (pp. 484). Oxford University Press, New York.

Nykopp. T. (2001). *Statistical Modelling Issues for the Adaptive Brain Interface*, Master Thesis. Helsinki University of Technology, 2001.

O'Donnell, R.D. & Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff, L. Kaufman & J.P. Thomas (Eds.). *Handbook of perception and human performance. Volume II, cognitive processes and performance.* (pp 42/1-42/49). New York: Wiley.

Paas, F.G.W.C., Van Merriënboer, J.G.J. & Adam, J.J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79, 419-430.

Parasuraman, R. (1985). Event-related brain potentials and intermodal divided attention. In *proceedings of Human Factors and Ergonomics Society Annual Meeting*, 29, 971-975.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception* (pp. 43.1-43.49). New York: Wiley.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors,* 29, 695-706.

Parasuraman, R., Bahri, T., Deaton, J.E., Morrison, J.G., Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems.* (Progress Report No. NAWCADWAR-92033-60). Warminster: Naval Air Warfare Center.

Parasuraman, R., Molloy, R., & Singh, I.L. (1993). Performance consequences of automation induced complacency. *International Journal of Aviation Psychology*, 3, 1-23.

Parasuraman, R., Mouloua, M., & Molloy, R. (1994). Monitoring automation failures in human machine systems. In M. Mouloua, & R. Parasuraman

(Eds.). *Human performance in automated systems: current research and trends* (pp. 45-49). Hillsdale, NJ: Earlbaum.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39,* 230-253. (Reprinted in N. Cooke & E. Salas (Ed.) *Best of Human Factors: Thirty Classic Contributions to Human Factors.* Santa Monica, CA: Human Factors and Ergonomics Society, 2008).

Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive Aiding and Adaptive Task Allocation Enhance Human-Machine Interaction. in M.W. Scerbo and M. Mouloua, Ed., *Automation Technology and Human Performance,* New Jersey: Lawrence Erlbaum Associates, Inc., pp. 119-123.

Parasuraman, R., Wickens, C. D., & Sheridan, T. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics,* 30, 286-297.

Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of workload. In P. A. Hancock & P. E. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 305-320). Mahwah, NJ: Erlbaum.

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science,* 4, 5-20.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, 2, 141-161.

Parasuraman, R., Cosenzo, K., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, 21, 270-297.

Parsons, H.M. (1985). *Automation and the individual*: *Comprehensive* and comparative views. Human Factors, 27, 99-112.

Partin, D. L., M. F. Sultan, C. M. Thrush, R. Prieto, and Wagner. S. J. (2006). Monitoring Driver Physiological Parameters for Improved Safety. *SAE World Congress*, Detroit, Mich.

Patten C.J., Kircher A., Ostlund J., Nilsson L., Svenson O. (2006). Driver experience and cognitive workload in different traffic environments. *Accident Analysis & Prevention*, 38(5), 887-894.

Paus T., Koski L., Caramanos Z., Westbury C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport*, 9(9), R37-47.

Pauzie, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems*, 2(4), 315-322.

Pedrotti, M., Lei, S., Dzaack, J. & Rötting, M. (2011). A data driven algorithm for eyeblink detection in low-speed eyetracking protocols. *Behavior Research Methods*, 43(2), 372-383.

Pei, X., Zheng C., Xu J., & Bin G. (2005). Adaboost for improving classification of left and right hand motor imagery tasks. In *Proceedings of 2005 First International Conference on Neural Interface and Control*, 10-13.

Picton T.W., Bentin S., Berg P., Donchin E., Hillyard S.A., Johnson R. J., Miller G.A., Ritter W., Ruchkin D.S., Rugg M.D., Taylor M.J. (2000). Guidelines for using event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37, 127-152.

Pigeau, R., Hoffmann, R., Purcell, S., Moffitt, A. (1987). The Effect of Endogenous Alpha on Hemispheric Asymmetries and the Relationship of Frontal Theta to Sustained Attention. In: AGARD Conference No. 432. *Electric and Magnetic Activity of the Central Nervous System: Research*

and Clinical Applications in Aerospace Medicine. Trondheim: NATO, p. 20.1-20.16.

Polich, J. (1987). Stimulus complexity, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, 68, 311 -320.

Polich J. (1990). P300, Probability, and Interstimulus Interval. *Psychophysiology*, 27(4), 396-403.

Polich, J. (2004). Clinical application of the P300 event-related brain potential. *Physical Medicine and Rehabilitation Clinics of North America*, 15, 133-161.

Pope, A . T., Bogart, E. H. and Bartolome, D. (1995). Biocybernetic system evaluates indices of operator engagement. *Biological Psychology*, 40, 187-196.

Porges, S. W. and Byrne, E. A. (1992). Research methods for measurement of heart rate and respiration. *Biological Psychology*, 34, 93-130.

Posner, M. E., and Peterson, S. E. (1990). The attentional system of the human brain. *Annual Review of Neuroscience*, 13, 25-42.

Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In: A. D. Milner & M. D. Rugg (Eds.). *The neuropsychology of consciousness*. San Diego: Academic, 91-111.

Prinzel, L. J., Freeman, F. G., Scerbo, M.W., Mikulka, P. J., & Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology*, 10, 393-410.

Prinzel, L.J., Pope, A.T., Freeman, F.G., Scerbo, M.W., & Mikulka, P.J. (2001). Empirical analysis of EEG and ERPs for psychophysiological adaptive task allocation. *NASA Technical Report*, TM-2001-211016.

Putze F., Jarvis, J.P., Schultz, T. (2010), Multimodal Recognition of Cognitive Workload for Multitasking in the Car. In *proceedings of 20th International Conference on Pattern Recognition*, ICPR 2010, Istanbul, Turkey.

Raabe, M., Rutschmann, R.M., Schrauf, M., Greenlee, M.W. (2005). Neural Correlates of Simulated Driving: Auditory Oddball Responses Dependent on Workload. In Schmorrow, D. D. (Eds.), *Foundations of Augmented Cognition* (pp. 1067-1076). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Publishers.

Ranney T. A. (1994) Models of driving behaviour: a review of the revolution. *Accident Analysis and Prevention*, 26(6), 733-750.

Rasmussen, J. (1983). Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC, 13, 257-266.

Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S. & Flemisch, F. (2009). The importance of driver state assessment within highly automated vehicles. *Paper presented at the 16th ITS World Congress*, Stockholm, Sweden, 21.-25. September 2009.

Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S. & Flemisch, F. (2010). Measures and Countermeasures for impaired driver's state within highly automated driving. *Transport Research Arena Europe 2010* (TRA2010), Brüssel, Belgien, 07.-10.06.2010.

Rechtschaffen, A., & Kales, A. (1968). *A Manual of Standardized Terminology, Techniques and Scoring System For Sleep Stages of Human Subjects*. US Department of Health, Education, and Welfare; National Institutes of Health.

Reid, G.B. & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P.A.

Hancock and N. Meshkati (Eds.). *Human Mental Workload,*. North Holland, Amsterdam, 185-218.

Rohmert, W. (1984). Das Belastungs - Beanspruchungs - Konzept. *Zeitschrift für Arbeitswissenschaft*, 38, 193-200.

Rouse, W.B., Edwards, S.L. & Hammer, J.M. (1993). Modelling the dynamics of mental workload and human performance in complex systems. *IEEE transactions on systems, man, and cybernetics*, 23, 1662-1671.

Russell, C. A., Wilson G. F., & Monett, C. T. (1996). Mental workload classification using a backpropagation neural network. In C. H. Dagli, M. Akay, C. L. P. Chen, B. R. Fernandez, & J. Ghosh (Eds.), *Intelligent engineering systems through artificial neural networks* (Vol. 6, pp. 685-690). New York: American Society of Mechanical Engineers.

Russell, C. A., & Wilson, G. F. (1998). Air traffic controller functional state classification using neural networks. In *Proceedings of the Artificial Neural Networks in Engineering (ANNIE'98) Conference* (Vol. 8, pp. 649-654). New York: American Society of Mechanical Engineers.

Sammer G, Blecker C, Gebhardt H, Bischoff M, Stark R, Morgen K, Vaitl D. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. *Human Brain Mapping*, 28(8), 793-803.

Sarter, N. & Woods, D.D. (1992). Pilot interaction with cockpit automation. I. Operational experiences with the Flight Management System. *International Journal of Aviation Psychology*, 2, 303-321.

Sayer, J. R, Devonshire, J. M., and Flanagan, C. A. (2007). Naturalistic driving performance during secondary tasks. In *Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Human performance in automated systems: Theory and applications* (pp. 37-64). Mahwah, NJ: Erlbaum.

Scerbo M.W., Freeman F.G., Mikulka P.J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science*, 4, 200-219.

Scerbo, M.W. (2007). Adaptive automation. In R. Parasuraman, & M. Rizzo (Eds.) *Neuroergonomics: The Brain at Work.* (pp. 239-252). New York: Oxford University Press.

Schandry, R. (2006). *Biologische Psychologie.* Weinheim: Beltz Verlag

Schieben, A., Temme, G., Köster, F., & Flemisch, F. (2011). How to interact with a highly automated vehicle: generic interaction design schemes and test results of a usability assessment. In D. de Waard, N. Gérard, L. Onnasch, R. Wiczorek, and D. Manzey (Eds.) (2011). *Human Centred Automation* (pp. 251-266). Maastricht, the Netherlands: Shaker Publishing.

Schier, M.A. (2000). Changes in EEG alpha power during simulated driving: a demonstration. *International Journal of Psychophysiology*, 155-162.

Schneider, W., Dumais, S.T. & Shiffrin, R.M. (1984). Automatic and control processing and attention. In R. Parasuraman and D.R. Davies (Eds.), *Varieties of attention*. (pp. 1-27). London: Academic Press.

Schultheis, H,, & Jameson, J. (2004). *Assessing cognitive load in adaptive hypermedia systems*: Physiological and behavioral methods. In W. Neijdl, P. De Bra (Eds.), *Adaptive hypermedia and adaptive web-based systems* (pp.18-24). Eindhoven, Netherland: Springer.

Schumacher, M. (2001). Vergleich verschiedener subjektiver Maße der Beanspruchung bei Realfahrten im PKW. Unpublished Diplomarbeit, Bayerische Julius-Maximilians-Universität Würzburg, Germany.

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J.R., Bromfield, E.B., McCarthy, D.C., Brandt, A., et al. (2007). Hippocampal and neocortical gamma

oscillations predict memory formation in humans. *Cerebral Cortex*, 17, 1190-1196.

Senders, J. W., Kristopherson, A. B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). The Attentional Demand of Automobile Driving. *Highway Research Record*, 195, 15-32.

Sheridan, T. B. (1995). Human-centered automation: oxymoron or common sense?. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics* (pp.1 - 6).

Sheridan, T.B., Parasuraman, R. (2006). Human-Automation Interaction. *Reviews of Human Factors and Ergonomics*, 1, 89-129.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience.* 18, 555-586

Sirevaag, E., Kramer, A.F., De Jong, R. & Mecklinger, A. (1988). A psychophysiological analysis of multi-task processing demands. *Psychophysiology*, 25, 482.

Smith, M.E., Gevins, A., Brown, H., Karnik, A., and Du, R. (2001). Monitoring task load with multivariate EEG measures during complex forms of human-computer interaction. *Human Factor*, 43(3), 366-380.

Son, J., Lee Y., & Kim M.H. (2011) Impact of traffic environment and cognitive workload on older drivers' behavior in simulated driving. *International Journal of Precision Engineering and Manufacturing*, 12(1) 135-141.

Spiers H.J., & Maguire E.A. (2007). Neural substrates of driving behaviour. *Neuroimage*. 36(1), 245-55.

Sterman, M.B. (1981). EEG-feedback: Physiological behavior modification. *Neuroscience and Biobehavioral Review*, 5, 405-412.

Sterman, M.B., & Mann, C.A. (1995) Concepts and applications of EEG analysis in aviation performance evaluation, *Biological Psychology*. 40(1-2), 115-30.

Sterman, M. B., Mann, C. A., Kaiser, D. A., & Suyenobu, B. Y. (1994). Multiband topographic analysis of a simulated visuomotor aviation task. *International Journal of Psychophysiology*, 16, 49-56.

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2003). Fatal distraction? A comparison of the cell-phone driver and the drunk driver. In *Proceedings of the second international driving symposium on human factors in driver assessment, training and vehicle design* (pp. 25-30). The University of Iowa Public Policy Center.

Sutton S, Braren M, Zubin J, John ER. (1965). Evoked-potential correlates of stimulus uncertainty. *Science.* 150(700), 1187-1188.

Tallon-Baudry, C., Bertrand, O., Hénaff, M.A., Isnard, J., & Fischer, C. (2005). Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cerebral Cortex*, 15, 654-662.

Teplan M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2 (2), 1-11.

Thatcher R.W., North D., Biver C. (2005). Evaluation and validity of a LORETA normative EEG database. *Clinical and EEG Neuroscience*. 36(2), 116-22.

Ting C.H.; Mahfouf, M.; Nassef, A.; Linkens, D.A.; Panoutsos, G.; Nickel, P.; Roberts, A.C.; Hockey, G. (2010). Real-Time Adaptive Automation System Based on Identification of Operator Functional State in Simulated Process Control Operations. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40 (2), 251 - 262

Törnros, J., and Bolling, A. (2006) Mobile Phone Use--Effects of Conversation on Mental Workload and Driving Speed in Rural and Urban Environments. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9 (4), 298-306.

Tsang, P. S. and Vidulich, M. A. (2006). Mental Workload and Situation Awareness. In *Handbook of Human Factors and Ergonomics* (Chapter 9, 243-268). Wiley, Hoboken, New Jersery.

Uchiyama Y, Ebe K, Kozato A, Okada T, Sadato N. (2003). The neural substrates of driving at a safe distance: a functional MRI study. *Neuroscience Letter*, 352(3), 199-202.

Van Boxtel, A., & Jessurun, M. (1993). Amplitude and bilateral coherency of facial and jawelevator EMG activity as an index of effort during a two-choice serial reaction task. *Psychophysiology*, 30, 589-604.

Van Winsum, W., Martens, M. H., & Herland, L. (1999). *The effects of speech versus tactile driver support messages on workload, driver behaviour and user acceptance* (Report No. TM-99-C043). Soesterberg, Netherlands: TNO Human Factors Research Institute.

Veltman, J.A. & Gaillard, A.W.K. (1993). Indices of mental workload in a complex task environment. *Neuropsychobiology*, 28, 72-75.

Veltman, J.A. & Gaillard, A.W.K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, 42, 323-342.

Veltman, J.A., Jansen, C. (2006). The role of operator state assessment in adaptive automation. TNO Human Factors Research Institute.

Verwey, W.B. (1990). *Adaptable driver-car interfacing and mental workload: a review of the literature* (Report V1041/GIDS/DIA/1). Haren, The Netherlands: Traffic Research Centre, University of Groningen.

Verwey W.B. (2000). On-line driver workload estimation. Effect of road situation and age on secondary task measures. *Ergonomics*, 43(2), 187-209.

Waterink, W., & Van Boxtel, A. (1994). Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology*, 37, 183-198.

Walter, H., Vetter, S.C., Grothe, J., Wunderlich, A.P., Hahn, S., & Spitzer, M. (2001). The neural correlates of driving. *Neuroreport*, 12, 1763-1767.

Walter S., Geffen G.M., Geffen L.B. (2001). The n-back as a dual-task: P300 morphology under divided attention. Psychophysiology, 38, 998-1003.

Walter, W.G., Cooper, R., Aldridge, V.J., McCallum, W.C., Winter, A.L. (1964). Contingent Negative Variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature*, 203, 380-384.

Ward, L.M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences*, 7, 553-559.

Weller, G., Schlag, B., Gatti, G., Jorna, R. & Leur, M.v.d. (2006). *Human Factors in Road Design State of the art and empirical evidence*. Report 8.1 RiPCORD-iSEREST.

Welke, S., Juergensohn, T., Roetting, M. (2009). Single-Trial Detection Of Cognitive Processes For Increasing Traffic Safety. In *Proceedings of the 21st International Technical Conference On the Enhanced Safety of Vehicles*. Stuttgart, Germany.

Wickens, C. D., Isreal, J., Donchin, E. (1977). The event-related cortical potential as index of task workload. In *proceedings of Human Factors and Ergonomics Society Annual Meeting*, 21 (4), 282-286.

Wickens C. D., Kramer A, Vanasse L, Donchin E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science*, 221(4615), 1080-1082.

Wickens, C.D. (1984). Processing resources in attention. In R.Parasuraman and D.R. Davies (Eds.). *Varieties of attention*. (pp.63-102). London: Academic Press.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance*. New York, Harper Collins Publishers Inc.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3(2),* 159-177.

Wickens, C. D., Lee, J., Liu, Y., & Gorden Becker, S. (2004). *An Introducation to Human Factors Enginering*. Second edition. New Jersey: Pearson.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors,* 50 (3), 449-455.

Wiener, E.L. (1988). Cockpit Automation. In E.L. Wiener & D.C. Nagel (Eds.), *Human factors in aviation* (pp. 433-461). San Diego, CA: Academic Press.

Wierwille, W.W. (1979). Physiological measures of aircrew mental workload. *Human Factors*, 21, 575-593.

Wierwille, W.W. and Casali, J.G. (1983). A validated rating scale for global mental workload measurement applications. In *Proceedings of the Human Factors Society - 27th Annual Meeting,* 129-133.

Wierwille, W.W. & Rahimi, M. & Casali, J.G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human Factors*, 27, 489-502.

Wierwille, W.W. & Eggemeier, F.T. (1993). Recommendation for mental workload measurement in a test and evaluation environment. *Human Factors*, 35, 263-281.

Wilson, G.F. (1992). Applied use of cardiac and respiration measures: practical considerations and precautions. *Biological Psychology*, 34, 163-178.

Wilson, G. F., & Fisher, F. (1995). Cognitive task classification based upon topographic EEG data. *Biological Psychology*, 40, 239-250.

Wilson, G.F., & O'Donnel, R.D. (1988). Measurement of operator workload with the neuropsychological workload test battery. In: Hancock, P., Meshkati, N. (Eds.), *Human Mental Workload*. North-Holland, Amsterdam, 63-100.

Wilson, G. F., & Russell, C. (1999). Operator Functional State Classification Using Neural Networks with Combined Physiological and Performance

Features. *In: Human Factors and Ergonomics Society Annual Meeting Proceedings*, 1099-1102.

Wilson, G. F. (2001). In-flight psychophysiological monitoring. In F. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory monitoring* (pp. 435-454). Seattle, WA: Hogrefe and Huber.

Wilson, G. F. (2002a). Psychophysiological Test Methods and Procedures. In S. G. Charlton and T. G. O'Brien (eds.). *Handbook of Human Factors Testing and Evaluation.* Lawrence Erlbaum Associates, Mahwah, N.J., 127-156.

Wilson, G. F. (2002b). An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. *The International Journal of Aviation Psychology*, 12(1), 3 -18.

Wilson, G. F., & Russell, C. A. (2003a). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human Factors*, 45, 381-389.

Wilson, G.F., & Russell, C.A. (2003b). Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. *Human Factors*, 45(4), 635-643.

Wilson, G. F., & Schlegel, R. E. (2004). Operator functional state assessment. Paris, FR, North Atlantic Treaty Organisation (NATO), *Research and Technology Organisation* (RTO), BP 25, F-92201, Neuilly-sur-Seine Cedex, France: 220.

Wilson, G.F. and Russell, C.A. (2004). Psychophysiologically determined adaptive aiding in a simulated UCAV task. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness, and automation: Current research and trends* (pp. 200-204). Mahwah, NJ: Erlbaum.

Wilson, G.F. and Russell, C.A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors*, 49 (6). 1005-1018.

Woods, D.D. (1994). Automation: Apparent simplicity, real complexity. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp.1-7). Hillsdale, NJ: Lawrence Erlbaum Associates.

Xie, B., & Salvendy, G. (2000a). Prediction of mental workload in single and multiple tasks environments. *International Journal of Cognitive Ergonomics*, *4*(*3*), 213-242.

Xie, B., & Salvendy, G. (2000b). Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. *Work & Stress*, 14(1), 74-99.

Yamamoto, S., & Matsuoka, S. (1990). Topographic EEG study of visual display terminal VDT performance with special reference to frontal midline theta waves. *Brain topography*, 2, 257-267.

Yang J., Yao S., Li J., Xu Y., & Ma Z. (1997). Application of mean EMG to the evaluation of human workload during manual operation. *Space Medicine and Medical Engineering*, 10(6), 425-429.

Yeh, Y.Y., Wickens, C.D. (1988). *Dissociation of Performance and Subjective Measures of Workload*. 30 (1), 111-120.

Yerkes R.M. & Dodson J.D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, 18 (5), 459-482.

Zijlstra, F.R.H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort.* Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.

List of Abbreviations

AA: adaptive automation

ACC: adaptive cruise control

ANN: artificial neural network

ANOVA: analysis of variance

ANS: autonomic nervous system

ATA: Adaptive task allocation

BCI: brain-computer interface

CDF: cumulative distribution function

CNV: contingent negative variation

CNS: central nervous system

DALI: driver activity load index

DLPFC: dorsolateral prefrontal cortex

DVE: driver-vehicle-environment

DWM-ATA: Damand-workload-matched adaptive task allocation

ECG: Electrocardiogram

EDA: electrodermal activity

EEG: Electroencephalogram

EMG: electromyogram

ERPs: event-related potentials

EOG: electrooculargram

HIP: Human Information processing

HR: heart rate

HRV: heart rate variability

IBI: inter-beat interval

IC: independent component

ICA: independent component analysis

IVTs: in-vehicle technologies

LCT: lane change task

LDA: linear discriminate analysis

LFM: logistic function model

LOA: level of automation

MATB: Multiple-Attribute Task Battery

MCH: Modified Cooper-Harper

NASA-TLX: National Aeronautics and Space Administration Task Load Index

NN50: the number of pairs of successive NNs that differ by more than 50 ms

OFS: operator's functional states

PASAT: Paced Auditory Serial Addition Task

PDF: probability distribution function

PDT: peripheral detection task

PNS: peripheral nervous system

RID: rate of information processing demand

RMSSD: the square root of the mean squared differences of successive NN intervals

RP: Bereitsschaftspotential or readiness potential

RSME: Rating Scale Mental Effort

RT: reaction time

SCP: slow cortical potentials

SDNN: the standard deviation of NN interval

SDANN: standard deviation of the average NN intervals calculated over short period

SNS: sympathetic nervous system

SVM: support vector machine

SWAT: subjective workload assessment technique

TWA: amplitude of T-wave

List of Tables

Table 4. 1 *Correlation (Pearson r) of the EEG parameters to other variables*.....*142*

Table 5. 1 *The averaged p values (standard deviation across subjects) for each task condition*.....*154*
Table 5. 2 *Comparison of the coefficients of variation between the z-scores and p-values*.....*156*
Table 5. 3 *The subjective load and corresponding p-values of theta and alpha power for subject 11*.....*160*
Table 5. 4 *Mean values (standard deviations) of the logistic function workload scores and the percentages of occurrences of the low, moderate and high workload in 9 task load conditions averaged over participants (N=24)*.....*163*
Table 5. 5 *Comparison of the n-back performance, theta power, alpha power, and workload scores between subject 7 and 11*.....*165*
Table 5. 6 *Comparison of correlation coefficients of the EEG parameters and regressed scores to other variables (SL, HR, HRV, RT, n-back error rate). The regressed scores showed better correlations with other workload indexes comparing with the fro-theta and par-alpha.*.....*166*

Table 6. 1 *The mean percentages (standard deviation) of low, moderate and high workload averaged over participant (N=20).*.....*193*

List of Figures

Figure 1. 1 *The problems of driving induced by increased complexity and simplicity, for which driver adaptive task allocation is proposed as the solution 4*

Figure 2. 1 *A closed-loop for adaptive task allocation 11*
Figure 2. 2 *Human Information processing model (Wickens et al., 2004, p. 122)..... 17*
Figure 2. 3 *Wickens' Multiple Resources Theory (adapted from Wickens, 2002)..... 18*
Figure 2. 4 *Inter-relationships of task load, workload, and performance (modified from Parasuraman & Hancock, 2001)..... 19*
Figure 2. 5 *Yerkes-Dodson law about the relation between workload and performance. 20*
Figure 2. 6 *De Waard's model concerning relations between task demand, workload and performance (adapted from De Waard, 1996) 21*
Figure 2. 7 *Wilson & Schlegel's model for task demands, workload, and performance (adapted from Greef et al., 2007)..... 22*
Figure 2. 8 *Workload-matched adaptive task allocation (modified from Parasuraman et al., 1999)..... 24*
Figure 2. 9 *A 3*3 metrix for the task demand and workload 26*
Figure 2. 10 *Demand-workload-matched adaptive task allocation (DWM-ATA) for human operator 27*
Figure 2. 11 *Workload assessment methods 29*
Figure 2. 12 *The typical time-domain ECG tracing of the cardiac cycle..... 43*
Figure 2. 13 *10-20 system for the standardized electrode placement 50*
Figure 2. 14 *Event-related potentials after stimulus 51*
Figure 2. 15 *EEG frequency components delta, theta, alpha, beta, and gamma 53*
Figure 2. 16 *Linking Michon's hierarchical model with Rasmussen's (1983) knowledge-rule-skill model (adapted from Weller et al., 2006)..... 67*

Figure 3. 1 *Lane Change Task: lane changes are promoted by the road signs..... 85*
Figure 3. 2 *Deviation between a normative model and the driver's behavioural data for performance assessment in LCT 86*
Figure 3. 3 *Experiment setup for the pre-study..... 87*
Figure 3. 4 *Correlation of subjective workload (left) and LCT performance (right) to speed levels in the pre-study 88*
Figure 3. 5 *LCT equipped with a driving box 90*
Figure 3. 6 *Comparison of mean deviation in both single task (left) and dual task (right) conditions..... 96*
Figure 3. 7 *Comparison of the PASAT performance in single and dual task conditions.. 96*
Figure 3. 8 *Independent components which demonstrate the typical components obtained in the present study (left: topography of the IC and right: the corresponding averaged*

signal). *97*
Figure 3. 9 *Artefacts-corrected ERP and steering angle in LCT* *99*
Figure 3. 10 *Spatial pattern of the ERPs in single task condition: (a) slow driving (60km/h), (b) moderate driving (80km/h), and (c) fast driving (100km/h). These results were based on the average of 24 participants.* *100*
Figure 3. 11 *Artefacts-corrected ERPs at different electrodes in single task conditions (N=24)* *101*
Figure 3. 12 *Spatial pattern of the ERPs in dual task conditions: (a) driving only at 80 km/h, (b) driving at 80 km/h plus p5, and (c) driving at 80 km/h plus p3. These results were based on the average of 22 participants.* *102*
Figure 3. 13 *Artefacts-corrected ERPs at different electrodes in dual task conditions (N=24)* *103*
Figure 3. 14 *Spatial patterns of theta, alpha, and beta in the single task condition* *104*
Figure 3. 15 *The changes in theta, alpha and beta powers with speed (N=24)* *105*
Figure 3. 16 *Spatial pattern of theta, alpha and beta in dual task conditions (N=24)* ... *107*
Figure 3. 17 *Comparison of changes in theta, alpha, and beta in dual conditions* *108*
Figure 3. 18 *Mean classification accuracy of ERPs and BPs for the single and dual task conditions* *110*

Figure 4. 1 *The overlaid projection of the LCT driving scene and n-back task. The digits were presented every 3 seconds with 1 second presence.* *123*
Figure 4. 2 *The procedures for the EEG data analysis. The means and standard deviations of the theta and alpha power were extracted. The means were calculated by averaging over data points within each task condition and participant. The standard deviations were calculated over data points within each task condition and participant.* *127*
Figure 4. 3 *The subjective loads in different task conditions averaged over participants (N=24). The error bar shows the standard error.* *129*
Figure 4. 4 *The mean deviation for LCT in different driving conditions averaged over participants (N=24). The error bar shows the standard error.* *130*
Figure 4. 5 *The n-back reaction time in different task conditions averaged over participants (N=24). The error bar shows the standard error.* *131*
Figure 4. 6 *The n-back error rate in different task conditions averaged over participants (N=24). The error bar shows the standard error.* *131*
Figure 4. 7 *The heart rate in different task conditions averaged over participants (N=24). The error bar shows the standard error.* *132*
Figure 4. 8 *The heart rate variability (LF/HF) in different task conditions averaged over participants (N=24). The error bar shows the standard error.* *133*
Figure 4. 9 *The topographies of the theta power (z-scores) in different task conditions averaged over participants (N=24).* *134*
Figure 4. 10 *The topographies of the alpha power (z-scores) in different task conditions averaged over participants (N=24).* *134*
Figure 4. 11 *Relative power spectrum density (the percentages to the total power of 4-30 Hz) (a) at frontal recording site (Fz); and (b) parietal recording site (Pz) averaged over*

participants (N=24). *135*
Figure 4. 12 *Statistical properties of short-term theta powers in low and high task load conditions for subject 10. (a) Histogram of the theta powers in low task load condition (b) histogram of the theta power in high task load condition. (c) Comparison of the regressed probability distributions of theta powers in the low and high task load conditions (d) Boxplot of the theta powers in the low and high task load conditions.* *137*
Figure 4. 13 *Statistical properties of short-term alpha powers in low and high task load conditions for subject 10. (a) histogram of the alpha powers in low task load condition (b) histogram of alpha powers in high task load condition. (c) comparison of the regressed probability distributions of alpha powers in the low and high task load conditions (d) boxplot of the alpha powers in the low and high task load conditions.* *137*
Figure 4. 14 *Mean z-scores of frontal theta (4-8Hz) for each of 9 task conditions. Z-scores were converted for each participant and averaged over 24 participants (N=24). The error bars show standard errors of the z-scores.* *138*
Figure 4. 15 *Mean z-scores of parietal alpha (8-12Hz) for each of 9 task conditions. Z-scores were converted for each participant and averaged over 24 participants (N=24). The error bars show standard errors of the z-scores.* *139*
Figure 4. 16 *Mean standard deviations of the short-term theta powers in different task conditions averaged over participants (N=24). The error bars show standard errors of this parameter.* *140*
Figure 4. 17 *Mean standard deviations of the short-term alpha powers in different task conditions averaged over participants (N=24). The error bars show standard errors of this parameter.* *140*

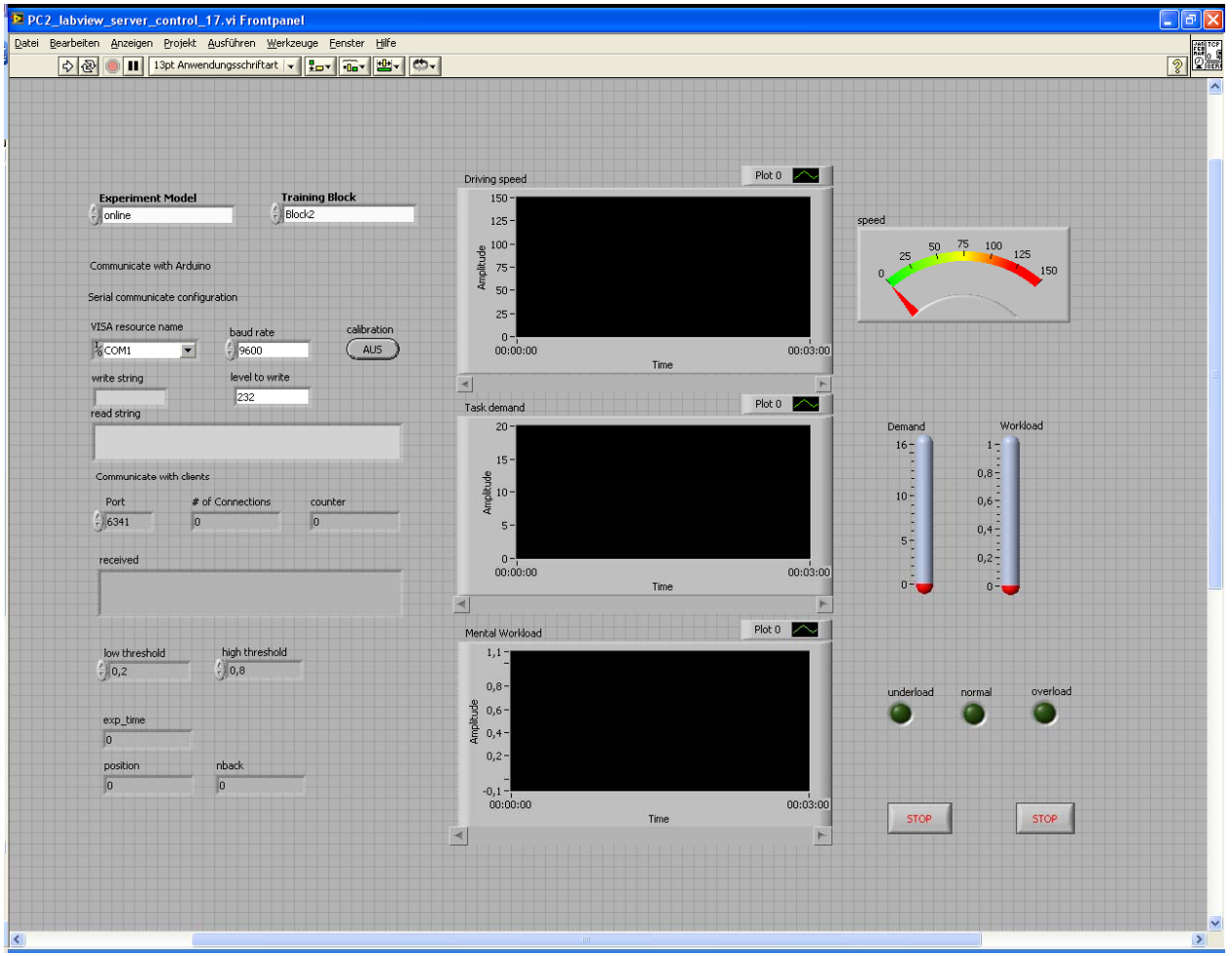
Figure 5. 1 *Distribution functions of theta powers for subject 4*..... *152*
Figure 5. 2 *Distribution functions of theta powers of subject 9* *152*
Figure 5. 3 *Distribution functions of alpha powers for subject 4* *153*
Figure 5. 4 *Distribution functions of alpha powers of subject 22* *153*
Figure 5. 5 *Comparison of the individual variation of the mean values, mean z-scores, and mean p values with each containing 24 data points (24 participants) for theta and alpha.* *155*
Figure 5. 6 *Logistic function curve* *158*
Figure 5. 7 *The comparison of the logistic function transferred workload indices with theta and alpha power in low (‘base’ & N0), moderate (’slow’ & N1) and high (‘fast’ & N2) for subject 11* *162*

Figure 6. 1 *EEG-based neural adaptive interface for driver adaptive task allocation*.. *172*
Figure 6. 2 *System overview in the third experiment. The whole system consisted of four computer and two projectors. A microcontroller was used as well to control the speed settings.* *175*
Figure 6. 3 *Labview interface for EEG-based closed-loop system in the third experiment. Speed, task demands, mental workload, and the mental workload states can be dynamically*

visualized driving the experiment. *177*
Figure 6. 4 *Task demand control strategy according to the DWM-ATA model* *180*
Figure 6. 5 *General comparison of task demand between the non-control and control task mode (N=20). The error bar shows the standard error.* *183*
Figure 6. 6 *Task demand comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.* *183*
Figure 6. 7 *General comparison of subjective loads comparison between the non-control and control mode (N=20). The error bar shows the standard error.* *184*
Figure 6. 8 *Subjective load comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.* *185*
Figure 6. 9 *General comparison of n-back error rate between the non-control and control task mode (N=20). The error bar shows the standard error.* *186*
Figure 6. 10 *General comparison of n-back reaction time between the non-control and control task mode (N=20). The error bar shows the standard error.* *186*
Figure 6. 11 *n-back error rate comparison between the non-control and control task mode (N=20). The error bar shows the standard error.* *187*
Figure 6. 12 *n-back reaction time comparison between the non-control and control task mode (N=20). The error bar shows the standard error.* *187*
Figure 6. 13 *General comparison of HR between non-control and control model (N=20). The error bar shows the standard error* *188*
Figure 6. 14 *General comparison of HRV(LF/HF) between non-control and control model (N=20). The error bar shows the standard error.* *188*
Figure 6. 15 *Heart Rate comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.* *189*
Figure 6. 16 *HRV (LF/HF) comparison between the non-control and control task mode in different n-back conditions (N=20). The error bar shows the standard error.* *190*
Figure 6. 17 *General comparison of EEG W.L. score between non-control and control model (N=20). The error bar shows the standard error.* *191*
Figure 6. 18 *General comparison of percentages of low, moderate and high workload between non-control and control model (N=20). The error bar shows the standard error.* *191*
Figure 6. 19 *EEG workload scores comparison between the non-control and control task mode (N=20). The error bar shows the standard error.* *192*
Figure 6. 20 *The n-back error rate (standard error across participant) according to DWM-ATA model (There was no n-back task in low task demand region. Thus, there was no performance data in this region)* *193*
Figure 6. 21 *DWM-ATA model for the driver-vehicle-environment loop* *199*

Appendix 1. The labview interface developed for experiment 3

This interface serves as the general controller of the experiment process as well as visualizes the data stream for the adaptive task allocation. The experiment model and training block could be configured in this interface. Meanwhile, the configuration of the TCP/IP communication, the communication between the Labviewservice and microcontroller, and workload thresholds could be set in this interface. The description of the visualization of the data stream is given in section 6.2.3. The original Labview code is included in the CD enclosed with this thesis.



Appendix 2. The Matlab toolbox for workload detection

A Matlab toolbox, workload-detection-toolbox 1.3, has been developed both offline and online EEG data processing. The original matlab codes were included in enclosed CD. Here, the main functions and short descriptions are given as following.

Offline processing:

do_sle_online_wrap(): This function initializes the Phypha toolbox including set the directory of the toolbox.

do_sle_onlin(): This function is used to train the logistic function model using the training datasets. The sub-functions including the dataset loading function *io_quickload ()*, data preprocessing function *bci_preproc()* are included in the Phypha toolbox.

sle_data_selection(): this function is used to exclude the data during the curve in the LCT.

sle_adaptive_theta_alpha (): since the center frequency of the theta activity and alpha activity is different from individuals. This function is used to find the center frequency for the theta and alpha activity.

sle_bp_decomposation(): this function is used to decompose the EEG signal into frequency bands.

sle_label_generation(): this function is used to generate the labels for the features extracted from *sle_bp_decomposation().* The labels indicate which condition the features belong to.

para_workload_train(): this function is used to train the logistic function model using the features. It includes several subfunctions: *sle_feature_choose (); sle_quantile_estimation (); sle_logit_regress ().*

sle_feature_choose (): In order to voiding that the features are dominated by a certain task level, the size of features should be equal for each task condition. This function is used to randomly select equal-sized the feature from each task condition.

sle_quantile_estimation (): This function is used to extract the p-quantiles using the predefined the p-values.

sle_logit_regress (): This function is used to train the logistic function model using the predefined workload scores and the p-quantiles. The coefficients could be obtained through the regression of this model.

sle_erp_artefacts_finding (): This function is used to automatically find the ocular as well as muscle artifacts induced ICs.

Online processing:

para_workload_test (): This function is used to download the EEG data set and the workload detection model during the online process. It includes several sub-functions: *utl_epochfilter()* included in the Phypha toolbox for EEG data preprocessing; *ext_sle_lafilter(); ext_sle_bp_decomposation (); ext_sle_workload_predict ().*

ext_sle_lapfilter (): This function is used to do the laplacian filtering of the preprocessed EEG data.

ext_sle_bp_decomposation (): This function is used to decompose the EEG frequency bands of the preprocessed and laplacian filtered EEG data.

ext_sle_workload_predict (): This function is used to quantify the workload using EEG frequency bands and the logistic function model.