Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)

Présentée et soutenue par :
Gautier DURANTIN

le 21/10/2015

Titre :

Brain imaging techniques to monitor the attentional state under ecological aeronautical settings

Contrôle de l'état attentionnel en conditions aéronautiques écologiques par imagerie cérébrale

---

École doctorale et discipline ou spécialité :

ED AA : Neurosciences – Signal, Image, Acoustique et Optimisation

Unité de recherche :

Équipe d'accueil ISAE-ONERA CSDV

Directeur(s) de Thèse :

M. Frédéric DEHAIS (directeur de thèse)
M. Arnaud DELORME (co-directeur de thèse)

Jury :

M. Simon THORPE - Président du jury
M. Hasan AYAZ - Rapporteur
M. Frédéric DEHAIS - Directeur de thèse
M. Arnaud DELORME - Co-directeur de thèse
M. Fabien LOTTE - Examinateur
M. Stéphane PERREY - Rapporteur

Contrôle de l'état attentionnel en conditions aéronautiques écologiques par imagerie cérébrale

L'état attentionnel de l'opérateur est un des précurseurs de l'erreur humaine dans les systèmes complexes. Cela est particulièrement vrai en aéronautique, où la sécurité dépend en premier lieu de la capacité à réagir rapidement et correctement. Les niveaux de complexité associés à la gestion de tels systèmes aboutissent à des niveaux de charge mentale et d'engagement de l'opérateur en constante variation, qui peuvent être prédicteurs de sa performance. Ce projet de recherche adopte une démarche de Neuroergonomie, et vise à estimer l'état attentionnel en conditions écologiques par l'utilisation de mesures cérébrales. Nous avons tout d'abord étudié le comportement de l'opérateur soumis à des niveaux de demande extrêmes à l'aide de mesures cérébrales et psycho-physiologiques. Les résultats de ces études nous ont conduits au développement d'un nouveau cadre théorique centré sur l'engagement de l'opérateur pour estimer son état attentionnel. De plus, nous avons étudié différentes techniques de traitement du signal de manière à rendre possible l'utilisation des mesures cérébrales en temps réel en situation écologique, en vue du développement d'interfaces cerveau-machine pour assister l'opérateur.

Brain imaging techniques to monitor the attentional state under ecological aeronautical conditions

The attentional state of operators is one of the main reasons for errors during human control of complex systems, and controlling these errors is critical especially in aeronautics, where errors are directly linked to safety and lives might be at stake. In particular, excessively high or low task demands encountered during the operation of such systems result in varying levels of mental workload and engagement which are linked with the operator performance. This research project adopts a Neuroergonomics approach and investigates the use of brain measurement techniques to monitor the attentional state of the operator under ecological conditions. We studied the behavior of the operator under both excessively low and high task demands with the use of multiple physiological and neurophysiological measurement techniques. Our results show that it is possible to use such techniques to characterize the attentional state. We then analyze the potential of real time application for such techniques. We investigated signal processing and analysis tools to improve the real-time usability of brain signals in ecological conditions, and proposed solutions towards the development of brain computer interfaces for assisting the human operator.

Acknowledgements

This PhD project funded by the Pôle de Recherche et d'Enseignement Supérieur (PRES) and the region Midi-Pyrénées would not have been possible without the contribution (direct or indirect) of many actors. This dissertation is dedicated to all of them, and I would like to thank them for their support through these few lines.

First, I would especially like to thank my two directors, Frédéric Dehais and Arnaud Delorme. They gave me the chance to start a PhD under their supervision, in 2012, even though I had no experience in the field. I am especially grateful for their trust, but also for their support and the precious advice during the project.

During three years I had the chance to meet wonderful colleagues. First, I would like to thank all the members of the DCAS department at ISAE-Supaéro, in particular Christian, Éric, Dominique, Patrice, Guillaume, Étienne, Jérême, Jérémie, Stéphane and Marie. then, all the member of the CerCo, and especially the CREMe team. Although I could not always spend as much time as I wanted with them, they contributed a lot to the great working atmosphere.

All the studies presented in this dissertation would not have been possible without all my colleagues from the Human Factors department at ISAE : Zarrin, Louise, Eve, Julia, Angela, Mickaël, Sébastien, Thibault, Kevin, Nicolas, Seva, Benoêt, and Francois. Special thanks go to Sébastien for not getting mad at me each time I came to his office with silly questions, to Thibault for sharing his office with me and for enduring the stress-related foot tapping, and to Zarrin for the precious support in english !

This PhD project was also the opportunity for me to meet great scholars. A special tribute goes to the late professor Raja Parasuraman, pioneer of Neuroergonomics. In a very humble way, this project is following his footsteps. Very special thanks also go to Daniel Callan at CiNet, Osaka University (Japan), who hosted me in his laboratory from October 2014 to January 2015. This fruitful internship made me discover a wonderful country and friendly people, and I do hope it was the beginning of a long collaboration. Finally, I would especially like to thank Dr. Hasan Ayaz and Dr. Stéphane Perrey, who accepted to review this dissertation, as well ad Dr. Fabien Lotte and Dr. Simon Thorpe, who accepted to be part of my thesis committee.

To conclude, I would like to thank my friends and family, who supported me all the way from the beginning of the project and long before. Those years would probably not have passed this well without them. Special thanks go to Julien, Edith, Fabrice, Aurélie, Éric, Thomas, Aline, Adeline, Alexandre, Sylvia, Alain, Léonore, Gaëtan, Arsène, Basile, Éléa and Auguste.

Contents

Contents 7

I Introduction 11

II State of the art 19

1 Performance and Task Demands 21
- 1.1 The concept of mental workload . . . 22
  - 1.1.1 Task demands and resources . . . 22
  - 1.1.2 Nature of mental resources . . . 23
  - 1.1.3 Cognitive control and resource allocation . . . 25
  - 1.1.4 Factors influencing resource allocation . . . 27
- 1.2 Mental Workload, Engagement and Cognitive Performance . . . 28
  - 1.2.1 Human behavior and engagement in highly demanding situations . . . 31
  - 1.2.2 Human behavior in low demanding situations . . . 34
- 1.3 Conclusions on mental workload and engagement in the evaluation of performance . . . 35

2 Neuroergonomics 37
- 2.1 History of Neuroergonomics . . . 37
- 2.2 Approaches in Neuroergonomics . . . 39
- 2.3 Measures in Neuroergonomics . . . 40
  - 2.3.1 Subjective measures . . . 41
  - 2.3.2 Objective measures : behavioral approach . . . 42
  - 2.3.3 Objective measures : Psycho-physiological approach . . . 43
  - 2.3.4 Objective measures : brain measurement techniques . . . 44
  - 2.3.5 Summary of the different measures . . . 50
- 2.4 Conclusions on Neuroergonomics . . . 50

3 Brain Computer Interfaces 53
- 3.1 Use and objectives of Brain Computer Interfaces . . . 53
  - 3.1.1 Active BCI . . . 55

3.1.2 Passive BCI . . . 56
3.2 BCI design . . . 57
3.2.1 Signal processing . . . 57
3.2.2 Control law design . . . 62
3.3 Remaining challenges in BCI design . . . 64
3.4 Conclusions on BCI . . . 65

III Contribution **67**

4 Mental overload detection **71**
4.1 Objectives of the experiment . . . 71
4.2 Simulated piloting task . . . 72
4.2.1 Protocol design . . . 72
4.2.2 Task description . . . 73
4.3 Procedure and Data Collection . . . 75
4.4 Data analyses . . . 76
4.5 Results . . . 77
4.6 Discussion . . . 81
4.7 Conclusion . . . 84

5 Mind wandering **85**
5.1 First experiment : assessing mind wandering with NIRS . . . 86
5.1.1 Objectives of the experiment . . . 86
5.1.2 Methods . . . 86
5.1.3 Results . . . 90
5.1.4 Discussion . . . 92
5.2 Second experiment : mind wandering with an ecological task . . . 95
5.2.1 Objectives of the experiment . . . 95
5.2.2 Methods . . . 95
5.2.3 Data processing . . . 97
5.2.4 Results . . . 98
5.2.5 Discussion of the current results and perspectives . . . 101

6 Signal Processing and Brain Computer Interface Methods **105**
6.1 Moving Average Convergence Divergence (MACD) filter . . . 106
6.1.1 Introduction . . . 106
6.1.2 MACD filter principle . . . 106
6.1.3 Experimental Methods . . . 107
6.1.4 Cerebral hemodynamics . . . 109
6.1.5 Results . . . 111
6.1.6 Cerebral hemodynamics . . . 112

6.1.7 Discussion and conclusion . . . 114
6.2 Application of the MACD filter to a BCI in a flight simulator . . . 115
6.2.1 Objectives of the study . . . 116
6.2.2 Methods . . . 116
6.2.3 Results . . . 124
6.2.4 Discussion . . . 126
6.3 Kalman filtering for signal improvement . . . 127
6.3.1 Kalman filter for EEG . . . 127
6.3.2 Kalman filter for fNIRS . . . 137

IV Conclusion **145**

7 General discussion **147**
7.1 Attentional dynamics . . . 147
7.1.1 The dynamics of engagement and disengagement . . . 147
7.1.2 A triad of concepts to understand human behavior . . . 149
7.2 Neuroergonomics : use of brain imaging signals in realistic conditions 150
7.3 Perspectives . . . 152
7.3.1 Extending the conceptual framework of behavior : inattentional deafness/blindness . . . 152
7.3.2 Modeling human behavior . . . 154
7.3.3 Towards a BCI to assist the operator . . . 155

Appendices **161**

A Kalman Filter **163**

B Kalman filter for EEG - Classification results **167**

C Kalman Filter design for fNIRS **169**
7.1 Chapitre 1 : Demandes de la tâche et performance . . . 175
7.2 Chapitre 2 : Neuroergonomie . . . 176
7.3 Chapitre 3 : Interfaces Cerveau-Machine . . . 177
7.4 Chapitre 4 : Surcharge mentale . . . 180
7.5 Chapitre 5 : Mind wandering . . . 181
7.6 Chapitre 6 : Traitement du signal et ICM . . . 183
7.7 Discussion . . . 185
7.8 Perspectives . . . 186

Bibliography **189**

I

Introduction

Introduction

Since the birth of aeronautics with the first controlled flight of the Wright brothers in 1903, airplanes have become bigger, faster, and safer. Thanks to decades of technological and operational improvements, airplanes are now considered the **safest mean of transportation**. Indeed, with an accident rate of 2.8 per million departures in civil aviation in 2014 (ICAO 2014) , the odds of being involved in an airplane crash resulting in at least one fatality were 1 in 3.4 million (PlaneCrashInfo.com 2015). Nevertheless, a poll released in April 2015 revealed that more than half of US citizens are anxious or afraid when they board on a plane (YouGov.com 2015), although 85% of them trust the pilot to avoid crashing[1]. However, intriguingly, if airplanes became safer, human error causal role in accidents appears to be more critical. Indeed, when looking at the causes of fatal accidents in civil aviation, human error is implied in around 60% of cases (see Fig.1). When considering all kinds of accidents, for civil and military aviation, this rate increases to 70-80% (Shappell and Wiegmann 2012).

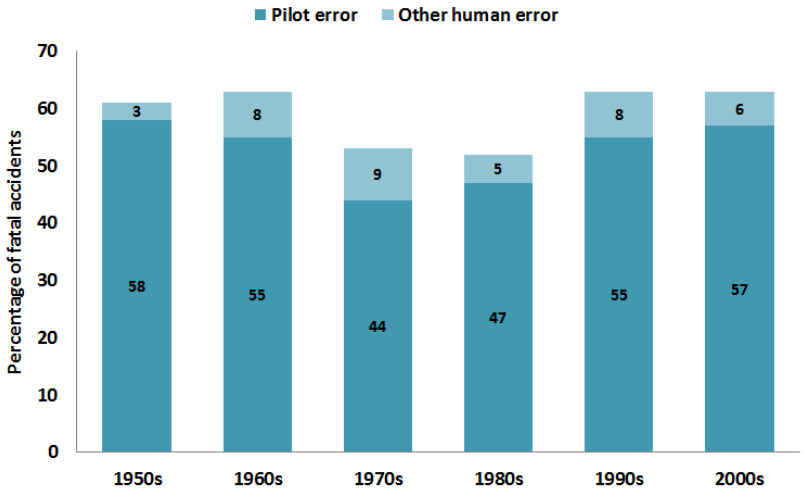


Figure 1 – *Percentage of fatal plane crashes attributed to human error per decade in civil aviation, from 1950 to 2010. Data extracted from (PlaneCrashInfo.com 2015).*

On February 4, 2015, when a TransAvia *ATR 72-600* crashed in Taipei, Taiwan, the investigation concluded that the pilot mistakenly throttled down the wrong engine fol-

[1]The fact that the poll was conducted immediately after the Germanwings crash on March 24, 2015 is a potential confound for this result. In this case, the pilot deliberately crashed the plane into the French Alps.

lowing a glitch with one of the two engines. In this case, the contribution of human error is evident, provided that the crew committed a **piloting error**. Contrarily, on August 10, 2014, when a Sepahan Airlines *HESA Iran-140-100* crashed in Tehran, Iran after it lost an engine during take off, evidence converged to a purely mechanical cause. However, investigation showed that the aircraft was carrying too much cargo and was largely overloaded for the prevailing temperature ($36°C$). The human factor is then also involved in the crash, probably linked to **organizational influences**. Finally, on July 1, 2002, when a *Boeing 757* and a *Tupolev Tu-154M* collided in mid-air above Überlingen, Germany, the investigation showed that the Air Traffic Control (ATC) officer informed the pilots of the airplanes too late of the collision risk, because he had been busy solving another problem, and his colleague was taking a break. In this case, once more, the human factor was involved, because of **unsafe supervision** linked to organizational influences.

These three examples exhibit the fact that the human factor plays a role at every stage of the accident. Accordingly, the comprehensive models developed to explain the causes of accidents (e.g. Reason (1990), Shappel and Wiegmann (2000)) put in relief the multifactorial nature of plane accidents. As shown in Fig. 2, the human may play a role both at the organizational, supervisory, and piloting stage. In practice, most of the accidents involve unsafe acts (i.e. piloting errors, see Fig. 1), so that the reasons of the errors committed by the pilot has become one of the main concerns in aviation safety.

The importance of human error in the failure of complex systems is not the prerogative of the aeronautical field. For instance, during the *Apollo 13* mission, it took no less than 54 minutes for the crew to understand that an explosion had occurred in the oxygen portion of the cryogenics system (Woods et al. 1994), due to the unexpectedness of the situation and to the complex information presentation. Similarly, in the medical domain, the complexity of surgical procedures as well as the pressure caused by life-threatening situations has been identified as one of the main causes of surgical errors (Bromiley 2008). Nevertheless, the culture of human error is especially predominant in the aeronautical field, compared to the others. As stated by Bromiley (2008), when commenting on the surgical accident that led to the death of his wife during routine surgery :

"In my world of aviation, error is accepted as normal, not poor performance or weakness. The trick is not to eliminate error completely but to maintain a culture which, through system, process and training, increases the probability of small errors being caught before they become errors that harm."

The model proposed in Fig.2 insists on the **preconditions for unsafe acts**, i.e. the **operational** (e.g. a risky situation, due to weather conditions for instance) and the **psychophysiological** conditions (stress, fatigue, etc...) likely to be at the origin of operational errors. In particular, an interesting aspect is the contrast in the **workload** imposed to pilots between very low demanding situations, and very highly demanding situations when

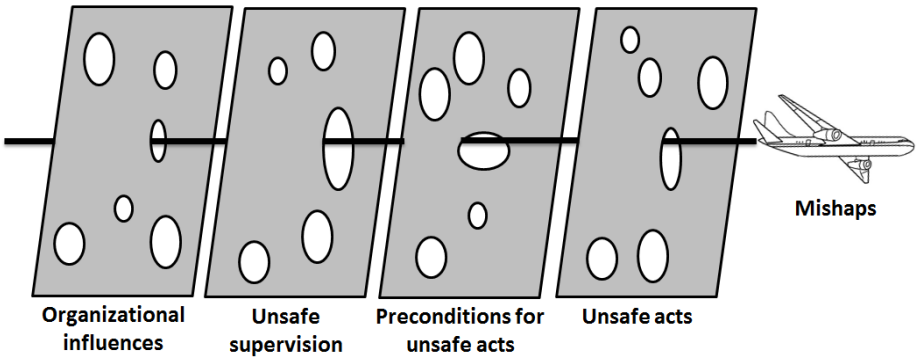


Figure 2 – *Reason's "swiss-cheese" model of the latent and active causes of aviation mishaps, illustrating the fact that only a conjunction of multiple causes leads to plane accidents. Adapted from Reason (1990).*

decisions have to be made and applied accurately and quickly. One of the reasons presented to explain this fact is that aviation procedures promote automated flight controls which are nowadays in use for the major part of the flight. Consequently, a poll conducted in 2013 among British Airways pilots revealed that more than half of them had already fallen asleep during a flight due to drowsiness and low demanding procedures (BALPA 2013). Even worse, a third of them admitted that they had woken up in the cockpit only to find that the other pilot was also sleeping. On contrast, accident reports show that the fate of the plane often depends on time-critical decision making abilities, as every second counts in a critical situation. As a consequence of these contrasting situations, the pilots must be able to switch very dynamically between **attentional states** adapted to varying levels of demands. However, the psycho-physiological state of the pilot coupled with the way information and controls are accessible in the cockpit does not always allow much flexibility. Consequently, operators can miss critical information or persist in erroneous strategies. Indeed, it is believed that the level of workload can impair the attentional processes needed for complex operating tasks. In the attentional resource allocation lies at the heart of the **operator functional state** framework (Hockey 2005), where sustained performance is assumed to be determined in part by the cognitive potential of the operator in relation to task goals and priorities. According to this framework, operators placed in highly demanding tasks will be able to sustain a good level of performance as long as the task remains predictable, but will fail to perform well in the event of an unexpected change, i.e. when the operator faces workload transitions. It is possible that this vulnerability period – when the operator may not be able to adapt to changes in task priorities – could be detected by assessing his/her **attentional state**.

A drastic way to mitigate human error would be to limit as much as possible the role of the human in the operational procedures, or in other words the pilotless plane concept. These considerations are driven by the recent expansion of automated flight controls and unmanned air vehicles for military and civilian purposes. Nevertheless, one should consider that the accident rate of these highly automated systems are very high: for example,

the Customs and Border Protection agency (BPC), which flies unmanned systems on border patrols, has a accident rate more than 353 times the commercial aviation accident rate (1 accident every 1850 flight hours) despite the use of a safety ground pilot (Haddal and Gertler 2010). Worst, another report from the MIT International Center for Air Transportation points out an accident rate 5000 times higher for UAVs than commercial aviation (Weibel and Hansman 2006). Ironically, if the technological and procedural improvements have led the accident rate to decrease drastically, the introduction of new technologies (and particularly the raise of automation) have created new types of accidents as the operators interact less with the system and trust automation too much (Funk et al. 1999). Moreover, these approaches deny the positive role of the pilot during flights (Pinet 2011). For instance, in 2009, a pilot successfully landed an US Airways *Airbus A320* on the Hudson River in New York City, after it had been damaged by multiple bird strikes. Similarly, in 2010, the crew of a Qantas *Airbus A380* successfully landed after the explosion of one of the reactors during the flight.

As a result, if we admit that human intervention cannot be eliminated from the operation of complex systems, research should focus on solutions to avoid human errors. However, the complexity of the cognitive mechanisms involved in operational situations usually requires an approach that goes further than behavioral observations and subjective ratings (Parasuraman et al. 1999). In fact, their understanding depends on our ability to describe the neural substrates and mechanisms responsible for human errors. In particular, as discussed earlier, taking into account the attentional state of the operator in systems deployment could be a way to anticipate human errors.

The rise of Cognitive Neuroscience in the 1960s started to shed light on the mechanisms underlying human attention and performance. The application of this knowledge to the understanding of real-world situations gave birth to the field of **Neuroergonomics** (Parasuraman 2003). The driving idea of this field is that Neuroscience can add significant contributions to the understanding of the mental state in operational situations, and that the evaluation of the brain at work using neuroimaging techniques can help us develop new solutions to improve the operator performance. This ambitious goal has two major challenges. First, applying brain imaging techniques to real-world situations supposes a less-controlled experimental environment than in laboratory experiments. Consequently, Neuroergonomics usually follows a gradual approach, ranging from well controlled but less realistic laboratory conditions to poorly controlled but realistic situations (see Fig.3). In addition, the application of brain measurements to complex environments faces technological limitations, as innovative signal processing and analysis techniques have to be applied for the data to be usable. Despite these difficulties, Neuroergonomics would aid in the avoidance of human errors. On one hand, the knowledge of the cognitive processes involved in critical situations could help us define operational recommendations more adapted to human cognition. On the other hand, the use of brain activity as an

input in Brain Computer Interfaces would provide a solution for assisting the operator in critical situations, by adapting the level of automation or sending feedback to the user. This contribution would be especially valuable as economical motivations and current research take us closer to the "Single-Pilot Operations Concept" (see Wolter and Gore (2015)), which holds that only one pilot would be required to remain at the controls for extended periods during the cruise phase while the other crew member would be in an area outside the flight deck.

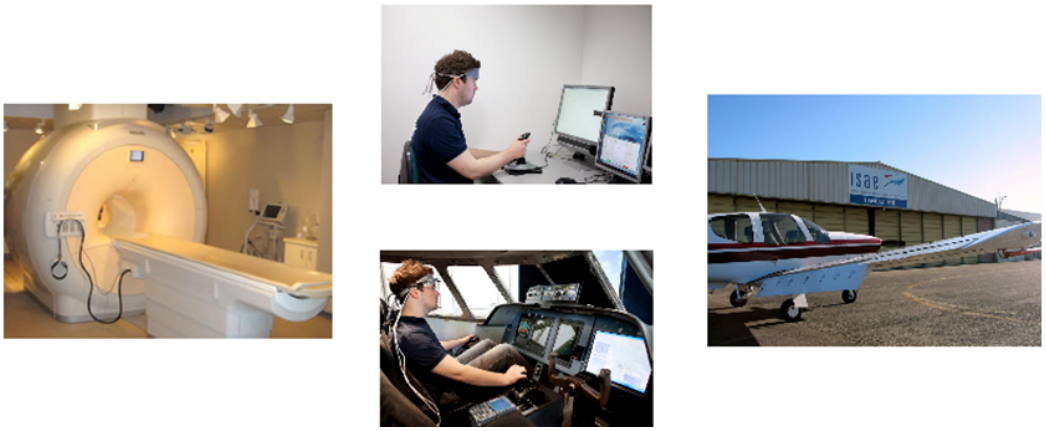


Figure 3 – *Illustration of the approach of Neuroergonomics as proposed by Parasuraman and Rizzo (2007): from highly controlled but less ecological situations (on the left) to less controlled but more ecological situations (on the right).*

This PhD project adopts a Neuroergonomics approach, and aims to investigate how the use of brain measurement techniques to assess the attentional state can lead to new **frameworks** and **solutions** towards the improvement of human performance in ecological conditions. Chapter 7.1 of this dissertation is dedicated to the presentation of the general concepts used to describe human cognition and performance when facing task demands, with a particular focus on extremal levels of demand. The next two chapters (7.2 and 7.3) will introduce the general concepts of Neuroergonomics and the tools it provides to improve human-machine interaction. In particular, the notion of Brain Computer Interfaces will be presented, conceptualizing the use of brain measures to improve interaction. Then, the last chapters (7.4, 7.5 and 7.6) of the dissertation will present the contributions of this PhD project, first concerning the use of brain measurement techniques to assess the attentional state under ecological situations in high (chapter 7.4) and low (chapter 7.5) demanding situations. The technological methods developed during the PhD project to allow the use and analysis of brain measures in a cockpit will be presented in chapter 7.6. The results obtained will be finally discussed in chapter 7.7, to provide a framework as well as innovative signal processing techniques for the use of brain measures as a mean to monitor the attentional state.

II

State of the art

Performance and Task Demands

1

- *Cognitive resources mobilized to face task demands define a level of mental workload.*
- *Mental workload is a descriptive variable, but does not account for the causes and types of resource allocation.*
- *The notion of engagement accounts for the management of workload conducted by the operator.*
- *Taken together, mental workload and engagement explain performance and executive failures.*

Descriptive models have been proposed to identify latent causes of human error, especially in the aeronautical field, and a large volume of research has highlighted the deleterious effects that high mental demands can have on operator performance. When operators must perform very various tasks, in dynamically changing environments, the limited nature of human mental resources supposes that it is impossible to perform well under very high task demands. It is therefore important for the operator to reallocate mental resources dynamically according to changes in task demands, but dynamic reallocation poses a challenge for human cognitive control.

In this chapter, the question of the link between task demands and human performance is addressed, as well as the different factors and mechanisms responsible for mental resources allocation. In particular, we consider the behavioral response to changes in task

demands, emphasizing the deleterious effects of both high and very low mental demands on operators.

1.1 The concept of mental workload

1.1.1 Task demands and resources

The execution of a task by an individual requires the involvement of several resources (physical, cognitive, attentional). One has to mobilize enough resources to capture the relevant information from the environment (task instructions, task stimuli, etc...), to use this information to perform an action, and to actually perform this action.

Many capacity theories put in relief on the *limited processing capabilites* of human brain (Kahneman 1973, Wickens 1981). Indeed, first structural theories of mind pointed out the existence of a *single channel bottleneck* in human cognition (Craik 1948), suggesting that it is impossible to perform two tasks simultaneously as well as separately. Although first theories postulated that *time* is the resource impossible to share between tasks (Craik 1948), Moray (1967) proposed the idea that the brain contains a *Limited Capacity Central Processing Unit* (LCCP), able to perform one operation at a time. According to this theory, the LCCP represents the bottleneck that limits human capacities. From this observation, *resource theories* emerged, putting in relief the competition between tasks for mental resources.

Capacity theories express the idea that when performing an action, we use **mental resources** extracted from the brain's limited capacities (see Fig.1.1). In particular, the idea of an increasing mobilization of cognitive resources to act towards the realization of one or several tasks defines the concept of **mental workload**, which could be seen as the level of cognitive resources necessary to perform a task.

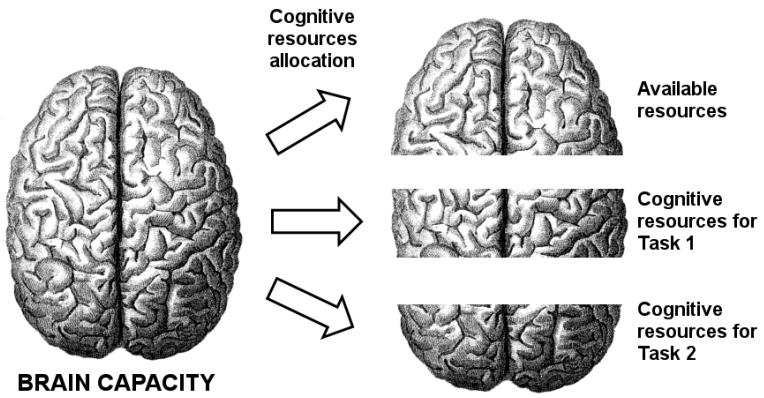


Figure 1.1 – *Schematic approach of brain capacity and resources. To perform a task, the brain disposes of a limited capacity, from which resources are extracted.*

Although the scientific community has not agreed yet on a definition of **mental workload**, its interpretation as a brain effort to perform a task is commonly accepted. For instance, Ayaz et al. (2012b) explained that *"mental workload reflects how hard the brain is working to meet task demands"*.

1.1.2 Nature of mental resources

The exact nature of the resources involved in the performance of a task remains unknown (Wickens 2002), making it ambiguous to evaluate the level of mental workload. Indeed, neurophysiological evidence converges to the fact that mental resources may have correlates at all levels of the cortical system : neuronal, glial and circulatory (Carpenter et al. 1999). At the neuronal level, mental resources could correlate with the quantity of neurons engaged in a task at a given time, but also their excitability levels or structural connectivity. At the circulatory and glial levels, mental resources also have a link with the state of glial cells and cortical vasculature, responsible for the appropriate metabolic and neurotransmitter functioning of the brain. All cortical cells, representing on one side dingle processing units (neurons) or active supporting units (neuroglia), would represent the mental resources mobilized for task completion and for which multiple tasks would have to compete. More generally, a common approach (Kahneman 1973, Hockey et al. 1986) postulates that mental resources can be viewed as an *energetical* concept, i.e. as a mental **effort** to perform a task. This view is also justified by the high energetic costs required by cortical cells during action. Therefore, measuring the exact level of mental resources involved at a given time and interpreting the results of neurophysiological measures remains complicated.

In addition, another question concerns the specificity of the resources to be allocated, and is at the origin of the two main resource theories. In Kahneman's **Resource Theory** (Kahneman 1973), the brain consists of a *unique pool* of resources, from which resources to perform the task are extracted. As schematized on Fig. 1.2 (left), this theory was directly inherited from the idea of a LCCP (Moray 1967), as the bottleneck constituting resource allocation stands for the limited brain capacity. In this model, a task would require a certain amount of resources to be performed correctly, leaving all the resources unused for other tasks.

Kahneman's resource model captured elegantly the limitation imposed by brain capacities. However, the greatest challenge to this theory came from the analysis of concurrent tasks being performed simultaneously. Indeed, instead of limiting the reflexion to a unique pool of generic resources, some studies (Kantowitz and Knight 1976, Wickens 1976) investigated the efficiency of resource allocation in multitasking in terms of the type of processing required to perform the concurrent task. These studies led to the evidence

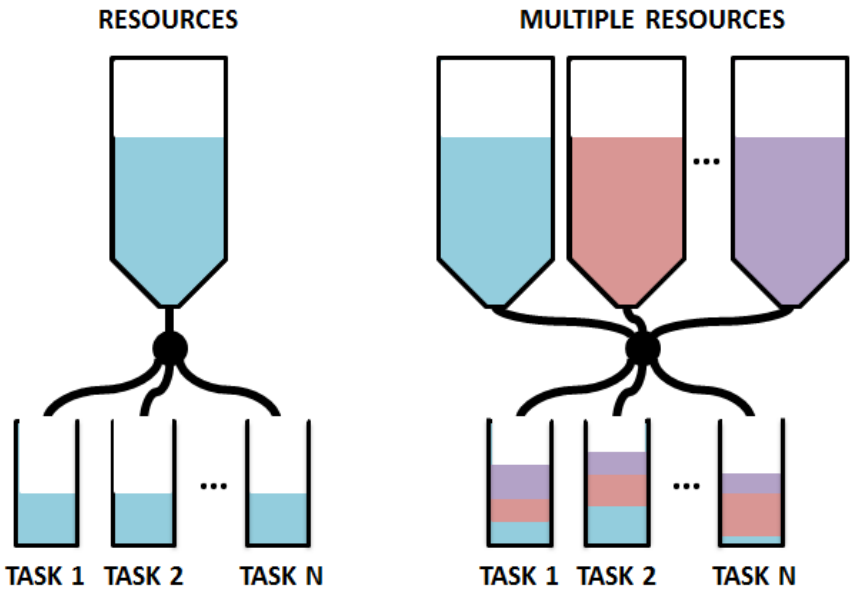


Figure 1.2 – *Schematic structural description of Kahneman's resource theory (left) and Wickens' Multiple resources theory (right). Both theories rely on the idea that resources are needed to perform a task. However, in Kahneman's theory, the resources are extracted from a single pool of resources, therefore limiting the resources available for other tasks. Wickens' theory proposes that the resources are organized in a multi-dimensionnal structure, i.e. as different pools of independent resources. The limitation imposed by the performance of a specific task would therefore strongly depends on the processing structures required.*

that dual-task performance was less affected when the two tasks required different processing structures, than when they used the same structure of processing. For instance, two visual tasks are likely to interfere when being performed simultaneously, so that the operator would not perform as well on both tasks than he would do on each task separately. However, this effect disappears when one of the tasks uses the visual modality, and the other uses the auditory modality (Kantowitz and Knight 1976). Those observations led to the hypothesis that the resources needed for those tasks would not come from a single pool of resources, but rather from different pools of independent resources.

Wickens (1980) proposed the **Multiple Resource Model** that postulates the existence of several resource pools (see Fig.1.2, right). Based on a review of concurrent tasks performance, he described the multiplicity of resources (Navon and Gopher 1979) as structured along four dimensions :

- **Stages** of processing : the resources required for the different stages of processing are different in nature. Indeed, perceiving or processing the information requires resources independent from the ones required for selecting and responding to the information.

- **Perceptual modalities** : the resources required for the different perceptual modalities (e.g. visual, auditory) are relatively independent, so that auditory tasks would barely interfere with visual ones, for instance.
- **Visual channels** : among the visually processed information, Wickens proposes a dichotomy between focal information (associated with foveal vision) and ambient information (associated with peripheral vision).
- **Processing codes** : Wickens proposes as the fourth multiplicity of resources the two types of processes associated with information processing. On one hand, spatial processes, and on the other hand categorical processes (usually verbal).

This arrangement of mental resources in a four-dimensional space and, more generally, the vision of multiple pools of resources rather than a unique pool helped Wickens develop one of the most comprehensive models of resource-linked human performance and workload. However, this subdivision of Resource theory also added more confusion to the mental resource concept, as the type of resources (attentional, visual, etc) mobilized for a task could be used to define a level of attentional load, visual load, and so forth. Progressively, such subdivisions of mental workload appeared in the literature (e.g. Lavie and Tsal (1994), Rees et al. (1997)), making the link between resources and behavior even more complex, and therefore questioning the relevancy of the resources approach as an explanatory approach to human behavior.

1.1.3 Cognitive control and resource allocation

Even if the exact nature and amount of cognitive resources remains difficult to evaluate, we understand that a key step towards the accomplishment of a task resides in their allocation. Specifically, an optimized process of cognitive resources allocation would help improve performance but also user's comfort (Stassen et al. 1990). Most of cognitive and attentional control theories (the central executive from Baddeley and Hitch (1974), or the Multiple Resource model from Wickens (1981), for instance) focus on the dynamic allocation of these cognitive resources, vital for the achievement of a task. The variety of factors influencing the way resources are allocated have exogenous (*bottom-up*) origins as well as endogenous ones(*top-down*), both influencing the level of mental workload.

Indeed, there is a fundamental difference between the level of cognitive resources imposed by a task, defining what is called *task load*, and the way the individual effectively allocates the resources to perform the task, which defines mental workload (Hilburn and Jorna 2001) and is directly related to the performance (Wickens 2002), according to this theory. On this basis, an individual who would not allocate enough resources relatively

to task demands would perform badly at a task, and if the level of task load exceeds the actual capacity of brain resources, the task would be impossible to perform.

The allocation of attentional and cognitive resources (which can be used to define a level of attentional load or mental load) has been considered by many models of attentional and executive control, emphasizing the role of the prefrontal cortex in the allocation of resources. The Executive control theory from Baddeley and Hitch (1974) supposes the existence of the **Central Executive**, whose main substrate is thought to be located in the prefrontal cortex (Ridderinkhof et al. 2004). Its role is to control and regulate cognitive processes, by allocating the resources to the various tasks. Similarly, the various comprehensive models of attention deal with the allocation of resources. For instance, Norman and Shallice (1986) proposed the definition of a **Supervisory Attentional System** (SAS), regulating the executive control functions in the brain, located in the prefrontal cortex. Posner's model of attention (Posner 1980), as well as Laberge's model (LaBerge 1997), also describe an **Anterior Attentional Control Network** (for Posner) and an Attention Control network (for Laberge), centrally involving the prefrontal cortex.

Other models describe a **cascade of top-down control** mechanism for the allocation of resources, such as the models proposed by Koechlin et al. (2003) or Banich (2009). First, **episodic control** allows the switch between different representations of a given task (i.e. different strategies of allocation). Then, **contextual control** and **sensory control** are responsible for the switch between different stimulus/response association, and the selection of the accurate responses. Similarly to the previously described models, the prefrontal cortex is thought to be one of the major correlates of the cascade of control mechanisms, especially for episodic and contextual control.

Several attempts at designing comprehensive models to understand workload variations from an internal point of view have been made. More generally, many models have focused on the evidence that mental resources allocation is done in a different way depending on the situation, which leads to different **modes** of resource allocation. For instance, Rasmussen (1983) proposed the "*Skill-Rule-Knowledge*" model, supposing that three strategies of resource allocation exist, and sharing common concepts with the *controlled vs. automatic* modes of cognition directed by the SAS (Norman and Shallice 1986). *Skill based* resource allocation is encountered in very experienced operators, when the task is well known. In this mode, the task will be treated quasi-automatically, with a minimum mental resource involvement (see for example Ayaz et al. (2012a)). *Rule based* mode is encountered at medium experience levels, when the task can be partially automatized as the reactions to specific events is already mastered. The behavior of the individual can in this case be explained by "If-then" rules (the ACT-R cognitive architecture proposed by Anderson (1996) relies on these principles). Finally, *Knowledge based* mode occurs when the

operator encounters a novel task, or has to try a novel approach, requiring the exploration of potential behaviors.

1.1.4 Factors influencing resource allocation

The question of the impact of dynamic workload changes on performance and on the way resources are allocated is at the center of **workload transition** research. On this topic, Wickens et al. (1993) provided a framework regrouping the five different types of factor likely to bring abrupt variations in the level of workload :

- **Task character** regroups the general aspects of the task likely to affect the behavior of the operator. For instance, we could consider on one hand a pilot facing an alarm signaling a major defective component, and on the other hand a train conductor facing an immediate danger. Even if the two situations are life threatening, the nature of the tasks is different enough to lead to radically different reactions from the two operators, with very different resource allocation.
- **Input** factor refers to the way information is presented to the operator, as well as the way the operator perceives these information. If the information is well presented, the level of resources needed for their gathering will be lower. On the contrary, complex or badly presented information will require higher level of resources (Wickens et al. 1983b).
- **Processing** factor refers to the way the individual processes the information and decides his reactions. It comprises all the internal processes affecting resource allocation (see below).
- **Output** factor refers to the different means of action the operator has to modify the system. It regroups all the commands, controls and interfaces designed in the system. As for the "input" factor, a well designed interface will lead to intuitive controls, requiring few resources to perform an accurate action (Wickens et al. 1983b).
- **Computer aiding and Automation** regroups the other actors involved to assist the operator in his tasks. Indeed, the implementation of automation should serve human-machine interaction and lower the mental workload of the operator (Parasuraman and Riley 1997), but misused automation can also lead to the transformation of a type of load into another and the increase of the operator's load (Hart and Sheridan 1984).

Three out of the five direct factors of workload transition proposed by Wickens and Huey concern the properties of the task (or of the system the human interacts with). Only

two of them ("*input*" and "*processing*") embrace the specificities and the state of the individual. Nevertheless, in practice, it is hard to know which of the relationship presented in Fig. 1.3 (positive linear curve, saturation curve or inverted-U curve) is likely to happen during a given task, because workload (i.e. the level of mental resources engaged in the task) cannot be estimated precisely with the sole properties of the task. Indeed, several factors at the individual level will impact their deployment. In particular, a variety of factors including the level of *stress*, *motivation*, *time pressure*, *fatigue* (Hart and Staveland 1988), as well as *expertise* (Bunce et al. 2011) or the *engagement* in the task (Lee 2014) lead to different resource mobilization even with constant task demands. In other words, task load is relatively independent from mental workload, so that mental workload should be defined as an interaction between the task and the individual performing the task (Hancock and Chignell 1988). As an analogy, Rohmert (1986) proposed to view mental workload as the deformation of a structure, and task load as the demand imposed on it (i.e. the mechanical constraints imposed by the external environment). From this point of view, the deformation experienced by the material would not only depend on external factors (i.e. task load), but also on a set of internal factors (e.g. properties of the material, shape of the structure).

1.2 Mental Workload, Engagement and Cognitive Performance

Although it is now clear that the concept of mental workload relies on the mechanisms responsible to resource allocation when facing varying task demands, the actual mathematical relationship between those two variables is still unclear. Its study can lead to different possible interpretations, summarized in Fig. 1.3 :

- A **positive linear curve**, stated for example in drive theory (Hull 1943, Spence 1958). In this model, increasing task demands would lead to the mobilization of proportional resources.
- A **saturation curve**, consistent with most attentional models (Norman and Shallice 1986, Baddeley and Hitch 1974, Wickens 1981), stating that the mobilization of resources would grow with task demands, until a maximal value is reached (in most theories of attention, this maximal value is flexible (Stokes et al. 1990, Young and Stanton 2002)).
- An **inverted-U curve** (Goldberg et al. 1998, Callicott et al. 1999), stating that the allocation of cognitive resources follows a quadratic law. In this model, higher task demands would lead to disengagement, i.e. a decrease in resources allocated to the task.

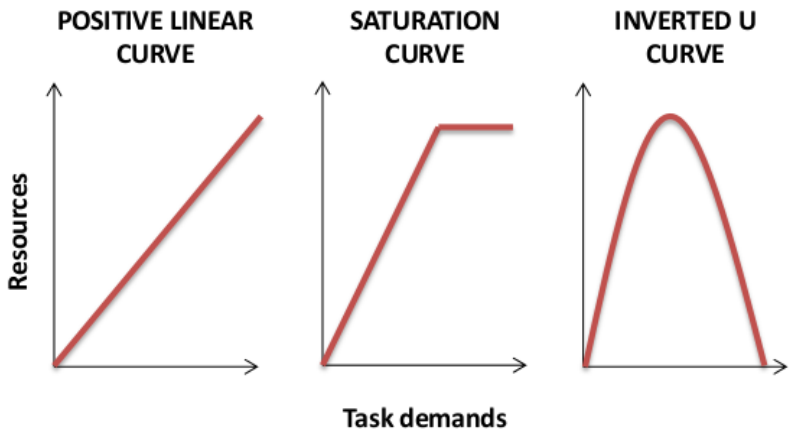


Figure 1.3 – *Representation of the different possible relationships between task demands and resource allocation. (left) Positive linear curve (Hull 1943) ; (center) Saturation curve (Stokes et al. 1990) ; (right) inverted-U curve (Goldberg et al. 1998, Callicott et al. 1999).*

On the same basis, Sullivan and Bhagat (1992) reviewed the relationship between task demands and cognitive performance. They exhibited the same potential mathematical relationships (linear curve, saturation curve, U-shaped curve), although the most commonly model referred to is the inverted-U shaped pattern linking mental workload and performance inspired from Yerkes and Dodson (1908) (Boyer 2014, Woods 2000).

Similarly, one could argue that mental workload is sufficient on a **descriptive level**, because the notion of cognitive resources is a key feature to understand human action. However, this concept remains highly ambiguous on an **explanatory level**, because the causes for resource allocation remain unknown. In summary, the concept of mental workload can bring an answer to the question ***What?*** (by quantifying a level of cognitive resources gathered to face task demands), but not to the questions ***How?*** (How are cognitive resources allocated, and based on which motivations ?) and ***When?*** (What is the dynamic link between task demands and resource allocation ?). In fact, the answer to these questions require a reasoning directed towards the individual rather than the task.

Indeed, as seen before, most models of workload describe it as the adaptation of the operator to meet task demands, provided that resource allocation depends on task demands. Alternatively, Sperandio (1971) proposed additional concepts. In his mind, mental workload should not be viewed as the result of an external demand applied on a individual passively adapting to it, but rather as something the operator actively manages, by a varying level of **engagement**. From a mathematical point of view, the level of cognitive resources involved in the task (defining the level of mental workload) can be considered as the **product** of a level of task demand by the engagement on the operator. This notion of workload regulation is also present in the work of Pope et al. (1995), defining the notion of engagement by stating that the interaction with systems requires the maintenance of sustained attention directed towards the task. Sperandio's *closed-loop* model of mental

workload describes the idea that the mental workload imposed by the strategies the operator picks influences both the task demands encountered in the future and the strategies the operator will choose in the future. Therefore, the changes in operator strategies (i.e. variations of engagement) occur as an answer to changes in mental workload rather than task demands. For instance (Segal and Wickens 1990), a pilot facing peaceful flying condition will probably keep in mind passengers' comfort and his own pleasure while piloting. On the contrary, when facing hard flight conditions leading to high workload levels, the same pilot will probably adapt his own objectives, leaving aside passengers' comfort to engage himself in the objective of security.

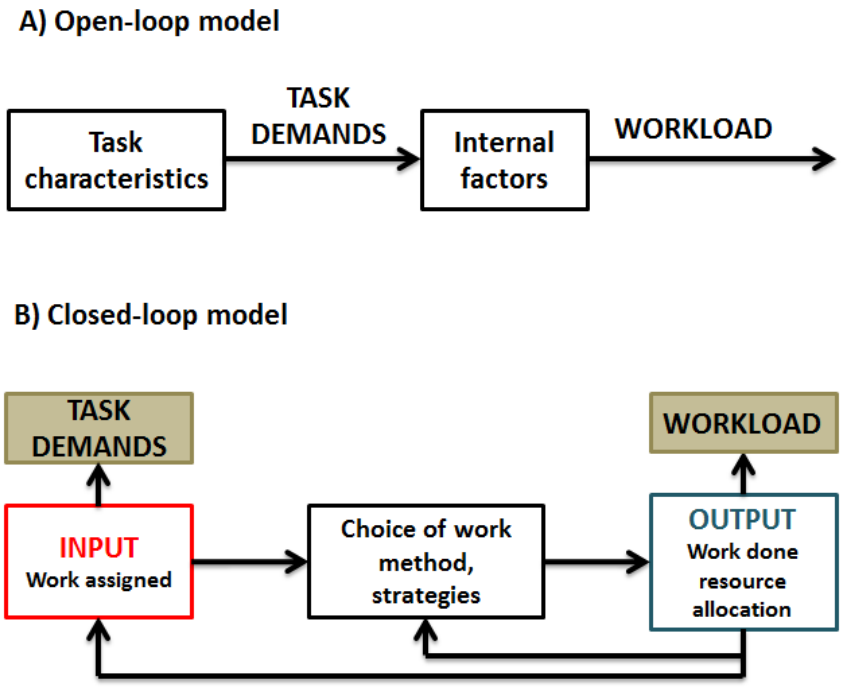


Figure 1.4 – *Systemic models of mental workload, adapted from Loft et al. (2007).*
A) Open-loop model of mental workload, consistent with most theories. The imposed task demand, depending on the characteristics of the task assigned, leads to the allocation of mental resources depending on internal factors.
B) Closed-loop model of mental workload, proposed by Sperandio (1971). In this model, strategies chosen by the operator are the intermediate between task demands and workload. Thus, the operator actively manages the level of workload to keep it within acceptable levels.

The two approaches are summarized in Fig.1.4, schematizing on one hand the open-loop approach, considering that resource allocation depends on task demands and a set of internal parameters, and on the other hand the closed-loop approach, considering resource allocation as a managed strategy, influenced not only by task demands and internal parameters, but also by the level of mental workload effectively exerted on the operator. This point of view puts into relief the fact that an individual, by managing the level of workload, **engages** himself/herself in the task, therefore explaining the existence of different modes in resource allocation as described in section 1.1.3 (e.g. controlled vs. automatic

mode). The use of operator engagement as a variable allows the explanation of resource allocation (Lee 2014). The management of workload and user engagement are served by the mechanisms of estimation of **cognitive costs** and **rewards** associated with a given strategy, playing an important role in adaptive behavior (Donoso et al. 2014).

The regulating effect of mental workload leading to different levels of **engagement** is particularly present when facing extreme workload levels (Sperandio 1971, Segal and Wickens 1990), as the operator may decide to disengage from the task (see section 1.2.1.3). Thus, understanding the different mechanisms involved in the adaptation to workload changes can benefit from the consideration of the operators facing such conditions (Boyer 2014). Indeed, understanding what happens on a behavioral and cognitive point of view when an individual deals with conditions of very high workload on the first hand and very low workload on the second hand can bring precious clues towards the understanding of the influence of the factors determining the mechanisms of resource allocation.

1.2.1 Human behavior and engagement in highly demanding situations

Highly demanding situations have been described as one of the main causes of human error in aviation. The rise of automation and computer aiding was initially developed as a way of decreasing the task demands imposed on operators.

Indeed, high mental demands are often associated with executive failures, such as erroneous strategies or strong decreases in performance, as synthesized on Fig. 1.5. The reasons to these behaviors can be found in the mechanisms occurring in highly demanding situations due to the engagement of the operator : Attentional tunneling, Perseveration and Disengagement.

1.2.1.1 Over-engagement at the attentional level : tunneling

Highly demanding levels of workload often occur in dynamical environment, with numerous sources of information (e.g. captors in the cockpit), leading to high levels of mental load both to capture the information (**perceptual load**) and to process it.

When facing such demanding situations, the operator may tend to focus his attention on a specific source of information, therefore paying less attention to the environment (Regis et al. 2014). This phenomenon, called **attentional tunneling**, may lead the operator to miss critical information (Dehais et al. 2011), and eventually cause a loss of **situation awareness**.

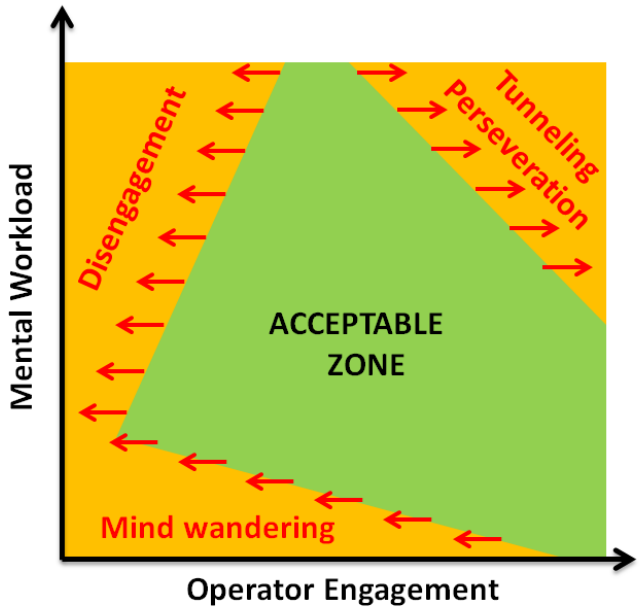


Figure 1.5 – *Operation patterns in function of mental workload (Y-axis) and operator engagement (X-axis). The acceptable zone for operation leading to reasonable performance is represented in green. In case of high mental workload (top), the operator can disengage gradually from the task (red arrows on the left) in case of low engagement. On the contrary, when user engagement is too high, mental workload can lead to attentional tunneling or perseveration (red arrows on the left, see 1.2.1.1. In case of low mental workload (bottom), the phenomenon of mind wandering leading to user disengagement can occur (see 1.2.2.1).*

The origins of this phenomenon rely in the attention orienting networks, as described by Posner (1980). Some studies showed that high levels of workload and stress may impair the networks responsible for attention orientation, mainly supported by the thalamic nuclei (LaBerge et al. 1992). As a consequence, the operator would be unable to disengage his attention from a given source of information.

States of attentional tunneling have often been associated with a reduction of the attentional span (Dehais et al. 2011). From this point of view we understand the term of *tunneling*, as the operator remains "locked" in a narrow visual area. However, it is important to note that the negligence of some sources of information can happen even without tunneling in a narrow visual area. Indeed, it has also been described in the phenomenon of **inattentional blindness** (Mack and Rock 1998) or **deafness** (Dehais et al. 2012). In her review, Lavie et al. (2014) describes that high task loads can make an operator virtually blind or deaf to some stimuli, so that he will remain focused on the same type of information.

1.2.1.2 Over-engagement at the executive level : perseveration

If at the attentional level, excessive mental demands can impair the ability of the operator to attend crucial sources of information (i.e. tunneling), they can also impair cognitive functions at the executive level. Indeed, high mental demands placed on operators and degraded environmental conditions may have an impact on the operator, limiting his capacity to adapt his strategies. This state of perseveration has been categorized by Dehais et al. (2010) as one of the causes of human error.

In their review, De Keyser and Woods (1990) and Woods et al. (1994) propose a classification of the situations in which the operator fails to adapt his behavior in demanding situations, named **fixation errors**. The pattern of behavior observed in those situations describes different types of fixation errors :

- The operator can have many strategies in mind and try them successively, jumping from one action to another, but seems to never entertain the correct strategy. This pattern is named ***everything but that***.
- On the contrary, the operator can be stuck on one strategy, seeming unable to shift or to consider other possibilities. These situations are named ***this and only this***.
- In the last situation, called ***everything's OK***, the operator fails to react to a change in his environment, despite the perceptual clues that the situation has changed and needs a reaction. This pattern is close to attentional tunneling described in 1.2.1.1, although the negligence of some sources of information can be voluntary in this pattern.

From the neurophysiological point of view, research suggest that the substrates at the origin of fixation errors rely in the basal ganglia. Indeed, Donoso et al. (2014) suggested that the nigrostriatal pathway of the dopaminergic system, already known to be involved in the initiation of voluntary motor actions, may be implied in the elaboration of new strategies. Stress induced by highly demanding situations has been shown to impair the functioning of the dopaminergic system (Arnsten 2000), therefore providing a potential explanation of the observed behaviors.

1.2.1.3 Mental overload and the process of disengaging

As discussed in part 1.2 with the work of Sperandio, the operator constantly manages his/her level of mental workload to maintain an acceptable level of work. In particular, the ratio between the amount of work actually accepted by the operator and task demands

(defined by the task itself) can be used to define a level of **engagement** in the task. When facing extremely high levels of workload, the operator may choose to disengage from the task (Lee 2014).

This pattern of disengagement, traditionally neglected by cognitive neuroscience in which one usually assumes that the operator is always fully dedicated to the task (Parasuraman and Rizzo 2007), is described as one strategic factor by Woods (2000). It relies on the fact that the operator constantly acts towards the realization of multiple goals (Bedny and Karwowski 2004) in a continuous process. In particular, these multiple goals can be associated with the development of multiple strategies (c.f. Donoso et al. (2014)), involving the estimation of the **cost** and **reward** associated with them. Thus, if one goal is considered too costly (because too demanding) compared to its reward (i.e. the importance of its outcome for the individual), the operator can shift his attention towards another personal goal. It is specifically in this case that the inverted-U shape mode of resource allocation described on Fig. 1.3 occurs.

1.2.2 Human behavior in low demanding situations

Research on cognitive performance usually focuses on highly demanding situations, because of the deleterious effect of mental workload on performance, as described in section 1.2.1 and represented by the right side of the inverted-U shaped curve linking workload and performance (see 1.2). Hart (2010) showed that the performance decrement induced by low workload (left part of the inverted-U curve) is usually less important than the one induced by high workload situations, therefore legitimating the focus on high workload in research.

However, the study of the mechanisms underlying performance in low demanding situations can give precious clues about the architecture of human cognition. Indeed, as we saw in parts 1.2 and 1.2.1.3, the notion of operator engagement is fundamental to understand performance, and is directly linked to the evaluation of costs and rewards associated with a given strategy. The concurrent analysis of various strategies is at the origin of the mechanism explaining performance decrement when facing low mental workload : mind wandering (see Fig. 1.5).

1.2.2.1 Mind Wandering : disengagement under low task loads

Low demanding tasks often elicit attentional lapses (Smallwood et al. 2012a) when subjects' attention drifts from the task towards self-centered matters. These attentional drifts causing low alertness are called mind wandering (MW) episodes, and people are gener-

ally unaware of when they occur. Avoiding these attentional drifts is not only a matter of willpower since even after years of practice focused meditation, meditators still experience these drifts regardless of their efforts to avoid them (Braboszcz and Delorme 2011). Although positive effects of mind wandering have been exhibited concerning creativity and incubation during problem solving (Mooneyham and Schooler 2013), studies showed that the attentional span and the cortical processing of the external environment is decreased during episodes of mind wandering (Smallwood et al. 2008, He et al. 2011), and it may jeopardize safety in operational situations such as driving a vehicle (Galéra et al. 2012) or flying a plane (Casner and Schooler 2013).

The generation of task unrelated thoughts during mind wandering exhibits the concurrent action of two distinct networks during task execution (Johnson and Zatorre 2006, Harrivel et al. 2013) : a **Task Positive Network**, represented by the Central Executive Network (Baddeley and Hitch 1974), and a **Task Negative Network**, represented by the **Default Mode Network** (Andrews-Hanna 2012), particularly active during mind wandering (Mason et al. 2007a, Christoff et al. 2009, Durantin et al. 2015). The opposition between phases of dominance of these two networks describes the dynamics of mind wandering (Hasenkamp et al. 2012). As discussed by Spreng (2012), the name "*Task negative network*" can be misleading, as this network plays an active role in goal-directed cognition, and not only a role of opposition to action. Indeed, it is believed that mind wandering enhances autobiographical planning (Baird et al. 2011), i.e. helps the individual to keep personal goals in mind. Therefore, if it is true that mind wandering may be seen as an executive failure, as it reduces the performance of an individual at a given task by causing disengagement, it may also have a functional role in goal-directed cognition, and therefore have a positive impact on other tasks that the evaluated one (McVay and Kane 2010).

1.3 Conclusions on mental workload and engagement in the evaluation of performance

As we discussed in this chapter, different mechanisms may affect human performance in operational situations. Those mechanisms, schematized on Fig. 1.5, are explained by two main factors : **mental workload** and **operator engagement**, respectively defining the level of cognitive resources involved in task performance, and the

Most studies focus on the evaluation of **mental workload** as a tool to improve human performance and avoid human error. However, this factor is not self-sufficient to account for all the phenomenons observed during task execution. Indeed, if both high and low levels of workload have been proved to impair human performance, usually schematized in an inverted-U shaped curve linking workload and performance, the explanation of

performance drop requires the introduction of the **operator engagement** as a variable. On one hand, the disengagement of the operator voluntarily (in case of very high mental demands) or involuntarily (in case of mind wandering) is one of the latent causes of human performance decrement. On the other hand, very high levels of engagement in a task observed in case of attentional tunneling or fixation errors also explain the loss in performance due to the persistence in erroneous strategies.

Neuroergonomics 2

- *Neuroergonomics is the application of Neuroscience to Ergonomics, and promotes an approach centered both on the task and the operator.*
- *The estimation of mental workload and engagement is a key step towards human-machine interaction improvement in Neuroergonomics.*
- *Brain measurement techniques have good estimation properties, and should be used in real-time to assess the state of the operator.*

In the previous chapter, we defined the notions of **mental workload** and **operator engagement** underlying human behavior and interaction with a task. Their estimation is one of the main tools to improve human-machine interaction. In particular, the important role of the prefrontal cortex has been emphasized, for its action in resource allocation and engagement. The ways these concepts are taken into account in systems design are studied in the field of **Neuroergonomics**. In this chapter, we define the principles and objectives of Neuroergonomics and the methods that can be used to reach them, with a focus on the methods of measurement of workload and engagement. Different measurement techniques will be presented in this chapter, with a comparison of their performance at mental workload and engagement estimation.

2.1 History of Neuroergonomics

Neuroergonomics is commonly defined as the application of neuroscience to ergonomics. It has its roots in two recent scientific domains : **neuroscience** and **ergonomics** (or hu-

man factors). Historically, these two fields emerged after World War II, following rather different paths.

On one hand, **Ergonomics** emerged from the behaviorist camp of the psychology from the 40s (Parasuraman and Rizzo 2007), and did not pay much attention to brain mechanisms underlying attention, execution of action and their motives. Indeed, it historically followed a procedural approach, considering the action for itself, without replacing it in a broader view (including the goals and motives of action - cf. chapter 7.1). The introduction of Cognitive Psychology in the 60s had an impact on ergonomists, whose interest in the cognitive processes underlying action grew ; however, this impulsion was not enough to get the domain out of its functionalist approach of mind, unrelated to its goals and motives.

On the other hand, **Cognitive Neuroscience** focused on the understanding of the neural processes involved in human cognitive, affective and motor processes. However, the field adopted from its beginnings a laboratory approach (using standardized laboratory tasks), giving little attention to real world settings.

Intriguingly, researchers have long criticized the opposition between on one side a field of ergonomics disconnected from the brain, and on the other side a neuroscience disconnected from the "real world" (Brehmer and Dörner 1993). More generally, even if the relative opposition between two different approaches of science (one centered on the individual and one centered on the systems) has always be noticed, few researchers have tried to combine the two approaches. For instance, concerning life sciences, the neurologist Oliver Sacks wrote in his book "The man who mistook his wife for a hat", that "*Charcot and his pupils, who included Freud and Babinski as well as Tourette, were among the last of their profession with a combined vision of body and soul, 'It' and 'I', neurology and psychiatry. By the turn of the century, a split had occurred, into a* ***soulless neurology*** *and a* ***bodiless psychology***"(Sacks 1998). Similarly, we could argue that until recently a brainless ergonomics science and a bodiless cognitive neuroscience coexisted.

The idea of filling the gap between Neuroscience and Ergonomics is at the origin of **Neuroergonomics**, which emerged at the beginning of the 21st century, mainly driven by professor Raja Parasuraman (Parasuraman 2003). As stated by Parasuraman and Rizzo (2007), their combination would help cognitive science to be liberated from "*a disembodied existence devoid of context, and provide it an anchor to the real world*". In fact, the driving idea of Neuroergonomics is that studying the brain performing ecologically valid tasks of everyday life (driving, walking, ...), helped by the knowledge on brain structure and functions as considered by neuroscience, can add significant knowledge both to the understanding of brain function and to the improvement of the devices used by the operators (a car, a plane, a computer, ...).

2.2 Approaches in Neuroergonomics

As explained in the previous section, reaching the goal of Neuroergonomics relies on the evaluation of brain activity while performing ecological tasks. We understand that this approach supposes to be able to evaluate both which tasks can be considered ecological enough, and how to evaluate the behavior of the individual performing this task.

First, the Neuroergonomics approach raises the question of ecological validity of a task (Brehmer and Dörner 1993). Indeed, the most significant contribution of the domain, compared to Neuroscience, is to explore human behavior in real-world settings rather than in artificial and standardized tasks (which will be called "laboratory tasks" in the following). However, it requires a proper limit between what can be considered as ecologically valid and what cannot.

However increasing the ecological validity of a task requires the involvement of more experimental parameters, leading to an increased dimensionality, and thus an increased difficulty to evaluate the performance and state of the operator. For instance, when piloting a plane, how is the performance defined ? Should we only consider the presence or absence of critical human errors, or should more parameters be taken into consideration ? Fig. 2.1 presents the different settings that can be used for an experimental approach :

- **Laboratory settings** use standard tasks, generally isolating one or two cognitive functions and parameters.
- **Micro-world settings** are computerized simulations recreating real operational conditions (and therefore requiring a variety of cognitive functions). Simulated environments allow us to control all the experimental parameters (e.g. acquisition parameters, difficulty, events...)
- **Real world settings** consist of recordings in realistic situations.

The holy grail of Neuroergonomics is the functional study of brain function in real-world settings. However, achieving this goal often requires the step-by-step investigation of all cognitive functions and parameters required in real-world situation. To this end, Neuroergonomics studies usually move gradually from laboratory approaches to real-world settings.

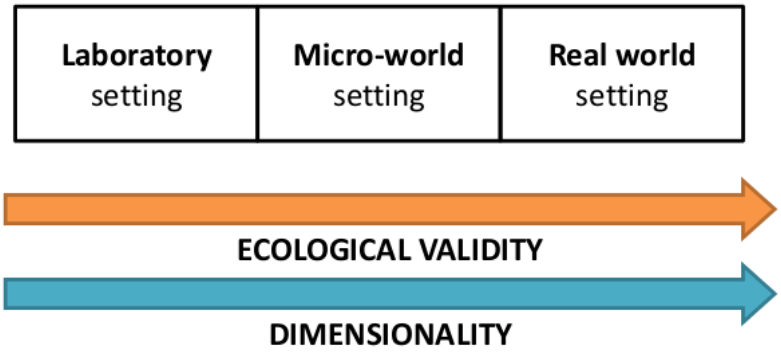


Figure 2.1 – *Variations of ecological validity and dimensionality induced by the different experimental approach : laboratory settings, micro-world settings or real world settings.*

2.3 Measures in Neuroergonomics

As discussed in Chapter 7.1, evaluating the behavior of an individual requires the evaluation of the levels of **mental workload** and **engagement** in the task. Taken together, these parameters describe the potential mechanisms underlying human performance (e.g. perseveration, mind wandering, attentional tunneling or disengagement, see Fig. 1.5). To this end, all the methods used to measure these two variables serve the goals of Neuroergonomics (Parasuraman 2003).

However, the challenge for Neuroergonomics resides in the choice and interpretation of these measures, the use of which should acquaint not only for the level of resources mobilized to face task demands, but also for the causes and temporal dynamics of their allocation. O'Donnell and Eggemeier (1986) provided the framework for the evaluation of mental workload measurement techniques, later precised by Wierwille and Eggemeier (1993). Based on their work, the different measurement techniques should be evaluated given six criteria, ranked by order of importance :

- ***Sensitivity*** refers to the ability of a measure to distinguish different levels of resources mobilized for a task or different levels of engagement. This criterion is therefore usually considered to be the most important quality criterion for a mental workload or engagement measurement. In particular, the sensitivity to resource level changes is particularly interesting to acquaint for the temporal dynamics of resource allocation (Boyer 2014).
- ***Diagnosticity*** refers to the ability to inform about the causes of mental workload and engagement. Often put in relation with multiple resources models (Wickens 2008), a measure of mental workload and engagement is considered *diagnostic* if, in addition to being sensitive to their variation, it can exhibit the contributions of specific types of resources (e.g. auditory resources, working memory resources,

attentional resources etc...). Although this criterion is often considered secondary for mental workload and engagement assessment, it is as important as sensitivity when it comes to Neuroergonomics study, as the potential of explanation for the causes of resource allocation gives precious clues concerning the engagement of the individual at a given time.

- ***Task intrusiveness*** quantifies the modifications of human behavior provoked by the presence of the measure while performing a task. To reach perfect ecological validity, an important issue in Neuroergonomics, one should use only non intrusive measures, so that the behavior of the operator will not be affected by the measure.

- ***Global sensitivity*** defines the ability of a measure to be sensitive to all levels of resources and engagement. Taken together with diagnosticity, these criteria define the *specificity* of a measure. Global sensitivity is a key characteristic if the measure has to be generic, that is if we want to apply it to various types of tasks. It is an important quality for a measure, however a globally sensitive measure also has to be diagnostic, otherwise its benefits will be lost when it comes to the characterization of the causes and modes of resources allocation.

- ***Transferability*** refers to the stability of the measurement features when the measure is repeated. It corresponds to the variability observed when applying the same measure both from one experimental session to the other (*inter-session transferability*), and from one subject to another (*inter-subject transferability*)

- ***Implementation requirements*** comprises all the limitations imposed by the measure. It depends on its cost, its complexity, but also on its real-time capability (i.e. its capacity to give regular measures) or signal-to-noise ratio (i.e. the amount of signal acquired by the device corresponding to mental workload).

2.3.1 Subjective measures

One of the easiest ways to evaluate mental workload and operator engagement is to ask him/her their impressions. This method, commonly implemented by using standardized questionnaires, is the cornerstone of **subjective evaluation** of mental workload. The NASA Task Load Index (TLX) (Hart and Staveland 1988) is one of the questionnaires widely used in neuroergonomics research, and relies on the subjective evaluation of workload along six criteria : *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort* and *frustration*.

In addition to their facility of use, studies suggest that subjective measures have high sensitivity and diagnosticity, and consequently give the most reliable and precise estimations (Chen et al. 2011). However, their main limitations reside in their task intrusiveness

and implementation requirements. Indeed, subjective assessments usually allow only one measurement during the whole task (at the beginning or at the end), or require to pause the experiment (as it is done in experience sampling, e.g. Christoff et al. (2009)), therefore negatively impacting user engagement. As a consequence, subjective questionnaires do not provide online and regular estimates of mental workload and engagement, but can be used as a complement other measurement methods (see for instance Hirshfield et al. (2011), Solovey et al. (2011)).

2.3.2 Objective measures : behavioral approach

An alternative way to assess user's mental workload and engagement is to directly measure his/her behavior. As described in section 1.3, provided that mental workload and engagement are related to the behavior (by explaining phenomenons such as attentional tunneling, fixation errors, disengagement or mind wandering), the technique would consist in inverting this relationship, i.e. deducing the operator's engagement and mental workload from the behavior.

The most straightforward behavioral measurement of the mental workload and engagement of an operator executing a given task is to assess the **performance** of the individual at this task (e.g. Cummings and Mitchell (2008)). Although simple in appearance, quantifying the level of performance of an individual is not always possible, especially in **data-limited** tasks (Norman and Bobrow 1975), i.e. tasks that do not require maximal engagement to reach maximum performance (For instance, driving a car can be performed equally well without fully engaging into the action of driving). Similarly, the interpretation of changes in performance can be hazardous, as a decrease in performance can be observed both in low and high workload conditions (see part 1.2).

Some studies also assess the mental workload and engagement experienced by an individual performing a task by measuring his/her performance at a **secondary task** (e.g. Brünken et al. (2004), Kaber et al. (2007)). The performance at a secondary task can give an estimation of the cognitive resources left available by the primary task. For instance, Ayaz et al. (2012b) could calibrate a model of mental workload experienced by an operator performing an Air Traffic Control (ATC) task by measuring the performance of the operator at a secondary memory task (N-back task) performed separately. The main limitation to this approach is the confound caused by user engagement in both primary and secondary tasks, leading to poorer sensitivity and diagnosticity : for instance, if the operator decides to engage himself/herself more in the secondary task, the interpretation of his/her performance as the reflect of the cognitive resources left free by the first task can be misleading. A good example of this fact occurs in case of the use of a *duration estimation task* as a secondary task (see Block et al. (2010)) : when asked to estimate the duration of an auditory

signal while performing a task, an operator can use two different strategies. On one hand, he/she can use **prospective judgment**, by engaging some cognitive resources in the temporal estimation task. On the other hand, he/she can use **retrospective judgment**, by not engaging any resources in the temporal estimation task and estimating the duration of the sound afterwards. These two paradigms of temporal estimation are affected differently by mental workload. Thus, in this case, operator engagement significantly modifies the accuracy of the interpretation deduced from performance measurement.

2.3.3 Objective measures : Psycho-physiological approach

2.3.3.1 Eye tracking measures

Measuring eye movements can give precious clues on his/her state and behavior while performing a task, as the visual modality is the main modality of attention, and to this end visual information represents one of the multiple resources the tasks are competing for (see section 1.1.2). Eye tracking measures rely on the estimation of eye movements, blinking events and pupil diameter which can be interpreted as correlates of CNS/ANS activity. Modern eye tracking devices, based on the technique of the corneal reflect, allow efficient measurement with very low intrusiveness. Finally, head-mounted devices allow the implementation of eye tracking measurements in real-world settings for details on the eye tracking technique). Poole and Ball (2006) proposed a review of the usability of eye tracking metrics to evaluate the interaction of an operator with a system, giving precious clues about his/her engagement and mental workload.

First, **Eye movements** are sensible to both the level of mental workload experienced by the operator and his/her engagement in the task. Indeed, the study of **fixations** (fixed position of the eye for a longer period of time, traducing the processing of visual information) and **saccades** (eye movements between two fixations, predominant in case of visual search) associated with the **areas of interest** (AOI) the operator looks at allows the distinction between phases of engagement in visual search or visual processing (Goldberg and Kotval 1999) and can be used to detect extreme engagement states such as attentional tunneling (Regis et al. 2014) or mind wandering (Uzzaman and Joordens 2011, Grandchamp et al. 2014). Intriguingly, if eye movements have been shown to be good indicators of engagement, their analysis can be misleading when it comes to mental workload sensitivity. Indeed, if both saccades and fixations are affected by variations of mental workload (Tsai et al. 2007), the way they are affected varies with the task being performed. For instance, depending on the task being performed and the engagement of the operator in it, an increase of mental workload can elicit longer fixations (e.g. de Greef et al. (2009), Ahlstrom and Friedman-Berg (2006)) or shorter ones (e.g. Schulz et al. (2011)).

Furthermore, **pupil diameter** can be used as a correlate of ANS activity, therefore wit-

nessing the effects of mental workload on it. Mental workload increases are associated with mydriasis, i.e. an augmentation of pupil diameter (de Greef et al. 2009). However, the influence of luminosity on pupil diameter usually confuses the interpretation of this metric, requiring extra processing to maintain good specificity of the estimation of workload (Peysakhovich et al. 2015).
Finally, increases in mental workload and engagement have been linked to decreases in **eye blinks** duration (Tsai et al. 2007, Ahlstrom and Friedman-Berg 2006), expanding the possibilities offered by eye tracking. However, if this method appears to be accurate to observe when extreme attentional phenomenons occur *a posteriori*, the precursors of these phenomenons are *a priori* inaccessible to eye tracking measurements.

2.3.3.2 Electrocardiography (ECG)

The close link between the heart and the brain of an individual is exhibited in all models of neurovisceral integration (Thayer and Lane 2009). Indeed, the heart influences the brain, as it is responsible for its energy supply (and therefore can be seen as one of the resources underlying task performance) ; and reciprocally, the brain modulates the heart activity through ANS. This activity can be derived from ECG measure, captured by electrodes placed at skin surface. Due to its simplicity of implementation and low cost, it has been used in a variety of situations from laboratory to real-world settings. If the easiest way to interpret ECG signal is to associate increases in **heart rate** (HR) with increases in mental workload (Haapalainen et al. 2010), power spectral analysis usually allows deeper interpretation less influenced by physical activity (Raya and Sison 2002). Indeed, Pomeranz et al. (1985) and Kamath and Fallen (1992) associated low frequency ($LF : 0.04 - 0.12Hz$) variability of heart rate with blood pressure control, i.e. sympathetic activity; and high frequency ($HF : 0.224 - 0.28Hz$) variations with respiratory sinus arrhythmia, i.e. parasympathetic activity. On this basis, the study of LF/HF ratio of **heart rate variability** has been proved to be sensitive to mental workload variations (Hjortskov et al. 2004, Murai et al. 2008, Miyake 2001).

2.3.4 Objective measures : brain measurement techniques

Behavioral and psycho-physiological measures usually lack accuracy and clarity of interpretation to account for functional failures. Indeed, if they usually provide a good observation of attentional failures (e.g. mind wandering, attentional tunneling or perseveration), the research for precursors of such phenomenons requires a deeper insight in the individual, by using cortical measures.

2.3.4.1 Electroencephalography (EEG)

Principle

Electroencephalography (EEG) is the oldest known method of brain activity measurement, as the first measure on human beings was done nearly a century ago (Berger 1929). As explained by Hämäläinen et al. (1993), during neuron activity, the **action potentials** transmission over the body of the neuron (from the dendrites to the axon) corresponds to an electromagnetic dipolar activity. More generally, when a large neural population is synchronized, the resulting electromagnetic dipole (equivalent to the sum of the dipoles representing each neuron involved) results in the generation of electric and magnetic fields measurable by placing electrodes on the scalp. However, the arrangement of neurons is primordial to be able to measure the resulting potential. Indeed, all the neurons have to be oriented in the same direction — what we call an *open-field structure*— to produce a large enough electromagnetic field. On the contrary, *closed-field structures* in which neurons are all oriented in different directions, especially present in subcortical areas, are not measurable using EEG. Another important point is that the electrodes placed on the scalp also measure environmental or muscular electric activity, which are usually much larger in amplitude than brain signals. These noises make EEG a hard technique to use and add complexity to data processing. Nevertheless the possibilities offered concerning the measurement of cortical neuronal activity and its cost have made EEG the most popular brain activity measurement device.

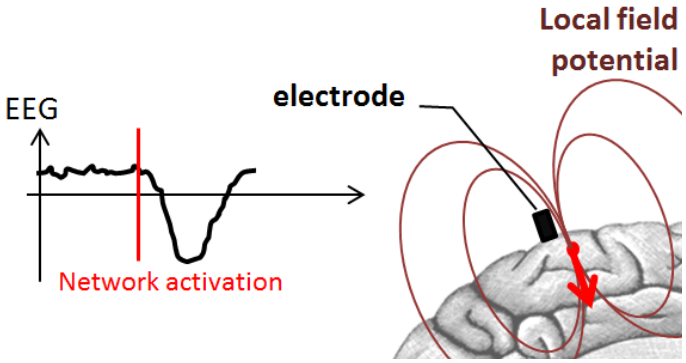


Figure 2.2 – *Principal of temporal EEG analysis. The activation of the neural network represented by the red dipole provokes a negative deflection in the local field potential measured at the electrode.*

First, the local synchronization and desynchronization of assemblies of neurons leading to variations in the EEG amplitude measured in a specific frequency band can be analyzed in terms of neuronal activity. Indeed, from an information theory point of view, desynchronized signals represent a maximum of information capacity (Thatcher et al. 1983). Similarly, desynchronized EEG corresponds to an activated level of cortical neurons (Steriade et al. 1990). Therefore, **Event-Related Desynchronization** (ERD) or **Synchronization** (ERS) can be interpreted in terms of an activated cortical area (Pfurtscheller

2001). This activity is non phase-locked to the event and highly frequency-band specific, which leads to the definition of different cortical rhythms (see Fig. 2.3).
By contrast, when presenting a stimulus to a subject, phase-locked reactions of the brain resulting in sequential activations of neuronal assemblies allow the observation of positive or negative deflections of the EEG signal (depending on the orientation of the dipole equivalent to the network), called **Event Related Potentials** (ERP) (Vaughan Jr and Arezzo 1988) (see Fig. 2.2). The observation of the location and shape of the ERP components gives precious clues on the processes underlying stimulus processing. For instance, the N100 component (negative deflection with a latency of 100 ms) has been linked stimulus perception in the primary and associative sensory cortices (Näätänen and Picton 1987) as well as selective attention (Hillyard et al. 1973) ; the P300 component (positive deflection with a latency of 300ms) has been related to activity of the frontal dopaminergic pathway in the context of early attention processes and of temporo-parietal noradrenergic pathways in the context of memory processing (Polich 2007).



Figure 2.3 – *Cortical rhythms frequency bands.*

Use in Neuroergonomics

EEG is commonly used as a technique of assessment of mental workload and user engagement, due to its relative portability compared to other cerebral measurement techniques and reliability. To this end, researchers use both techniques of analysis of EEG signal (Antonenko et al. 2010, Berka et al. 2007), i.e. on one hand the **event-related potentials** (ERP) observed consequently to neuron networks activations elicited by stimuli, and on the other hand the **cortical rhythms** resulting from local neurons synchrony.

First, mental workload modulates the amplitude of the ERP components elicited by the presentation of auditory stimuli (in addition to the performance of a task) (Allison and Polich 2008). In particular, the diminution of the N100 and P300 components amplitudes associated with increased mental workload has been used as a method of estimation of mental workload (Kramer et al. 1995). Furthermore, some studies suggest that these components can be used as detectors of user's attentional disengagement caused by extreme mental workload situations, such as inattentional deafness (Giraudet et al. 2015) or mind wandering (Smallwood et al. 2008). However, although it can give detailed information on the operator's state, event-related techniques for EEG analysis are hardly useable in ecological situations, due to the necessity to introduce additional auditory stimuli unre-

lated to the task, and also to the difficulty of ERP components' extraction from one single auditory trial (Berka et al. 2007). This latter difficulty causes ERP analysis to be used mainly for *a posteriori* characterization of mental workload and engagement.
Alternatively, the analysis of spectral power density (SPD) associated with cerebral rhythms' power bands is easier to implement in real-world situations (Wang 2010, Antonenko et al. 2010), although still largely affected by environmental noise. Mental workload has been shown to trigger increases in alpha and theta power bands (Gevins and Smith 2005, Berka et al. 2007), making those power bands correlates of mental workload. On the same basis, Prinzel et al. (2000) developed an estimator of user engagement (defined as his/her capacity to maintain sustained attention directed towards the task) in a task implemented using beta power divided by alpha plus theta power. This led to a clear distinction between estimates of mental workload (called *EEG-workload* metrics) and engagement (called *EEG-engagement* metrics) in the study of metrics based on EEG spectral analysis (Stevens et al. 2007). In particular, some vigilance studies have focused on the ability to detect by those techniques executive failures such as disengagement (Eoh et al. 2005, Braboszcz and Delorme 2011). Taken together with ERP components, those metrics resulted in the development of estimators of mental workload and engagement (Trejo et al. 2005).

2.3.4.2 Functional Near Infrared Spectroscopy (fNIRS)

Principle

The idea of the existence of an adaptive mechanism of **regulation of blood supply** to brain areas has been first presented more than a century ago (Roy and Sherrington 1890), and has been confirmed later on by many studies (Friedland and Iadecola 1991). In fact, through a mechanism called **neurovascular coupling** (Pasley and Freeman 2008), neuronal activation implies a local augmentation of Cerebral blood flow (CBF) together with a null to moderate augmentation of Cerebral Metabolic Rate of Oxygen ($CMRO_2$) (Leon-Carrion and Leon-Dominguez 2012). The main hypothesis to explain this mechanism requires to look a bit closer at the neuronal metabolism. Indeed, although aerobic metabolism represent the major part of ATP production in the human organism, this metabolism is nearly maximally used by the neurons at rest to maintain electrochemical gradients (Sokoloff 1993). Consequently, the additional ATP production required by the neuron when it is activated requires the transitory use of anaerobic metabolism to meet increasing energy demands (Fox et al. 1988), thus explaining the moderate increase in $CMRO_2$. Simultaneously, local neuronal activity stimulates astrocytes metabolism(Pellerin and Magistretti 1994), and their central role in neurovascular coupling via endfeet on capillaries induces local vasodilation, therefore explaining the CBF increase (Yu et al. 2010). Even if the actual

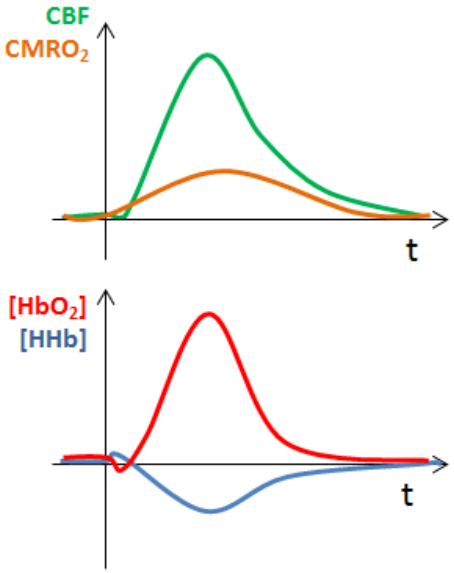


Figure 2.4 – *Illustration of the* ***hemodynamic response*** *following local neuronal activity. Neuronal activity induces a moderate increase of the cerebral metabolic rate of oxygen ($CMRO_2$), which triggers local vasodilation through neurovascular coupling mechanism, locally leading to higher cerebral blood flow (CBF). This mechanisms corresponds at the blood oxygenation level to an increase in oxygenated hemoglobin concentrations (HbO_2) and a decrease in deoxygenated hemoglobin (HHb).*

neurovascular coupling seems to imply other mechanisms, as some studies observed that neuronal activity can forwardly stimulate vasodilation (Attwell and Iadecola 2002), the gap between CBF and $CMRO_2$ increases is commonly regarded as the motor of hemodynamic response.
Locally, the changes in CBF and $CMRO_2$ following neuronal activation result in a local increase in **oxygenated hemoglobin** (HbO_2) concentration, as more blood is supplied to the area, and a local decrease in **deoxygenated hemoglobin** (HHb) concentration, as the augmentation of cerebral blood volume (CBV) with none or little increase in oxygen consumption results in greater dilution of HHb (see Fig. 2.4).

The observation of the concentration changes is the cornerstone of functional magnetic resonance imaging (fMRI), as HbO_2 and HHb have different magnetic susceptibility (Ogawa et al. 1993), and fMRI uses this to provide a measure of their concentration. On the same basis, functional near infrared spectroscopy (fNIRS) uses the property that HbO_2 and HHb have different optical characteristics to provide the same measure. Indeed, HbO_2 and HHb have different light absorbance depending on the wavelength, so that the near infrared light transmittance of the brain measured by placing a light emitter and a light sensor (the combination of which forms an **optode**) can be directly linked to the local concentrations of these chromophores by using a modified Beer-Lambert law (Villringer and Obrig 2002) (see Fig. 2.5). The simplicity of the method makes it a very good candidate for neuroergonomics, however two methodological limitations remain : first, the sensor positioning has to be done carefully to ensure good contact and emitter-detector distance, as it impacts the optical pathlength (i.e. the path followed by the light when interacting with tissues) (Van der Zee et al. 1990) ; then, the very high melanin absorbance makes fNIRS harder to use in subjects with dark skin pigmentation or in hairy areas of the scalp (Wassenaar and Van den Brand 2005).

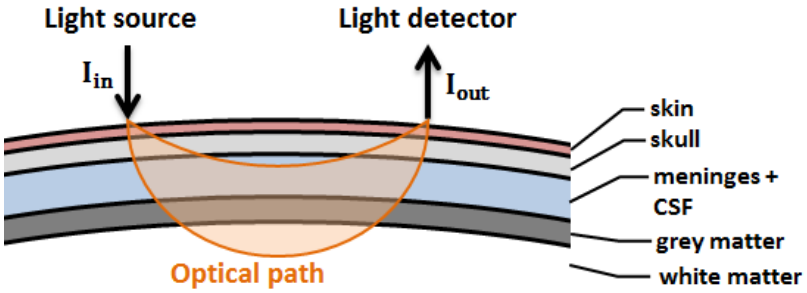


Figure 2.5 – *Propagation of light in the human tissues and NIRS functioning. The light travels between the source and the detector through a banana-shaped optical path, due to its interaction with tissues. The Beer-Lambert law states that for a given wavelength, the transmittance of the area corssed by the light (defined by $log(I_{in}/I_{out})$) is proportional to the concentrations of HbO_2 and HHb in the optical path.*

Use in Neuroergonomics

Functional near infrared spectroscopy (fNIRS) has gained momentum over the last two decades thanks to its promising results. Indeed, the principle of this technique makes it a cost-efficient, portable and robust measure, usually less affected by external noise and motion artifacts than EEG (Derosière et al. 2013). This device therefore seems to be adapted for real-world situations (Perrey 2008). Nevertheless, the neurovascular coupling mechanism linking brain areas recruitment and hemodynamics is subject to high variability between subjects (Jasdzewski et al. 2003), and is also a slow mechanism (Friston et al. 1995), leading to poorer transferability and temporal resolution.

In a variety of studies in neuroergonomics ranging from laboratory task to realistic settings (e.g. (Mandrick et al. 2013, Takeuchi 2000, Harrison et al. 2013)) the variations of the concentrations of **oxygenated hemoglobin** (HbO_2) and **deoxygenated hemoglobin** (HHb) have been associated to changes in mental workload and operator engagement. In fact, the recruitment of cortical executive regions associated with higher mental workload and engagement results in higher HbO_2 and lower HHb concentrations (Ayaz et al. 2012b). In addition to the good sensitivity, one of the main advantages of fNIRS is the high functional precision offered by its spatial resolution. Indeed, the recruitment of executive regions of the PFC is measurable using this technique (e.g. working memory and DLPFC (Ayaz et al. 2007), workload during ATC supervision and recruitment of the frontopolar cortex (Ayaz et al. 2011), left hemisphere recruitment during landing (Takeuchi 2000)), offering a greater precision concerning the type of resources mobilized for a task, i.e. a good diagnosticity.

Although neuroergonomics research in fNIRS mainly concerns the study of the PFC, because of the presence of hair on the other parts of the scalp, the predominant role of this brain area in executive control (as seen in chapter 7.1) make fNIRS a good candidate

to study mental workload and operator engagement (Derosière et al. 2013). Concerning the detection of precursors of executive failures, fNIRS has also shown potential to detect both disengagement linked to overload (Izzetoglu et al. 2007, Durantin et al. 2014a) and attention lapses (Harrivel et al. 2013, Durantin et al. 2015), and its functional resolution also offer correlates of various modes of resource allocation linked to internal parameters such as expertise (Bunce et al. 2011).

2.3.5 Summary of the different measures

Table 2.1 summarizes the measures presented in this section, using the criterias proposed by Wierwille and Eggemeier (1993) (see also Kramer (1991) for a review). In particular, **specificity** regroups the criteria of diagnosticity and global sensitivity to account for the ability of one measure to inform about the type of resources mobilized and the causes of this mobilization. The only criterion not represented in the table is the *transferability* criterion, as transferability techniques still represent a major challenge for signal processing, no matter what type of measure is used (Pan and Yang 2010, Power et al. 2012).

As shown in the table, cortical measurements such as fNIRS or EEG seem to be the most promising ones, as they exhibit both diagnosticity concerning the type of resources implied (and sometimes of the modes of allocation of those resources), and sensitivity to mental workload and user engagement. However, the use of multimodal approach can be used in order to give maximum information on the operator's state.

2.4 Conclusions on Neuroergonomics

In this chapter, we introduced the concept of Neuroergonomics, and described how this field can benefit from the use of brain imaging techniques (such as fNIRS or EEG) and the study of human behavior in ecological situations. However, measuring ANS and CNS activities to evaluate mental workload and engagement is not sufficient for a neuroergonomics approach if the estimates are not used to facilitate human-machine interaction. **Adaptive automation** (Sheridan 2011, Kaber et al. 2007) and **cognitive counter-measures** (Dehais et al. 2011), which consist in inferring the human operator's cognitive state from different measurement techniques (Wilson and Russell 2007) and then adapt the nature of the interaction or give feedback to the operator to avoid cognitive bottlenecks (St. John et al. 2004), are key concepts of the neuroergonomics approach. To this end, Neuroergonomics proposes the use of **Brain Computer Interfaces** (BCI) as a tool to build adaptive systems (Parasuraman 2003). However, one must be able to use all the measurement tech-

	Sensitivity	Specificity	Intrusiveness	Implementation Requirements
Subjective measures			Require to interrupt the operator	No regular and online measures
Performance	Maximum performance can be reached without maximum resource allocation	Misleading due to effects of both high and low workload		Sometimes hard to evaluate
Performance (secondary task)	Depends on subject engagement in both tasks		The secondary task can affect the primary one	Forces the use of multiple tasks paradigms
Eye Tracking		Do not account for the causes of resource allocation		
ECG		Do not account for the causes of resource allocation		Easily affected by electromagnetic noise
EEG		The types of resources involved are hard to distinguish online (requires ERP analysis or ICA)		Some components are hard to evaluate online (ERP), signal quality usually bad in real world settings
fNIRS				Lower temporal resolution

Table 2.1 – ***Summary of measures characteristics for the assessment of mental workload and operator engagement in Neuroergonomics.*** *(see details in the text of each section)*

niques that were described in this chapter in real-time and in real environment, therefore raising new challenges for system design.

Brain Computer Interfaces 3

- *BCIs use brain inputs to restore/replace (active BCI) or improve (passive BCI) a function.*
- *The key steps of BCI design include signal processing and control law design.*
- *Although most current BCIs use EEG, fNIRS has the potential to improve usability of BCI.*
- *Technological challenges still remain in BCI design and require innovative signal processing techniques.*

In the previous chapters, we insisted on the importance to measure the **mental workload** and **engagement** of an operator in order to interpret his/her performance. In particular, the approach used in **Neuroergonomics** aims at using the knowledge from cognitive neuroscience to improve human-machine interaction. **Brain Computer Interfaces** (BCI), which will be described in this chapter are one of the main tools to achieve this goal. After the presentation of their main objectives, we will focus on the key steps impacting BCI performance, and describe the current state of the art in BCI design using fNIRS and EEG. Finally, this chapter will explain the remaining challenges to improve the usability of BCI.

3.1 Use and objectives of Brain Computer Interfaces

A comprehensive way to introduce the systems comprising a BCI is to compare them with "classical" controlled systems. Fig. 3.1 shows the usual representation of controlled systems (e.g. Franklin et al. (2006) for theory), receiving a order (*reference*) and producing an *outcome*. The structure of such systems can be divided into three main parts :

- The **controller** transforms the order (usually abstract) into an actual command (with a physical meaning : e.g. current intensity, mechanical effort) which is sent to the system.
- The **plant** transforms the command into an outcome. Together with the controller, they form the forward loop.
- The **sensors** are used in the feedback loop to estimate the outcome of the system. This measure is then fed back in the forward loop to modify the order given to the controller (in order to adapt the strategy to reach the reference).

For instance, if the controlled system is an oven, then the order given (reference) would be the temperature to be reached. The function of the system relies on the *heating element* (plant), which is supposed to transform a current (command) into heat (outcome). In this example, the sensor used in the feedback loop would be the *thermostat*, used to check if the temperature inside the oven has reached the target. The difference between the measured temperature and the target one (reference) can then be used as a new command for the system. Finally, the controller would be a *logical block* transforming the difference between the two temperatures into a command for the heating element (turn on the heating element if the temperature is below the target, turn it off if the temperature is above).

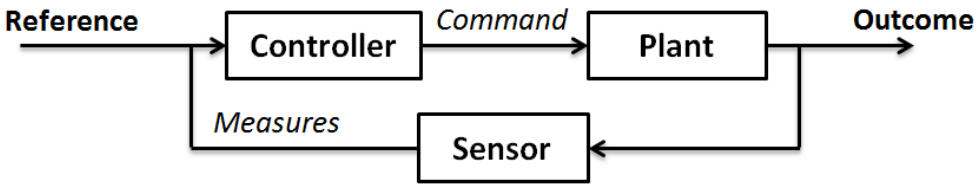


Figure 3.1 – *Canonical form of controlled systems, dividing the different components in three groups : controller, plant and sensor.*

The role of the human in such **automated systems** is just to provide them with the reference, i.e. the order to apply (e.g. choose the temperature of the oven, in the example above). The three structural blocks defining the system (controller, plant, sensor) are independent of the human during normal functioning mode, allowing fully automated functioning. In an **interactive system**, the operator would be part of the controller block (i.e. the action of the system would not be fully automated and would require human intervention) and/or the sensor block (i.e. the operator would constantly evaluate the state of the system and modify the order it gives to it in order to reach his/her goal). However, such systems usually require a physical input from the operator (e.g. manual input, or speech). The fundamental idea of **BCI** is to replace those physical inputs by brain signals, that can be intentionally or unintentionally produced, leading to a distinction between **active** and **passive** BCI (Zander et al. 2010).

3.1.1 Active BCI

In the absence of medical methods to repair the damage done by some disorders (e.g. amyotrophic lateral sclerosis, brainstem stroke, spinal cord injury) on the neural pathways through which the brain interacts with its external environment (e.g. neuromuscular channels), researchers have tried to figure out ways to **restore** the functions lost consequently to those disorders (Wolpaw et al. 2002). The first development of BCI aimed at restoring such functions to disabled persons (e.g. Travis et al. (1975), Kuhlman (1978), Birbaumer et al. (1981)). Their objective was to *actively* control interfaces using brain inputs (e.g. move a cursor on a screen), therefore they were named **active BCI**.

In a traditional controlled system (see Fig. 3.1), the control is realized by the feedback loop, which allows the modification of the strategy to reach the goal. The idea of **active BCI** is to put the human inside the feedback loop, in order for him to control the system. As shown on Fig. 3.2, an external loop including the human is added to design an active BCI. The role of the human is to **observe and take a decision** concerning the system. For instance, if the goal of the active BCI is to control a cursor on a screen, the human has to observe the current position of the cursor and decide where he/she wants to move it. The decision taken by the human is then estimated using **brain measurement techniques**, and transformed into a command for the system.

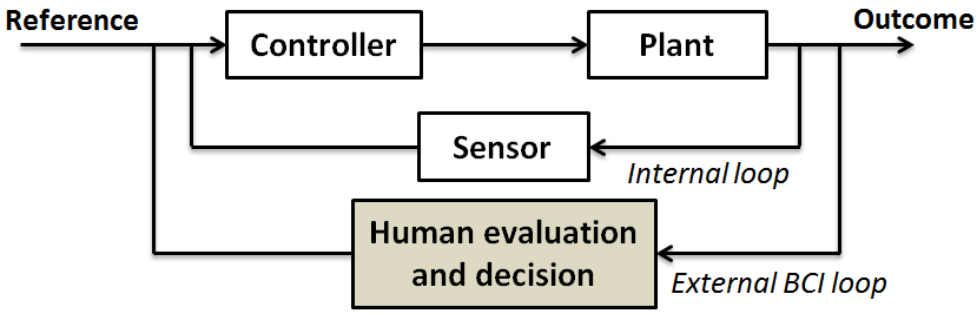


Figure 3.2 – *Structure of an active BCI. An external control loop is added to allow the operator to actively observe and take decisions concerning the system (e.g. modify the goal of the system). The decisions of the operator are then evaluated by brain measures and fed back into the system.*

If BCI were first considered for their therapeutic possibilities (to **restore** a function), the idea of using them to **replace** a function, i.e. to control a device with one's mind, rapidly emerged in science-fiction fantasy (e.g. Thomas (1977), Venditti and Weldele (2012)) and in research (see Wolpaw et al. (2002) and Beverina et al. (2003) for a review). Several attempts at designing mind-controlled devices (e.g. LaFleur et al. (2013), Kim et al. (2014)) or video games (e.g. Lalor et al. (2005)) for healthy users emerged, however two major limitations remain. Indeed, on the first hand, as the system needs regular inputs, such applications would be possible only using rapid neurophysiological dynamics such as EEG (Coyle

et al. 2004). On the second hand, a greater problem concerns the additional workload generated in mind-controlled devices compared to manually-controlled devices (Brouwer et al. 2013). In fact, for healthy people, controlling devices with intentionally generated brain activity rather than manually has been associated with greater control difficulties and loss of presence (Groenegress et al. 2010).

3.1.2 Passive BCI

In line with Neuroergonomics, an alternative approach in BCI design is the idea of using BCI to **enhance** the interaction between human and machines and not to replace it (Parasuraman 2003, George and Lécuyer 2010). This kind of systems views brain activity monitoring as a mean of estimation of the operator's state (**engagement** in the task **mental workload** experienced) which can be used to modify the system with a **passive** input from the operator. They are called **passive BCI** (Cutrell and Tan 2008).

As shown on Fig.3.3, human is the controller of such systems, i.e. he/she effectively operates the system to reach his/her own goals (for instance, driving a car or piloting a plane relies on such human-machine interactions). The role of **brain measurement techniques** in these systems is to estimate the **state** of the operator (e.g. mental workload, engagement), which can be used either to give **feedback** to the operator or to the system.

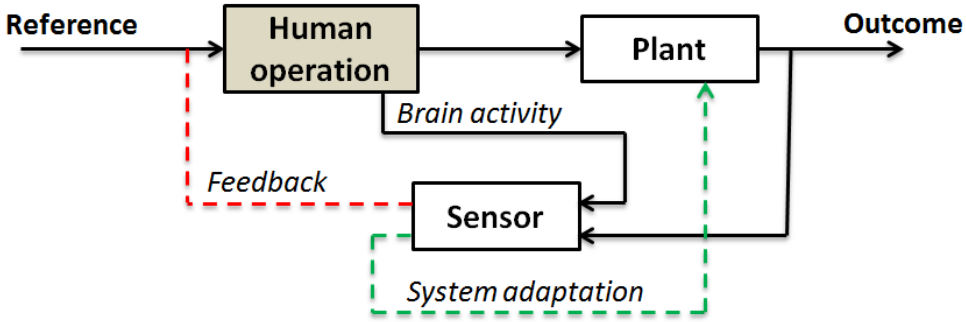


Figure 3.3 – *Structure of a passive BCI. Human plays the role of the controller, and sensors include estimation of the state of both the system and the operator (using brain measurement techniques). The information collected by the sensors are used to give feedback to the operator (in red) and/or to adapt the functioning of the system (in green).*

If the action of state estimation is directed towards the operator, one of the potential uses is **Neurofeedback**. By giving the operator feedback on his/her mental state, such systems enhance operator performance (Gruzelier et al. 2006, Egner and Gruzelier 2003) and can be used as therapeutic methods for psychiatric disorders (Fuchs et al. 2003, Dohrmann et al. 2007, Halas and Pelaez 2015). Another way of implementing such operator feedback is the use of **cognitive countermeasures**, defined as means to mitigate cognitive bias by modifying perceptual information presentation (Dehais et al. 2003). Knowledge concern-

ing the operator state by the use of passive BCI could help provide such countermeasures in a timely and accurate manner (de Visser and Parasuraman 2011).

If the action of state estimation is directed towards the system itself, the main lever for improved performance is adaptive automation (Taylor et al. 2010). The management of the level of automation permitted by the knowledge of the operator state allows the management of the level of workload and engagement of the user, therefore improving his/her performance and avoiding executive failures (e.g. attentional tunneling, disengagement) (Kohlmorgen et al. 2007). In general, such adaptive systems are favored compared to active BCI because the brain signals used as an input are not intentionally produced, and therefore passive BCI do not require additional effort from the operator (Zander and Kothe 2011, Brouwer et al. 2013)

3.2 BCI design

With the gain of momentum of BCI over the last decades due to their applications in rehabilitation, reeducation, safety and performance enhancement, several attempts have been made to formalize the concepts underlying BCI design. In particular, several software platforms have been designed to provide the necessary tools for BCI development, e.g. *BCI2000* (Schalk et al. 2004), *OpenVibe* (Renard et al. 2010) or *BCILAB* (Kothe and Makeig 2013). Although most of these platforms were originally designed for EEG-based BCI, a growing body of research has investigated the potential of fNIRS to replace or complement EEG in BCI systems, in particular due to its spatial resolution and sensitivity to idle modes (see Naseer and Hong (2015b) for a review).

Mason et al. (2007b) proposed a decomposition of the BCI design process into several steps, summarized on Fig 3.4. First, the **signal processing** step comprises all the processes applied to raw signals to clear them from undesired signals, leaving only the relevant brain activity (e.g. correlates of mental workload or user engagement). Then, the **command law design** consists in the definition of the appropriate feedback for the system using the brain signals (definition of a control signal in case of an active BCI, definition of cognitive countermeasures, adaptive automation measures or neurofeedback in case of passive BCI).

3.2.1 Signal processing

The signal processing step is usually considered as the most critical in BCI design (McFarland et al. 2006), as the quality of the removal of all undesired components of the

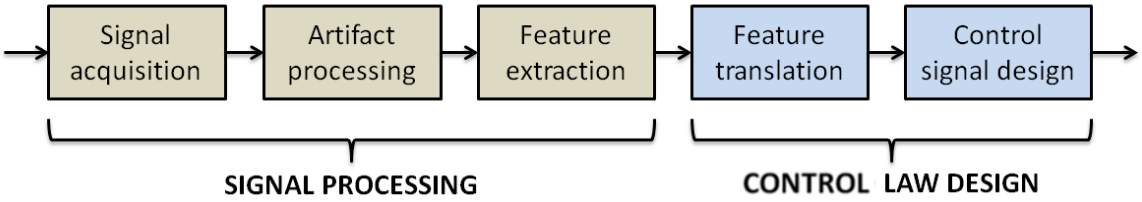


Figure 3.4 – *Different steps in the BCI design process, adapted from Mason et al. (2007b). First, the **signal processing** step includes raw signal acquisition, and its processing to extract only the desired signal. Then, the **command law design** step consists in the interpretation of the extracted signal to produce a control signal (which will be fed back to the operator or the system)*

raw signal without removing any part of the desired brain signal affects critically the BCI performance (see Fig.3.5 for a schematic). As discussed by Bashashati et al. (2007), four crucial steps influence the quality of the features usable for BCI, and represent the greatest challenges for BCI design :

- The quality of the brain measurement device (hardware) and of measurement procedures.
- The efficiency of the artifact removal techniques.
- The potential signal improvement techniques applied to the cerebral signal.
- The accuracy of the features extracted from the signal.

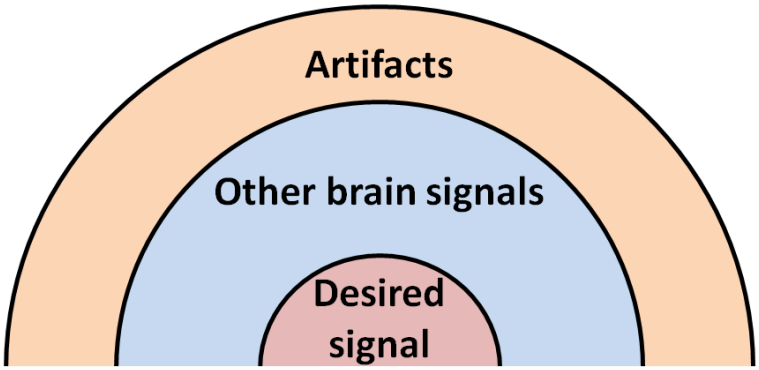


Figure 3.5 – *The signal processing step consists in "digging out" the desired brain signal, first by clearing brain signal from artifacts (coming from external sources), and then by extracting the relevant features from the brain activity.*

3.2.1.1 Brain measurement procedures to improve signal usability

BCI design always has to compromise between the usability of a measure and its precision. Some measurement techniques, very invasive and non portable, such as fMRI, have

quickly been left aside in BCI research, and most BCI research over the last decades has been conducted using EEG devices, thanks to their portability. There have been several attempts to make the EEG technique more portable and more convenient to use in real-world situations (for instance with the introduction of dry electrodes devices, e.g. Popescu et al. (2007), Grozea et al. (2011)). The progressive use of fNIRS to design BCI has been also introduced as an alternative to EEG, thanks to its easiness of use and good functional resolution (Sitaram et al. 2007) (although its low temporal resolution makes it a better candidate for passive BCI). In particular it is believed that fNIRs would be good candidate for idle mode detection, i.e. to detect instants when the operator does not request any command (Tomita et al. 2014). Most recent BCI attempts obtain the best performance by hybridizing the two systems (e.g. Pfurtscheller et al. (2010a), Fazli et al. (2012a)).

Nevertheless, brain measurement procedures should already bear in mind the need to avoid artifacts, i.e. the part of the signal not reflecting brain activity (Bashashati et al. 2007). Indeed, most environmental artifacts (e.g. electromagnetic noise, coupling loss between the device and the scalp) can be avoided by preliminary considerations (adaptation of the environment to avoid electromagnetic noise, enhancement of the device structure...), whereas some internal artifacts (e.g. muscular activity) can be avoided by warning the operator.

3.2.1.2 Artifact removal techniques

Artifacts represent the part of the data which does not correspond to brain activity. They can have an internal origin (i.e. come from the subject him/herself) or an external one (see table 3.1 for a summary). Concerning **EEG**, external artifacts include electromagnetic line noise and electrodes movements ; internal artifacts include muscle activity, eye movements and blinks and heart rate (Iriarte et al. 2003). Concerning **fNIRS**, external artifacts include ambient light intensity, sensor movements and linear trends (drift) in the signal (Jang et al. 2009); internal artifacts include vasomotion, heart rate and motion artifacts (Mandrick et al. 2013, Derosière et al. 2013).

Stationary artifacts, i.e. artifacts whose dynamical properties do not change over time, are usually removed using time-series analysis and bandpass linear filtering (Huppert et al. 2009). For both EEG and fNIRS, most studies focus on linear numeric filter with **finite impulse response** (FIR) for most offline analysis. Indeed, FIR filters are always *stable* and with *linear phase*, two critical properties in signal processing, but usually require higher orders (i.e. more data) to perform well. This latter characteristics does not make them a good candidate for real-time implementation in BCI, as high order filters would induce delays in signal processing. On the contrary, most online studies in EEG and fNIRS (for BCI applications) use **infinite impulse response** filters because of their high

	Stationary	Non Stationary
EEG	heart rate, regular muscular activity, eye movements and blinks, line noise, electrode displacement	large muscle punctual activity, electromagnetic interference
fNIRS	vasomotion, heart rate, drift, ambient light, sensor movements	Large motion artifacts

Table 3.1 – *Stationary and Non Stationary artifacts in EEG and fNIRS signals*

performance with low order, even though their stability and phase linearity have to be checked.

Non stationary artifacts, whose dynamical properties change over time, represent one of the most challenging parts of BCI design (van Erp et al. 2012). Indeed, the impossibility to design a unique and stationnary filter to remove those artifacts, often coupled with the non-linearity of such artifacts, requires the use of adaptive filtering (e.g. Zhang et al. (2009) for fNIRS and Correa et al. (2007) for EEG) or estimation methods (e.g. Kalman filtering (Izzetoglu et al. 2010)). Some algorithms were also developed to remove such artifacts using subspace techniques (i.e. trying to find stationary and/or linear subspaces in the data), and they represent a growing body of research in the field of BCI (e.g. von Bünau et al. (2009), Teixeira et al. (2008)) although they usually require calibration.

3.2.1.3 Signal improvement techniques

Once the signal has been cleaned from its artifacts, only brain activity should remain (on Fig.3.5, the first layer has been removed). However, the extraction of the desired brain activity for BCI design can be enhanced. As a comparison, we can view the artifact processing step as a mean to decrease the noise in Signal/Noise ratio (S/N), but this ratio can still be improved by increasing the desired signal predominance. Such signal improvement techniques usually rely on **a priori knowledge** on brain dynamics by using **correlation or convolution** techniques, e.g. the hemodynamics response function for fNIRS (Cui et al. 2010, Kamran et al. 2012) or the event related potentials for EEG (Quiroga and Garcia 2003). Online improvement has also been tested using **estimators** relying on modeled dynamics such as Kalman filter (e.g. Georgiadis et al. (2005a) for EEG and Abdelnour and Huppert (2009), Hu et al. (2010) for fNIRS, although these methods still need improvment concerning the estimation parameters (Aqil et al. 2012).

3.2.1.4 Feature selection

Feature extraction from brain data depends both on the objectives of the BCI and the type of classification algorithm to be used (see 3.2.2). The **features** represent the portion of data (or a mathematical quantity describing the data, e.g. mean, variance...) from which the control signal will be extracted. EEG BCI, representing the main part of current BCI research, are usually divided into four types, depending on the type of features they use (Amiri et al. 2013):

- **Event-related desynchronization or synchronization** (ERD/ERS) BCI (e.g. Pfurtscheller and Da Silva (1999)), which use the cortical rhythms resulting from local neuronal synchrony or asynchrony. Therefore, the main feature used for this type of BCI is the power spectral density (PSD) of EEG data in one specific band (e.g. *alpha rhythm* for meditation, *mu rhythm* for motor imagery).
- **Steady-State Visually evoked potentials** (SSVEP) BCI, based on the frequency training of the visual areas resulting from the presentation of a visual stimulus at a given frequency (e.g. Vidal (1973)). Similarly to ERD/ERS BCI, the feature used here is the PSD of EEG data at in a given frequency range (corresponding to the stimulus presentation frequency).
- **P300 ERP component** BCI, based on the elicitation of P300 when attending to a stimulus. As the goal of such BCI is to detect the presence or absence of a component, all the data points following the arrival of the stimulus are commonly used as features (e.g. Jin et al. (2012)).
- **Slow cortical potentials** BCI, based on the elicitation of slower ERP (e.g. Allison et al. (2012)). If the principle is akin to P300 BCI, and therefore usually requires to use all data points as features, it has been shown that the number of features required for calibration is significantly higher for this type of BCI as the amplitude of the ERP components on which they rely is lower than P300's. (Amiri et al. 2013)

Concerning fNIRS, similarly to EEG frequency components (ERD/ERS or SSVEP), most current BCI research focuses on the level of hemodynamic response, and therefore on the **amplitude of the** HbO_2 **or** HHb **changes** (e.g. Coyle et al. (2007), Naito et al. (2007)), therefore some BCI use only their mean level of concentration as a feature (e.g. Nagaoka et al. (2010)). However, some researchers have focused more precisely on the shape of the hemodynamic response, using all data points following a given stimulus as features (e.g. Sitaram et al. (2007)), or alternatively all the **statistical moments** describing the hemodynamics response shape : mean, variance, skewness, kurtosis (e.g. Tai and Chau (2009), Gateau et al. (2015)).

3.2.2 Control law design

Once the features containing the relevant information have been extracted from the signal, one has to mine useful information from them and then convert them into a feedback for the system or the operator. This step is similar to the **control law design** step in physical systems design, and consists of two phases : feature classification and feedback.

3.2.2.1 Feature classification

An optional step to be able to give feedback in a timely and appropriate manner is to associate the features extracted from brain data with cognitive states or functions (e.g. level of mental workload or engagement, state of attentional tunneling, etc), if those features can not be exploited directly to give the feedback (see next section). This association between a feature vector and a mental state is done by **classification techniques** (see Müller et al. (2004) and Lotte et al. (2007) for a review), such as for instance linear discriminant analysis (LDA) and support vector machines (SVM) (Cortes and Vapnik 1995). These methods rely on two phases, schematized on Fig. 3.6 :

- A **training phase**, used for the mathematical estimation of the frontier separating data into classes. Also called **calibration**, this phase consists in the acquisition of data whose labels (i.e. the cognitive function or state to categorize) are already known, in order to compute a separation model from them (some methods do not require a priori knowledge on the labels : see *unsupervised learning* in the next paragraph).
- A **testing phase** consisting in the application of the previously computed model to new data to predict their label. A correctly classified trial is a **hit**, otherwise it is called a **miss**. The total amount of hits defines the **accuracy** of the classifier.

Reducing the complexity and time of calibration is one of the main challenges to BCI design (Lotte et al. 2007). To this end, one of the potential solutions in to use **unsupervised learning techniques**, which do not separate the data according to their label but according to the differences among their components, therefore not requiring to know the labels of calibration data beforehand. These techniques allow greater flexibility and adaptability of the BCI (Gan 2006, Lotte and Guan 2010).

The main limitation to this approach is the very **high dimensionality** observed in brain signals (Rakotomamonjy et al. 2005). Indeed, the number of data samples needed for the training phase increases significantly as the size of the feature vector increases. In

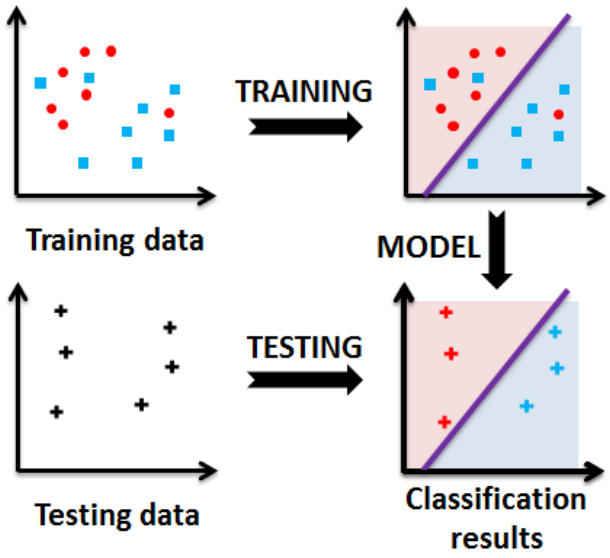


Figure 3.6 – *Example of classification process for linear separation (LDA) of two classes (blue and red). The classification process is :*
***Training** : Use training data, whose labels (blue or red) are known, to mathematically define the separation model.*
***Testing** : Apply the model to non classified data (testing data) to predict their label.*

practice, it is advised that the number of data samples for the training phase should be five to ten times the dimension of the feature vector (Jain and Chandrasekaran 1982) to have a good calibration of the model. This requirement, known as the **curse of dimensionality** (Friedman 1997), is one of the critical limitations to BCI performance (Lotte et al. 2007), as it is often not possible to collect enough data within reasonable calibration time. However **dimensionality reduction techniques** which extract only the main components present in the data to reduce their dimensionality (e.g. principal components analysis, independent components analysis) can be applied in order to limit the influence of this fact (Müller et al. 2004).

3.2.2.2 Feedback

Giving feedback to the system or the operator is the last step to "close the loop" of the BCI, but requires the transformation of the information decoded from brain activity into a control signal for the system. On one hand, it is possible to **use directly brain features as a control signal**, especially in case of **active BCI**. For instance, BCI based on the evaluation of ERD/ERS (cortical rhythms) use directly the PSD of the EEG in a specific bandwidth as a control signal (e.g. *BrainBall* (Jackson and Mappus 2010)). Similarly, active BCI based on the level of hemodynamic response can use directly the signals collected from fNIRS (e.g. *Dolphin trainer* (Nishimura et al. 2010)). In this case, no classification step is required.

Alternatively, some active BCI rely on classifiers to threshold the brain signals, in order to produce a *on/off* signal (or any type of binary input) for the system (e.g. SSVEP-based control interfaces (Lalor et al. 2005, Kim et al. 2014)).

On the other hand, most **passive BCI** rely on classifiers to detect critical states when intervention on the system is needed. First, as presented in section 3.1.2, **adaptive systems** have been designed with such techniques to adapt the level of **difficulty** (Rani et al. 2005) or **automation** (Pope et al. 1995, Kohlmorgen et al. 2007) of the system operation procedure when the BCI input is classified as critical (too low or too high engagement or mental workload), in order to improve user performance and engagement. Then, **cognitive countermeasures**, i.e. modifications in the system's interface can be applied as a way to improve user's immersion (Bos et al. 2010, Girouard et al. 2013) or mitigate conflicts between the human and the system (Dehais et al. 2011). Finally, the unintentionally generated brain activity used by a passive BCI can also be used to improve the performance of an active BCI, for instance by correcting errors in the operator's commands (Ferrez and Millán 2008, Dal Seno et al. 2010) or monitoring the performance obtained from the active BCI (Liu et al. 2013).

3.3 Remaining challenges in BCI design

BCI (and in particular passive BCI) now appears as one of the most efficient way to improve user performance and engagement both in aeronautics (Thurlings et al. 2010) and space (Coffey et al. 2010). However, implementing BCI for real-world applications still requires progress in the efficiency of BCI (Moore 2003), not only concerning the facility of utilization of brain measurement devices in real-world settings. Wolpaw et al. (2002) and Popescu et al. (2008) identified three remaining challenges for BCI design : **Information Transfer Rate**, **Synchronicity** and **BCI illiteracy**.

First, due to the long processing time and quantity of data required by the BCI design (in particular for the feature classification phase), the quantity of information fed back into the system is highly limited, leading to bad **Information Transfer Rate** (ITR). Most BCI systems are relatively low bandwidth systems, with ITR usually ranging from 5 to 25 $bits/minute$ (Wolpaw et al. 2002). Even if better ITR become possible with some paradigms (SSVEP paradigms have reached up to $83 bits/minute$ ITR (Parini et al. 2009, Lin et al. 2015)), reaching good levels of ITR for all types of BCI still remains one of the greatest challenge (in particular for fNIRS-based BCI (Naseer and Hong 2015a)), to allow regular feedback to the system. To reach this goal, BCI research requires **innovative methods for signal processing** (McFarland et al. 2006) and **classification** (Lotte et al. 2007).

Then, most BCI now give the control of timing and speed of the task to the system

(Popescu et al. 2008), so that the feedback is given in an **synchronous** way. However, non invasive and efficient BCI implementation requires the development of **asynchronous BCI** (or self-paced BCI) : the computer is no longer in control of timing and speed, the operator is. The development of asynchronous BCI has been increasingly addressed in the recent years (see Borisoff et al. (2006) for instance), but the main milestone towards their implementation is the detection of **idle modes**, i.e. the moments when the operator is not engaged in the task (Parini et al. 2009, Wei et al. 2009). In particular, the introduction of fNIRS and its growing utilization in the domain of BCI has been made possible thanks to its better ability to detect such idle modes, as variations of the oxygenation level allow a good detection of on-task and off-task periods with few data (Coyle et al. 2004, Tomita et al. 2014, Durantin et al. 2014c).

Finally, most studies point towards the fact that even BCI that have high accuracy on average cannot be used by everyone (Popescu et al. 2008), as an estimated 15 to 30% of the subjects (called BCI illiterates) exhibit very low accuracy (Blankertz et al. 2009, Allison and Neuper 2010). This problem is known as the **BCI illiteracy** problem, and still represents a great challenge for next generation BCI. If some studies have found that intrinsic neurophysiological factors may be at the origin of BCI illiteracy (Kübler et al. 2004, Blankertz et al. 2010), as well as the high dimensionality and non stationarity of brain signals, it was also suggested that very high remaining noise could be at the origin of this phenomenon (Sannelli et al. 2008). Solving this problem requires two approaches : one centered on the individual, by providing a training to the BCI users that will optimize the user's strategies (Vidaurre and Blankertz 2010), and one centered on the system itself, as better signal processing and classification procedures should help us reduce inter-subject variability.

3.4 Conclusions on BCI

In this chapter, we demonstrated that BCI represent an efficient solution to improve human-machine interaction, in line with the neuroergonomics approach. Indeed, we saw that estimating the level of mental workload and/or engagement of the operator with brain measurement techniques can be used to design passive BCI. They can be used for instance to adapt the level of automation of a system or to give feedback in order to optimize the level of performance as well as the level of comfort and safety experienced by the operator. However, the development of such systems faces several challenges. First, the brain measurement techniques should be able to detect the changes associated with the executive failures described in chapters 7.1 and 7.2. Then, the signal processing techniques applied to these measures as well as the types of features extracted from them have to be optimized. Accordingly, the BCI will give efficient results with minimal calibration times,

and will be usable by the majority of people. As an answer to those challenges, the next chapters will describe the studies led during this PhD project. We first studied the ability of fNIRS to detect the executive failures associated with operator's engagement and mental workload. Then, we implemented signal processing techniques for signal processing and feature extraction, in order to design a BCI.

III

Contribution

In the previous chapters, we described the importance to evaluate both the level of mental workload (describing a level of cognitive resources mobilized to face task demands) and the level of engagement of an operator in a given task (see chapter 7.1). These two components are both indispensable when it comes to the understanding of behavior. We also discussed in chapter 7.2 the use of psycho-physiological and brain measurement techniques (especially EEG and fNIRS) to assess these characteristics, in line with the approach of Neuroergonomics. In particular, we identified that fNIRS is a promising technique for the evaluation of the state of the operator as well as the investigation of the causes of human behavior. Finally, we saw in chapter 7.3 that the use of such measures in real-time represents a great challenge for interactive systems design in the field of BCI, as we need to translate a brain activity into a command for a system.

The contributions of this PhD project follow the Neuroergonomics approach, and aim at investigating the psycho-physiological and neural correlates underlying mental workload and engagement, as well as those underlying the executive failures (e.g. disengagement, mind wandering). The first objective is the main topic of the two next chapters.

First, we focused on cognition in highly demanding situations, and investigated with fNIRS and ECG the neural correlates of mental overload. In line with the Neuroergonomics approach, we used an ecological task for this experiment in order to reproduce real world conditions.

Then, we investigated the neural correlates of cognition in low demanding situations, with a particular focus on the phenomenon of Mind Wandering. As this state had never been investigated with fNIRS, the first study we led used a controlled experiment, to assess the potential of this metric to capture mind wandering related brain activity. The results of this study left open the question of the applicability of the results in ecological situations, and suggested the use of a multimodal approach to define more precisely the state of MW. To this end, we designed a second experiment, using fNIRS and Eye Tracker in an ecological task.

Finally, the last chapter focuses on the signal processing and classification techniques that were developed during the PhD project, and which constitute its second objective. Indeed, as the Neuroergonomics approach requires the use of brain signals in real-world situations and the development of BCI, the objective was to investigate some methods usable in real-time to improve the signal collected from the participants and evaluate their mental state. The innovative methods we used include the Moving Average Convergence Divergence (MACD) filter, adapted from economic market analysis, and the Kalman filter, whose principle relies on the dynamical models simulations to improve signal quality.

Mental overload detection

4

- *We developed an ecological task with varying task demands, replicating remotely operated vehicle control.*
- *Central nervous system and Autonomous nervous system activity were measured using fNIRS and ECG.*
- *The most difficult condition elicited a decrease of oxygenation at optode 3 in the DLPFC, together with a reduced LF/HF ratio.*
- *The results suggest that highest task demands corresponded to specific strategies of resource allocation, characteristic of disengagement.*

As discussed in chapters 7.2 and 7.3, the detection of executive failures (that may lead to performance decrement) using objective measures is a key step to be able to provide help in a timely and accurate manner to the operator. In chapter 7.1, we insisted on the importance of **mental workload** and **operator engagement** as precursors of such failures. The following chapter presents our study published in Durantin et al. (2014b) and Gagnon et al. (2012) focusing on high mental demands.

4.1 Objectives of the experiment

The objective of the study was to explore CNS and ANS responses when mental capacity is exceeded in a realistic, engaging Remotely Operated Vehicle (ROV) task, in an attempt to better understand the precursors of executive failure. In particular, one of the main

goals was the assessment of fNIRS as a good indicator of mental workload level and/or operator engagement under very high mental demands, which had previously been done in laboratory settings with increasing levels of difficult in a N-back task by Izzetoglu et al. (2007). Simultaneously, we wanted to use other measures of mental workload, as a way to compare their sensitivity and capacity to detect overload. To this end, we used ECG as an indicator of ANS activity, and subjective NASA TLX (Hart and Staveland 1988) as a way to have a reference level for the difficulty levels induced by our ecological task.

4.2 Simulated piloting task

4.2.1 Protocol design

In our attempt to reproduce conditions that induce mental overload for pilots and ROV operators, we took care to preserve ecological validity while using a well-controlled laboratory protocol. We designed a simulation of a ROV operation task that specifically involved psychomotor and working memory (WM) abilities, as several studies have proved these cognitive functions to be highly correlated with complex task performance (Causse et al. 2011a, Ayaz et al. 2012a).

On this purpose, we chose to use a micro-world approach (see chapter 7.2), using a *shoot'em up* type game with horizontal scrolling (see Fig. 4.1 for an example) to reproduce typical piloting actions, leading to a task that would be both immersive and with a high functional fidelity (Jentsch and Bowers 1998, Kearney and Pivec 2007).



Figure 4.1 – *Screenshots from the Atari game "1943 : the battle for midway" (left), a classical shoot'em up with vertical scrolling from 1988, and from the Super Nintendo game "Gradius III" (right) from 1990, with horizontal scrolling.*

In this type of game, the participant views the plane he/she is piloting from above. The screen "scrolls" horizontally, giving the impression that the plane moves forward with a constant speed. Thus, the participant has only one degree of freedom to pilot the plane.

Usually, the participant has to fly by following some instructions (e.g. avoiding some obstacles, or shooting at other aircraft).

The experimental setting was designed to reproduce at a functional level the requirements associated with piloting. With this approach, we manipulated two factors (control difficulty and processing load) to generate different levels of workload, in order to keep a high degree of experimental control (Brehmer and Dörner 1993). An important reason that led us to choose to represent these aspects of piloting, additionally to the fact that they are critical in piloting situations (Oron-Gilad et al. 2006), is the ability of frontal fNIRS to capture the associated neural correlates (Ayaz et al. 2012a;b).

The use of these two factors to manipulate the workload led us to the expectation of an increasing perceived level of task difficulty, as described on table 4.1.

	Low processing load	**High processing load**
Easy control difficulty	+	++
Hard control difficulty	++	+++

Table 4.1 – *Expected levels of task difficulty induced by the variations of the two factors (processing load and control difficulty).*

It is also important to note that the choice of a horizontal scrolling game (like in *GradiusIII*, see Fig.4.1) rather than the commonly used vertical option (like in *1943* on Fig.4.1) was made to spread out the different components (piloted plane, target planes and cue) across the screen. This arrangement of information made it easier to capture saccadic ocular activity with an eye-tracker [1]. The computerized simulation was coded using Java.

4.2.2 Task description

The subjects were instructed to follow a target aircraft with the piloted one, by minimizing the distance between them, as a player of a *shoot em up* game would do. Visual stimuli were presented on a DELL 21" monitor placed one meter from the participants (see Fig.4.2 for a screenshot).

The own aircraft was located at about 24cm from the left of the screen, and target aircrafts were moving approximately 2cm to 4cm from the left. The target aircraft was indicated to the subject by a visual cue presented on the right-hand side of the screen

[1] A SMI RED500®fixed eye-tracker was originally intended to be used, but was removed from the experiment because of the interference between fNIRS and eye-tracking measurements, as they both rely on infrared light. We focused on the isolation of the two measurements later, when designing the fNIRS/Eye Tracker Mind Wandering Experiment.

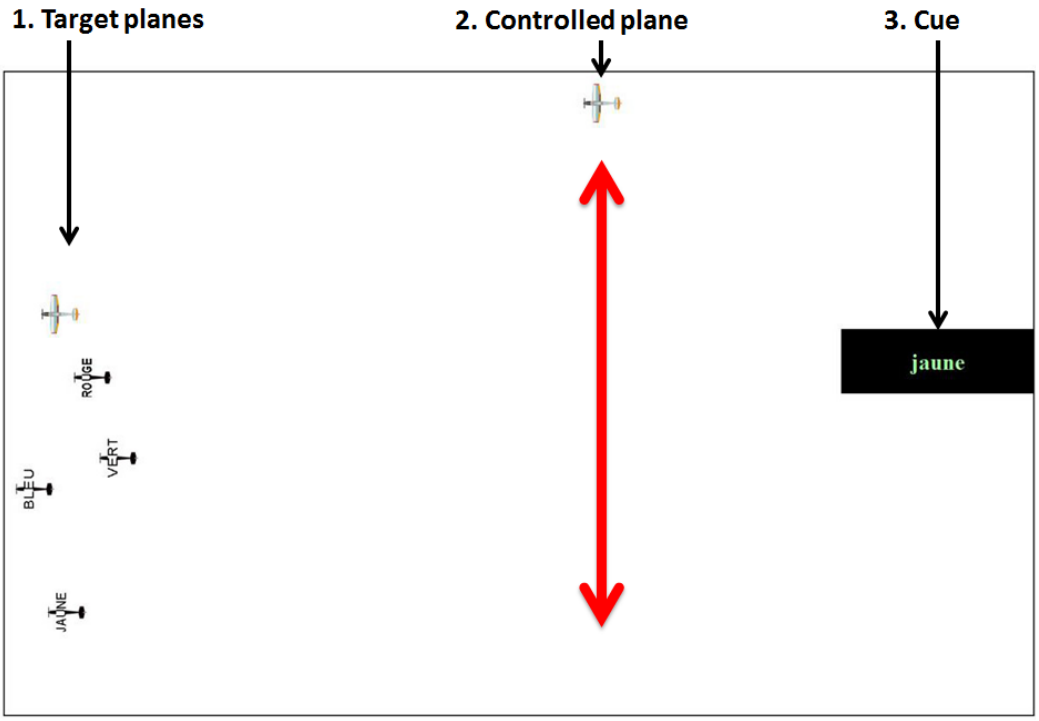


Figure 4.2 – *Screenshot from the task. The participant had to control the plane (2), to follow the target plane (1) indicated by the cues (3) presented regularly. The red arrow indicates the possible movements of the controlled plane.*

(approximately 38cm from the left). The cue selection rule was inspired by the Stroop paradigm in order to reproduce high-level cognitive functions such as inhibition and cognitive control, critical in piloting operations (Oron-Gilad et al. 2006).

A cue consisted of a color name written with red, blue, green or yellow ink (see Table 4.2 for examples). The color name and the ink color were pseudo-randomly chosen so that they could either be congruent or incongruent. Target aircraft had a color name written on their wing, and participants were asked to follow the plane corresponding to the ink color of the given cue. Moreover, 20 percent of the cues contained a word which was not a color name (e.g., "read", "grin"...). In these cases, the participants were asked to follow the unnamed plane (regardless to the color of the ink). A new cue was presented for 1.6 seconds every 8.6 seconds. The interval between two presentations is referred to as an "object tracking phase".

Participants completed four experimental sessions resulting from the manipulation of two levels within two factors: difficulty of control and processing load. There were two levels of difficulty of control (easy or hard) manipulated by varying the strength of the crosswind (no crosswind in the easy condition, strong crosswind in the hard condition) and the inertia of the plane (low vs. high). The processing load was varied in terms of working memory, with an N-Back-like sub-task. It has been shown that processing load can be varied by manipulating the value of N, which is the number of items to

Cue	Target plane	Reason
	Blue	Ink color = blue
	Yellow	Ink color = yellow
	Unnamed plane	Text = not a color name

Table 4.2 – *Cue examples*

be maintained and manipulated in working memory. Subjects had to target the aircraft corresponding to the last cue presented (N; low load condition) or the cue before (N-1; high load condition). The combination of the two factors yielded a $2x2$ repeated-measures design with four conditions:

- low processing load / easy control
- low processing load / hard control
- high processing load / easy control
- high processing load / hard control

During the experimental sessions, participants also had to stop a randomly-initiated auditory alarm, by pressing a button on the joystick. Before the start of the experiment, subjects were taught to recognize a 2000ms duration, and were asked to stop the alarm 2000ms after it began.

4.3 Procedure and Data Collection

Twelve volunteers gave their consent to participate in the study (mean age = 25; SD = 5.25; 10 males ; 10 right-handed). The volunteers were fully informed about the experimental protocol. They were given financial compensation for their part in the study. A picture of participant setting is shown on Fig. 4.3.

Participants were first trained in the piloting task. The training session consisted of 10 object tracking phases for each processing load level, and was performed at both the easy



Figure 4.3 – *Task setting. The participant sat in front of the monitor where the task was displayed, and interacted with the task using a joystick.*

and difficult levels of control. After the training session (10 minutes approximatively), participants performed the four experimental sessions consecutively (the order of the experimental sessions was counterbalanced across participants). Each session comprised 40 object-tracking phases and lasted approximately six minutes. During each session, hemodynamics of the PFC (i.e. inferred by variations of the oxygenation level) was recorded at a sample rate of 2Hz using a fNIR100 (Biopac®) equipped with 16 optodes, and the acquisition software COBI Studio®. Heart rate data was recorded at a sample rate of 2048 Hz using the Biopac®ECG.

After each session, participants filled out the NASA-TLX (i.e. subjective workload). At the end of the experiment, they also filled a questionnaire to give their general impressions about the task.

4.4 Data analyses

Behavioral

Using the distance between the piloted and the target planes, we determined whether or not the operator followed the correct plane; that is, during any particular object tracking phase, when the mean distance between the piloted plane and the target exceeded 3.1 times the wingspan of the planes, the current tracking phase was considered a failure. A performance score was then computed for each tracking task over the experimental session.

fNIRS

For each optode, fNIRS data (HbO_2 concentration relative to a 10 seconds baseline) was normalized, and mean HbO_2 concentrations for each condition were calculated over the session, using MatLab®. Visualization of the data was performed using fnirSoft®.

ECG

Using ECG data, mean LF/HF Ratio was estimated over each session using Fast Fourier Transformation (FFT) of the ECGLAB toolbox (de Carvalho et al. 2002) for MatLab®.

Statistical analysis

Repeated-measures ANOVAs with within subjects factors *processing load* (low vs. high) and *control difficulty* (easy vs. hard) were carried out to test whether the effects of control difficulty and processing load were statistically significant on the different measures (heart rate variability, HbO_2 concentration on each optode, object tracking performance, and NASA-TLX score for global workload index and for each of the criteria).

4.5 Results

Performance at the tracking task

Performance on the object tracking task (success rate of planes followed correctly) is shown in Fig.4.4 . A repeated measures ANOVA showed a significant decrease in success rate with processing load, $F(1,11) = 50.69$, $p < 0.001$, and a trend of decrease with control difficulty, $F(1,11) = 5.55$, $p < 0.05$. This effect corresponded to a lower success rate for high processing load and for hard control difficulty. Moreover, there was a significant interaction $F(1,11) = 4.85$; $p < 0.05$. Thus, the performance decreased significantly with high processing load compared to low, but only within the hard control condition (in the easy control condition, processing load did not induce any significant modification of the subjects' performance)

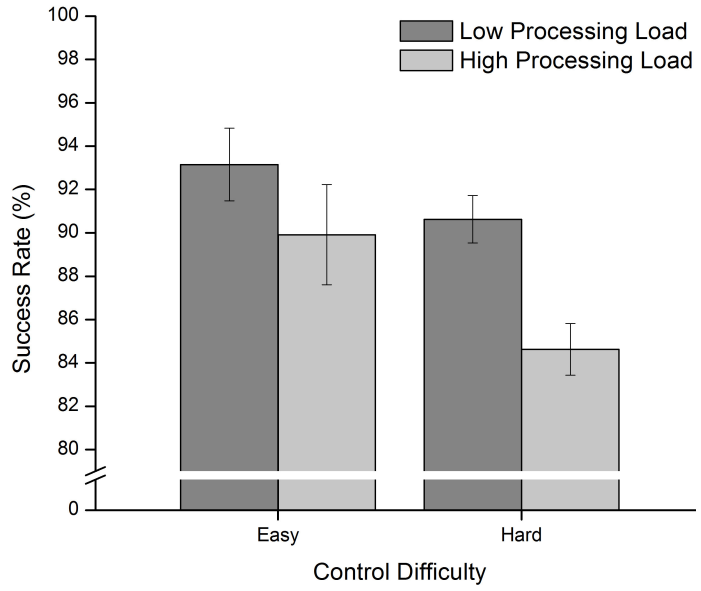


Figure 4.4 – *Success Rate in percent for the four experimental conditions. Error bars represent the standard error of the mean.*

Perceived workload

The ANOVA conducted on the NASA-TLX data with Bonferroni-Holm correction for multiple testing revealed significant main effects of processing load, $F(1,11) = 25.01$, $p < .001$, and difficulty of control, $F(1,11) = 4.70$, $p = .053$, indicating a significant increase in the perceived workload with high processing load or high control difficulty. However, the two-way interaction was not significant, $F(1,11) < 1$.

Mental demand ($p < 0.001$) and *Frustration* ($p < 0.05$) dimensions of the NASA TLX test were the most affected by the increase in task difficulty. Fig. 4.5 shows how variations of experimental parameters affect perceived mental load.

In the questionnaire at the end of the experiment, 9 participants out of 12 explicitly insisted on the fact that the task became particularly difficult for them when both the control difficulty (the crosswind leading the plane to respond less intuitively to the pilot's commands) and the processing load were high (requiring for the subjects to have the cues constantly in mind).

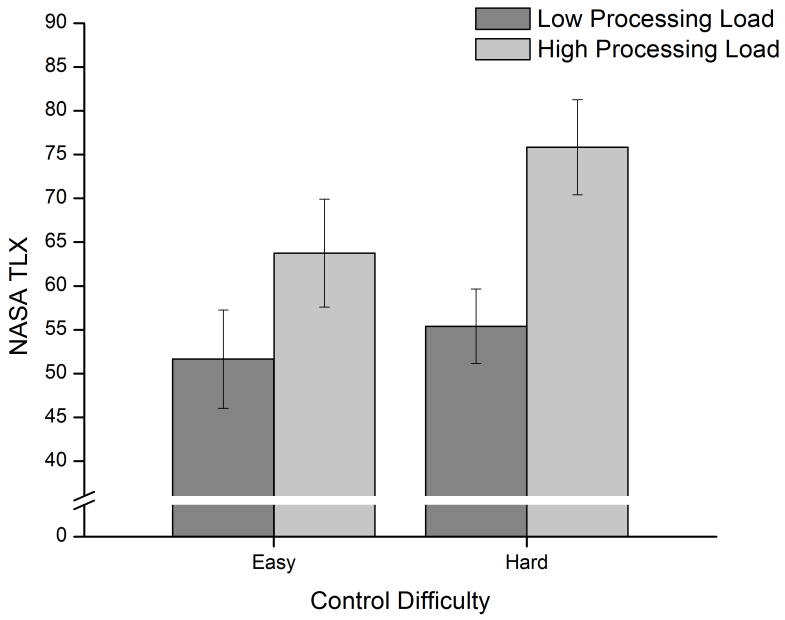


Figure 4.5 – *NASA TLX subjective workload rating for the four experimental conditions. Error bars represent the standard error of the mean.*

fNIRS

The mean HbO_2 level measured on each of the 16 optodes for each of the experimental sessions is shown on Fig.4.6.

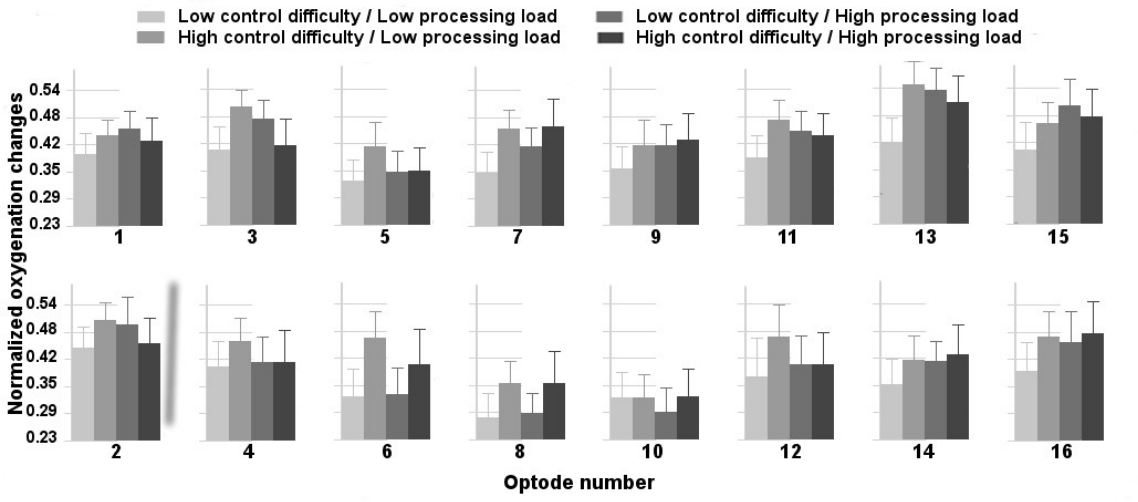


Figure 4.6 – *Normalized oxygenation data for each optode over the four sessions. The decrease in oxygenation during the last session, although being significant only at optode 3, is visible on many optodes, especially those in the dorsolateral prefrontal cortex (optodes 1, 3, 13 and 15).*

A main effect of control difficulty ($F(1,11) = 5.82$; $p < 0.05$) was observed on optode

6, showing an increase in the level of HbO2 with an increase in piloting difficulty. A significant two-way interaction was also visible on optode 3 ($F(1,11) = 5.11$; $p < .05$), revealing an increase in HbO2 concentration for the higher processing load in the easy control condition, but a decrease for high processing load in the hard control condition. The result is shown in Fig. 4.7 for optode 3, located in the left dorsolateral prefrontal cortex. However, although the effect is significant only on optode 3, the same tendency of variation (increasing HbO_2 with higher perceived workload, but decrease for the hardest condition) is visible on all optodes in the right and left dorsolateral prefrontal cortex (see Fig.4.6 for the data from all optodes). No effect was found on HHb data.

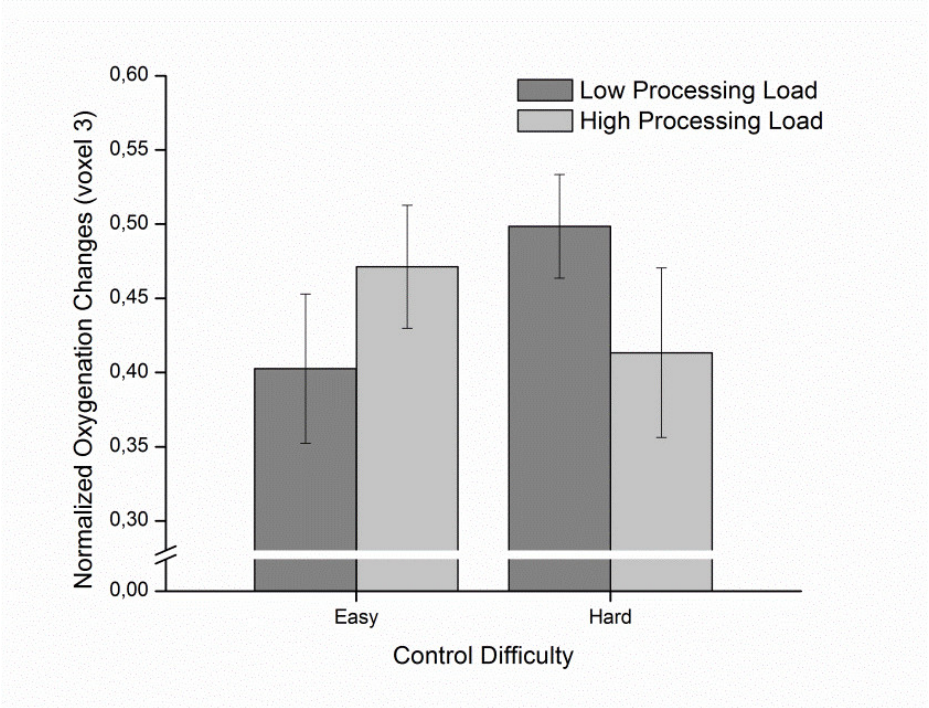


Figure 4.7 – *Normalized mean HbO_2 changes across the four experimental conditions on optode 3. The error bars represent the standard error of the mean.*

ECG

Average LF/HF ratio for each session was extracted from ECG data. Fig.4.8 shows average LF/HF ratio under each experimental condition. The ANOVA exhibited a significant two-way interaction ($F(1,10) = 9.45$, $p < .05$). Within the easy control condition, LF/HF ratio was greater with high than low processing load; however, at the high level of control difficulty, LF/HF ratio was then lower with the high than the low processing load. This result is similar to the fNIRS results that showed the same pattern (a decrease of average HbO_2 concentrations for high processing load in the hard control condition).

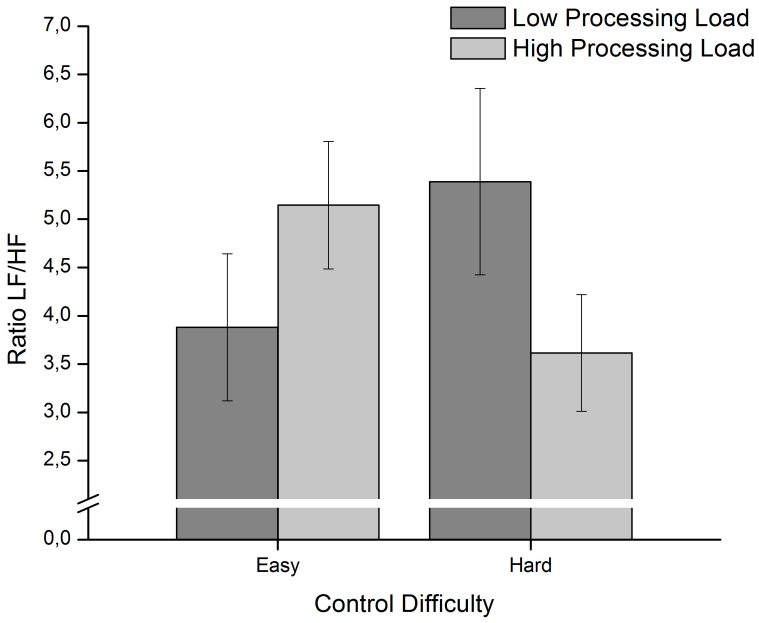


Figure 4.8 – *Average LF/HF Ratio across the four experimental conditions. The error bars represent the standard error of the mean. The LF/HF Ratio in function of experimental sessions shows the same pattern of evolution as HbO_2 concentration measured by fNIRS in the dorsolateral prefrontal cortex (see Fig. 4.7).*

4.6 Discussion

The objective of the experiment was to investigate potential measures of mental workload in a simulated environment, especially when mental demand exceeds cognitive resources. The use of a multimodal approach with three different measures (self-report, HRV and fNIRS) allowed us to compare the contributions of different methods to the evaluation of mental workload and the detection of executive failures.

The first issue of the study, as discussed in part 4.2.1, was to come up with a laboratory task which would be sufficiently engaging to elicit different levels of mental demand and simulate piloting with a high functional fidelity. Subjective NASA-TLX scores demonstrated that subjective mental workload increased across experimental conditions. In particular, from the operators' perspective, there were real differences in terms of frustration and mental demand across the different task conditions. The subjective reports from the subjects at the end of the experiment also emphasize the increasing mental demand as the control difficulty and the processing load increased.

Processing load and control difficulty both impacted upon aircraft tracking performance, as shown on Fig. 4.4, especially with a significant drop of the success rate in the hardest condition compared to all the other sessions. Moreover, fNIRS results suggest our

task successfully involved high-level executive functions as changes in oxygenation level of some areas of the PFC were observed. First, control difficulty modulated oxygenation level in the optode 6 (close to the anterior medial PFC). This result is similar to previous studies as Ayaz et al. (2012a), realized with the same fNIRS device, that exhibited anterior medial PFC activations during aircraft's trajectory supervision. Though we did not find a main effect of processing load on PFC activation, our results revealed that interactions exist between subsets of task difficulty as showed by changes in oxygenation levels of optode 3 within the left DLPFC. Again, the findings are similar to previous neuroimaging studies showing that this area is particularly sensitive to mental workload, both in controlled WM laboratory tasks (Smith and Jonides 1997, Owen et al. 2005, Schreppel et al. 2007) and ecologically valid tasks simulating piloting operations (Ayaz et al. 2012a, Menda et al. 2011).

Indeed, variation of oxygenation level in optode 3, a part of the left DLPFC, tended to follow the inverted U-shaped pattern (see Fig. 4.9): HbO2 concentration increased progressively across difficulty levels but then decreased significantly during the most difficult condition. This finding is particularly interesting given that the DLPFC is considered to be a major anatomical correlate of the central executive (Baddeley 2003), a region that plays a key role in task supervision and cognitive control (Qin et al. 2009). This decline in DLPFC activity corresponded to the poorest performance, and was associated with the highest subjective mental load and the strongest feelings of frustration.

Altogether, these results suggest that the volunteers were unable or unwilling to mobilize cognitive resources despite task demand, and may have faced mental overload. The correlation between left DLPFC activity and performance showed that this was particularly true for some of the participants, as those with the lowest levels of activation exhibited the poorest performance during the hardest experimental condition. This pattern of DLPFC lower activation associated with deleterious effects on performance has previously been observed using fMRI or EEG during stressing (Qin et al. 2009) conditions or working memory overload (Van Snellenberg et al. 2015). The pattern revealed with fNIRS is similar to previous results obtained with laboratory cognitive tasks manipulating WM load (Izzetoglu et al. 2007).

A possible explanation of the pattern observed on hemodynamics is related to the role played by memory updating in multitasking. Indeed, as stated by Wickens (Wickens et al. 1983a) : "*the resources on which this updating activity depends seem to be limited in their availability, and, when deployed in the service of one task, their availability to be of service to other tasks is reduced*". Along these lines, load sensitive brain areas have been shown to elicit either transient or sustained activations over time; a pattern that is often taken to suggest a distinction between areas involved in active maintenance and those involved in time-limited cognitive activities such as memory updating (Cohen et al. 1997, Rowe and Passingham

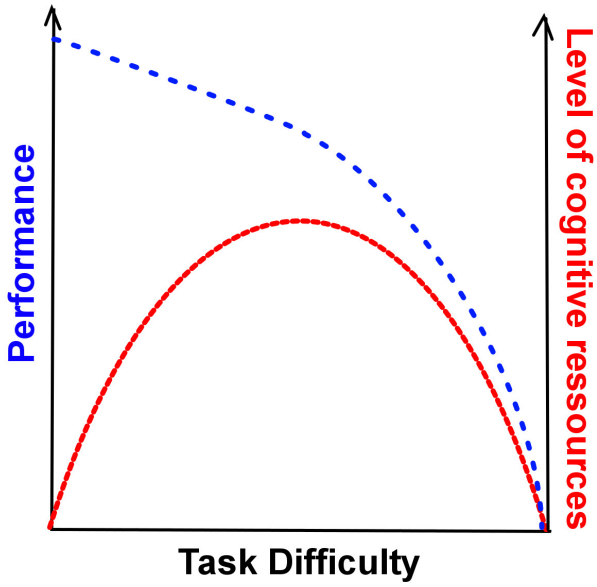


Figure 4.9 – *Evolution of the level of cognitive resources allocated and performance in function of task difficulty, as observed in this task.*

2001, Barch et al. 1997). Thus, the possible role of the dorsolateral prefrontal cortex in transient activations during multitasking (Rowe and Passingham 2001) could explain the decline of its activity during the most difficult condition in our experiment. However, the implication of the DLPFC either in the maintenance of information (Barch et al. 1997, Cohen et al. 1997) or in updating (Rowe and Passingham 2001, Koechlin et al. 2003) remains uncertain. Further work is required to investigate whether activations observed in the DLPFC were induced sustainably or transiently. More importantly, in accordance with our objectives, the observed inverted U-shaped pattern of DLPFC activity in response to increased mental workload could also be explained by the fact that this study manipulated task difficulty to the extent that it **exceeded participants' mental resources**. Although it could be argued that this apparent disengagement reflects a lack of motivation leading the participant to somehow abandon the task, in our experiment, the volunteers were constantly attempting to adjust the aircraft trajectory even in the most difficult condition, suggesting that their disengagement occurred as a consequence of mental demand increase instead of motivation. Moreover, it is worth noting that HRV, assessed through the LF/HF ratio, followed a similar pattern with a significant decrease during the hardest condition. This diminished influence of the sympathetic nervous system on the ANS suggests reduced catabolic activity and lower mobilization of mental resources to deal with the situation. As we saw in chapter 7.2 that this sensor provides a good indicator of mental load, this measure provides supplementary evidence to our fNIRS findings in favor of a **deleterious mechanism on resource management induced by mental overload**. Such a pattern of physiological response to mental overload has already been observed on other physiological responses such as pupil diameter (Granholm et al. 1996), suggesting a

change in ANS action to face mental overload. Moreover, it tends to confirm that mental overload could result from a lower level of both cerebral and ANS activities. Interestingly enough, and although no correlation between HRV data and PFC modulation was found here, it is worth noting that such correlations between cardiac metrics and prefrontal neuroimaging data have been found previously (Gianaros et al. 2004, Napadow et al. 2008). Indeed, these results provide a direct observation of the interactions between the PFC and cardiac activity as postulated by previous studies (Boucsein 2000) or by neurovisceral integration models (Thayer and Lane 2009), although our results do not allow us to determine whether these interactions are mediated by a common factor, or whether there is a causal link between neural and cardiac activity.

4.7 Conclusion

This study shows the usability of fNIRS to assess user's mental workload and engagement during ecological experiments and the possibility of using complementary behavioral (cognitive) and physiological measurements to derive the operator's functional state. Ultimately, although executive failure (here, disengagement linked to mental overload) had an impact on the psychophysiological and cerebral objective measurements, the fact that high perceived mental load could be associated with a related decline of the central and peripheral nervous systems might represent an issue for adaptive automation perspectives. Indeed, in the case of our study, the analysis of the level of cardiac and prefrontal activities does not allow us to formally discriminate the easiest from the hardest condition. To this end, one has to take into account multiple factors such as operator engagement and performance, as emphasized in chapter 7.1. This must be taken into account in the development of future adaptive automation by considering other measurements such as behavioral performance (e.g., operator's reaction time and actions on the user interface, eye movement) as well as the state of the global system (e.g., failure).

Mind wandering

5

- *We conducted two experiments to investigate the neural correlates of mind wandering with fNIRS.*
- *The controlled experiment confirmed the ability of fNIRS to detect the activations of the DMN during mind wandering.*
- *The ecological experiment, using both fNIRS and Eye Tracker, suggested the implication of both the DMN and the locus cœruleus network of vigilance in the generation and maintenance of mind wandering.*
- *The results suggest that fNIRS, combined with Eye Tracker, would allow online characterization of the states of cognition under low demanding conditions.*

In the previous chapter, we focused on high mental demands placed on operators and were able to observe and characterize the executive failures associated with such situations. However, as we discussed during chapter 7.1, very low mental demands also have deleterious effect on performance, as mind wandering causes the operator's attention to drift away from the task. The objective of this chapter is the characterization of those events using objective measures, first in a very controlled task, then in a more ecologically valid task. Indeed, the approach of Neuroergonomics is crucial when it comes to char-

acterization, as mind wandering episodes are especially present in real-world situations (Killingsworth and Gilbert 2010).

5.1 First experiment : assessing mind wandering with NIRS

5.1.1 Objectives of the experiment

The deleterious aspects of mind wandering on performance have been highlighted in chapter 7.1. Similarly to the experiment led in the previous chapter (7.4), the objective is to characterize the occurrence of executive failures in operational situations by using objective brain measurement techniques. Mind wandering has already been characterized using eye tracking (Smilek et al. 2010, Reichle et al. 2010), and was associated with variations in eye blinks and movements, or in EEG (Smallwood et al. 2008, Braboszcz and Delorme 2011), as it was associated with reductions of the cortical processing measurable using ERP amplitudes and cortical rhythms. Concerning its neuroanatomical substrates, the recruitment of the DMN during MW was studied using fMRI (Christoff et al. 2009, Hasenkamp et al. 2012) (see also chapter 7.1). In fNIRS, even if the possibility to measure the recruitment of the DMN has been put in relief by some studies (e.g. Harrivel et al. (2013)), their measure in the context of mind wandering has never been studied to my knowledge. To that end, one of the principal objectives of the study was to assess the ability of fNIRS to detect cortical activations in the context of MW, which would make it a suitable technique for the monitoring of the attentional state in neuroergonomics. The study was led in collaboration with two ISAE students, Marc Mari-Mari and Alvaro Miedes, who helped me during task coding and data collection. The results of the study were published in Durantin et al. (2015).

5.1.2 Methods

5.1.2.1 Task

Design

To study MW in a controlled environment, the first challenge was to come up with a task simple enough to study MW. Another objective of the experiment was to dispose of enough MW samples to train a classification model, without having a too long experiment. Studies on MW have usually involved sustained attention paradigms such as breath counting or go-no go tasks. One example of a go-no go task includes the popular Sustained Attention to Response Task (SART), wherein single digit numbers appear one at a

time on a computer screen: subjects are instructed to press a button whenever a number other than the target number (3) appears. As described by Manly et al. (1999), the SART is simple so subjects' attention frequently leaves the primary task, in addition to being sensitive to the tendency for participants to automate their behavior. During episodes of MW, subjects tend to press the button systematically (even when the number 3 appears). By counting the number of successive missed targets, it is possible to assess the time and frequency of MW episodes. Indeed, As outlined in Smallwood and Schooler (2006), investigations of the experience of MW during the SART support that, first, blocks in which this phenomenon occurs are associated with faster response times than are blocks in which attention is directed towards the task. Second, high levels of MW reported with retrospective questionnaires are associated with a tendency to make an error during periods of task disengagement.

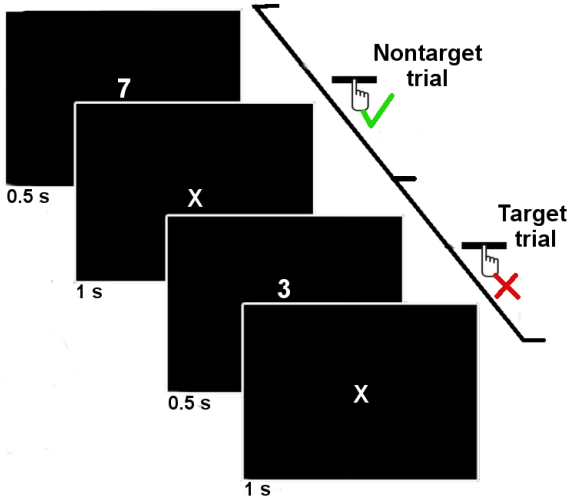


Figure 5.1 – *Time course of two trials of the SART protocol. The digits were presented 500ms with an onset-to-onset interval of 1500ms. The subjects were asked to press the spacebar when a nontarget digit was presented, and not to respond when a target digit (3) was presented.*

Procedure

Subjects were asked to perform a computerized SART task. The SART task consists of a simple go/no-go task in which a single infrequent target digit is presented (here the digit 3) amongst frequent non-targets digits (1–9). The computer screen was placed approximately 70 cm from the participants' head. Each digit was presented for 500 ms on the computer screen and then replaced by a fixation mark (X) for 1000 ms (see Figure 5.1). Digits appeared in white on a black background and were approximately 3 cm high in Arial font. The participants were asked to press the spacebar of the computer keyboard for non-target digits, and not to press it if the digit was a target (3). The target trials for which the participants inaccurately pressed the spacebar were considered *SART Errors*.

In the rest of this chapter, we will designate by the term *SART No errors* the other target trials, i.e. when the participant did not press the spacebar (as requested).

The experiment consisted in the presentation of two blocks of 198 trials, 22 of them (11%) being target trials. The proportion of target trials was chosen in accordance with Manly et al. (1999), as in their experiment MW occurrence was higher with this ratio. Targets presentation was pseudorandom, and ensured that two target trials would not follow each other. The mean interval between two target trials was 11.6 s (SE: 3.35 s). Preliminary to the experiment, for training purposes, subjects performed 18 practice trials (2 of which were targets). Subjects were told to respond as accurately and as quickly as possible to digit presentations.

5.1.2.2 Participants

Twenty-three male students from ISAE school of engineering gave their consent to participate in the experiment (21 right-handed; age range 21–24 years, mean age: 22.6). All participants reported normal or corrected vision, and none were suffering from neuropsychological problems. Prior to the experiment, subjects were provided instructions for the SART task. No explanations were given about the phenomenon of MW, in order to avoid this having an impact on their performance. At the end of the experiment, subjects were asked to complete a questionnaire where they could report their opinion about the task and whether they experienced any difficulties. In the questionnaire, we also asked subjects whether they had any thoughts that were unrelated to the task, to retrospectively test for MW meta-awareness.

5.1.2.3 Data acquisition

During the experiment, hemodynamic data from the PFC were recorded using a fNIR100 device (Biopac Inc.). The device consists of four light-emitting diodes (LED) sources of $730nm$ and $850nm$ (LED current: $12mA$), and ten detectors. The source and detectors are separated by 2.5 cm, resulting in 16 optodes uniformly placed on a rectangular grid on the forehead. Data were collected with a sampling frequency of 2 Hz. A baseline of 10 s at the beginning of the experiment was used to calibrate the device. Optodes 1, 3 and 5 were defective (due to two defective detectors) and removed from all subjects.

5.1.2.4 Data processing

The 95% confidence intervals for the number of SART errors committed during the task were calculated using the SingleBayes method from Crawford and Garthwaite (Crawford and Garthwaite 2007), and the probability density function of the apparition of SART errors during the experiment was estimated using Statistica®. To this end, we used periods of 30 s and proceeded to a Weibull model estimation, a probability distribution commonly used in survival analysis for event density probability function estimation (Kennedy and Gehan 1971).

The density function describes the probability that a SART Errors happen at a given time, and is mathematically defined as follow:

$$pdf(t) = \frac{\mathcal{P}(t < SARTError < t + dt)}{dt} \tag{5.1}$$

Hemodynamic data recorded by fNIRS were processed using the EEGLAB toolbox for MatLab®(Delorme and Makeig 2004). Interestingly enough, even if EEGLAB is originally intended for EEG data, it proposes a large panel of processing options (ranging from artifact processing to statistics) for event-related analysis. As data processing will focus on the concentration dynamics around the target events (*SART Errors* or *SART No Errors*), this toolbox allowed efficient processing. Continuous data were first high-pass filtered using a short non-linear Infinite Impulse Response (IIR) filter of order 6 and a cut-off frequency of 0.02 Hz. We extracted epochs from continuous data for the SART Error and SART No Error conditions, starting 30 s before stimulus apparition, and stopping 10 s after. The significance of the hemodynamic changes mapped using fNIRS was tested at the topographical level, using the Montecarlo Statistics and using the Cluster Correction for multiple comparisons to correct for multiple optodes measurements (Maris and Oostenveld 2007).

5.1.2.5 Classification

As one of the main objectives of the study was to assess the ability of fNIRS to predict executive failures such as mind wandering, the data processing stage included a classification step. We used Linear Discriminant Analysis (LDA) for the classification of MW epochs as compared to epochs in which subjects were concentrated. Non-target epochs were not considered here. We considered only MW epochs, with either a response (correct detection) or no response (incorrect detection). LDA is a simple classifier that is robust and fast to compute. As explained on Fig. 5.2, we used a 10-fold cross-validation approach to test our model. This means that there were 10 iterations to the algorithm. For

each iteration, 90% of the data were used for learning the LDA weights and the remaining 10% of the data were used to test the model. The portion of data used for testing was always different from the previous iterations.

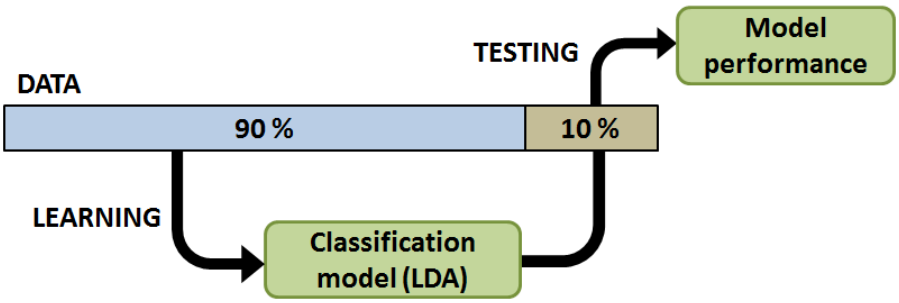


Figure 5.2 – *Cross validation process : Data were randomly divided into ten parts. 90% of the data were used to train the model, then we tested the model on the 10% remaining data to obtain the performance of the classifier. This process was repeated ten times by permuting testing and training data.*

5.1.3 Results

5.1.3.1 Behavioral

Over the 44 target trials of the SART task, the participants made a mean of 12.7 errors (standard deviation: 7.0), which represents 29% of the target trials. Figure 5.3 shows the probability density function of the apparition of SART errors during the two sessions of the experiment. This figure exhibits the increasing SART Errors density over time, showing that the subjects committed more SART errors as the duration of the experiment increased.

At the end of the experiment, 19 out of the 23 participants stated in the questionnaires that they found it difficult to stay focused on the task and had irrelevant thoughts during the experiment. The four subjects who did not mention this fact committed an average of 10.8 errors (SD: 6.45), whereas the 19 subjects conscious of MW committed on average 13.1 errors (SD: 7.37). Although the small sample size for the first group of four subjects did not permit group level statistical testing, the individual comparison of each of these subjects with the group of subjects who reported MW revealed that none of the subjects who were not aware of MW committed significantly less SART Errors than the 19 others.

5.1.3.2 Hemodynamics

Figure 5.4 shows the topography of HbO_2 concentration under both *SART Error* and *SART No Error* conditions, from 15 s until 5 s before the apparition of the stimulus. After

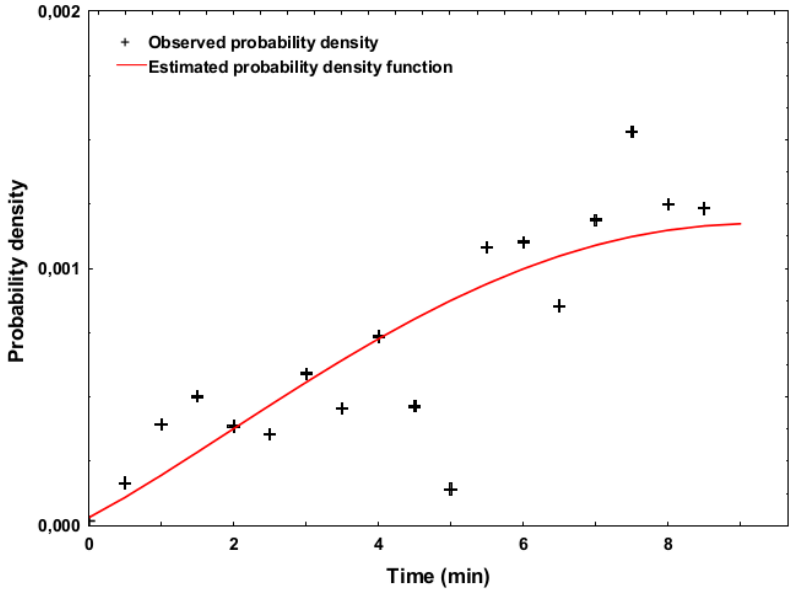


Figure 5.3 – *Estimated probability density function representing the occurrence of SART errors across time (Weibull model, estimated over the two blocks of the experiment using unweighted least-squares regression).*

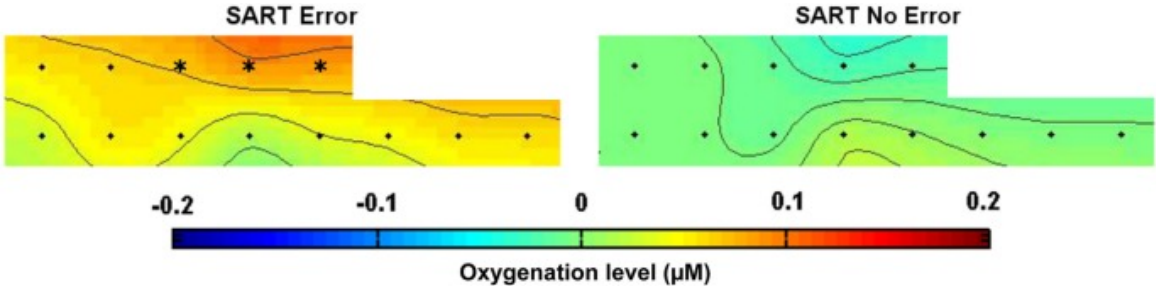


Figure 5.4 – *Topography of HbO_2 concentration over the prefrontal cortex during SART Error (on the left) and SART No Error (on the right) conditions, averaged $[-15s; -5s]$ before the apparition of the target stimulus across all subjects. The color code represents the level of HbO_2 concentration changes relative to baseline (in μM). Optodes exhibiting significant differences (all in the mPFC) are marked with a * (significance level = 0.01 after correction for multiple comparisons)*

correcting for multiple comparisons, our analysis revealed significantly higher levels of HbO_2 measured at optodes 7, 9 and 11, located in the dorsomedial prefrontal cortex, preceding SART Errors (associated with MW episodes). Figure 5.5 represents the temporal dynamics associated with optode 9, and shows that the greater activation observed in the medial prefrontal cortex (mPFC) before SART Errors returns to normal before the arrival of the stimulus. No significant variations relative to SART Errors were found on the deoxy-hemoglobin (HHb) signal.

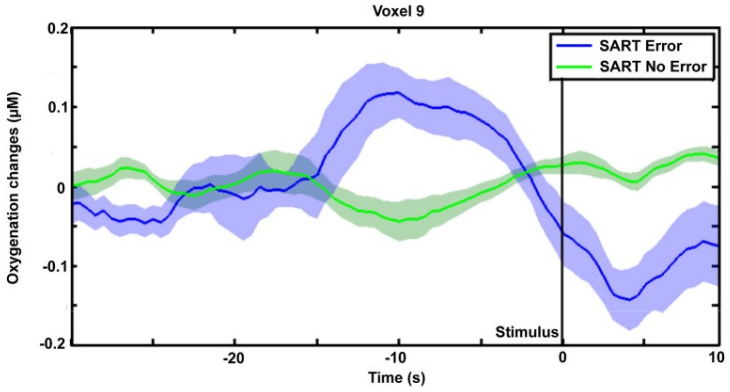


Figure 5.5 – *Variation of HbO_2 concentration (averaged trials across subjects) on optode 9 (in the mPFC) preceding SART Error (in blue) and SART No Error (in green) trials. Shaded areas represent the standard error of the mean for each condition.*

	Mean	**Standard Error**
Accuracy	56%	2.70%
Sensitivity	52%	6.33%
Specificity	62%	4.82%

Table 5.1 – *Mean mind wandering episodes classification performance across subjects.*

5.1.3.3 Classification

We used data from only 11 subjects to classify SART Error and SART No Error trials, keeping only the subjects who made at least 10 errors, so we would have enough trials to train and test the classifier. In accordance with the hemodynamics results (see section 5.1.3.2),the classifier used the average HbO_2 level at optodes 7, 9 and 11 as regressors for classification between SART error and SART No error trials. The classification performance results are summarized in table 5.1, and give the **accuracy** (percentage of correctly classified trials), the **sensitivity** (percentage of SART Error trials correctly classified) and **specificity** (percentage of SART No error trials correctly classified) of the LDA classifier. Although the mean accuracy level is only 56% (SE:2.70%), 7 subjects out of 11 had a classification accuracy superior to 60% 'see individual results on Fig. 5.6). A Wilcoxon sign test showed that this result was unlikely to occur by chance ($p < 0.016$; degree of freedom of 10).

5.1.4 Discussion

We showed that MW occurred during the SART, as 19 out of 23 participants retrospectively reported MW meta-awareness during the experiment. First, the high number of errors

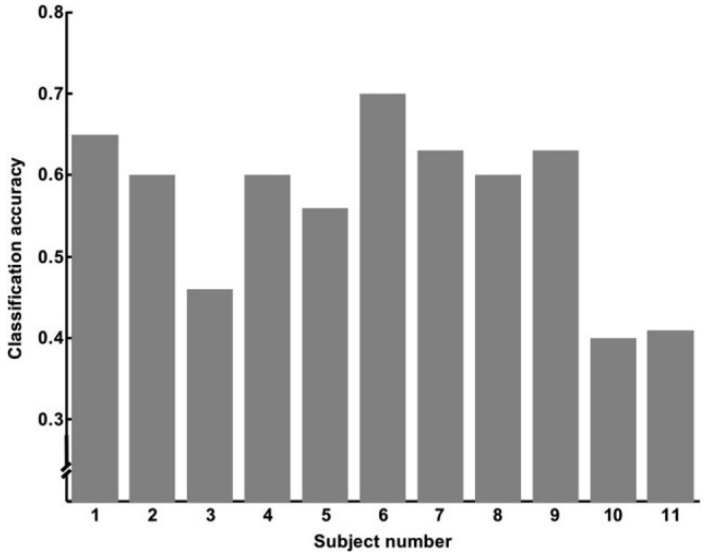


Figure 5.6 – *Accuracy per subject obtained for the classification of SART Error vs. SART No Error trials.*

made during target trials is consistent with SART characteristics regarding MW (Manly et al. 1999). Notably, the increasing occurrence of MW episodes with the time spent on the task as shown in Figure 5.3 is also consistent with previous findings using sustained attention tasks (Mackworth 1948). However, contrary to previous research (Smallwood and Schooler 2006), the retrospective awareness of MW could not be associated with a higher number of SART errors. Indeed, although 4 out of 23 subjects did not report experiencing MW, it is noticeable this was not associated with the absence of SART Errors (mean number of errors: 10.8), therefore showing that SART Errors occurred during the experiment, even without MW meta-awareness.

Our results showed that it is possible to characterize responses in which subjects are not paying attention to the task, compared to periods where they are attending based solely on the fNIRS signal. In fact, we observed significant differences in the fNIRS signal measured in the mPFC (part of the DMN) preceding the cue presentation for correct vs. incorrect trials. Notably, the medial prefrontal activations we observed prior to SART errors (see Figure 5.4) were similar to activations observed previously in fMRI with the same task (Christoff et al. 2009). These results, in conjunction with previous neuroimaging MW investigations (Mason et al. 2007a, Hasenkamp et al. 2012), support the feasibility of using fNIRS to detect DMN activity (Harrivel et al. 2013), and its implication in the phenomenon of MW. Our results did not show implication of the Executive Network areas such as dorsolateral prefrontal cortex (DLPFC) during SART errors, as revealed by two previous fMRI studies (Christoff et al. 2009, Stawarczyk et al. 2011). In these studies, DLPFC activation was related to MW meta-awareness but not to SART Errors. However, our experimental paradigm was not designed to measure MW meta-awareness, as SART Errors appeared even in the absence of MW meta-awareness, which did not allow to state about DLPFC activations. Moreover, the absence of left DLPFC activations could be due

to the fact that the optodes in this area were defective and had to be removed from the study.

An important contribution of our study concerns the temporal dynamics of the activations observed in the mPFC. This area did not show sustained activation during MW episodes, as the level of HbO_2 measured faded even before target stimulus apparition (see Figure 5.5). Similarly, previous investigation of MW using SART protocol in fMRI only found DMN activations immediately preceding MW occurrences (Hasenkamp et al. 2012, Christoff et al. 2009). This observation would suggest that the contribution of the mPFC to the DMN is strongest at the beginning of MW episodes, which would be consistent with the role of the mPFC in the development of the DMN during sleep as shown by functional connectivity MRI studies (Sämann et al. 2011). Hence, this transient activation of the mPFC could suggest that this brain area plays a role in switching from a concentrated state to MW. However, further investigation using self-caught MW protocols and analyzing the temporal dynamics of brain activation would be needed to support this hypothesis.

Nevertheless, if fNIRS was able to characterize the executive failures happening in case of low task demands, some limitations remain concerning the interpretation of SART Errors. Although previous research has demonstrated that performance on the SART task is mainly determined by the capacity to endogenously sustain attention (Manly et al. 1999, Smallwood and Schooler 2006), there remains an ongoing debate as to whether SART errors could be related to impulsivity in subjects' responses and the absence of motor inhibition (Helton et al. 2009, Stevenson et al. 2011). If previous investigations including both SART and experience sampling experiments performed online support that SART errors are linked to MW (Christoff et al. 2009); our experiment cannot conclude on this point. Indeed, despite the high rate of retrospective assessment of MW, the occurrence of SART Errors even without meta-awareness during the experiment suggests that MW alone cannot account for the totality of the SART Errors. Further investigation is needed concerning the presence or absence of left DLPFC activations, as the optodes in this area had to be removed, and to confirm that the activations measured relate only to MW.

Classification accuracy showed that single-trial classification returned relatively poor results, although significantly better than chance, indicating that real time detection of such events using only fNIRS signal would be difficult to achieve. These results, in addition to be imputable to the relatively low number of training samples, could in part be due to the high inter-subject variability observed with fNIRS signals (Jasdzewski et al. 2003). Moreover, the temporal closeness between target trials, due to their rate among all the trials (11%) may have potentially jeopardized single-trial classification. Although further investigation using a modified protocol is needed to eliminate the potential confound, the low classification accuracy we obtained does not make fNIRS a good candidate to detect

MW in real time, when used alone. Nevertheless, the significance of the classification compared to chance suggests that the fNIRS signal could complement other methodologies such as Eye Tracking or EEG to improve classification performance, and predict MW before subjects become aware of them.

5.2 Second experiment : mind wandering with an ecological task

5.2.1 Objectives of the experiment

The first experiment aimed at testing the ability of fNIRS to detect the neural activations in a context of MW. It showed that fNIRS was able to detect the DMN activations corresponding to MW occurrence, however was not efficient at predicting them by using formal classification when used alone. Furthermore, fNIRS was tested using a very controlled experiment, and still requires to be applied to more ecological settings. Therefore, the main objective was to characterize mind wandering with an experiment closer to realistic settings to test for reproducibility of the results observed in the first experiment, and to confirm them by using another measurement technique (here, we chose Eye Tracking, as it had previously been used to characterize MW (Franklin et al. 2013)). Task coding and data collection were done with the help of four ISAE students who worked under my supervision : Melanie Vadillo, Charlotte Dietrich, Romain Roche and Camille Gontier.

5.2.2 Methods

5.2.2.1 Task design

The protocol used in this study was inspired from Shaw et al. (2013), and was a modified version of the SART protocol used in the previous experiment (see section 5.1). The subjects assumed the role of Unmanned Aerial Vehicle (UAV) controllers monitoring the flight pattern of a squadron of four UAVs projected on a computer monitor (see Fig. 5.7). Critical signals were cases in which one of the UAVs was flying in an inappropriate direction relative to the others so that a collision could occur. In accordance with the SART paradigm, subjects were asked to "validate" the safe flight patterns (by pressing the spacebar when a non-target stimulus appeared) and to reject critical flight patterns (by not pressing the spacebar when a target stimulus appeared).

The display was updated 60 times per minute : a randomly chosen stimulus was displayed for $1000ms$, followed by a mask for another $1000ms$. In case of a non-target

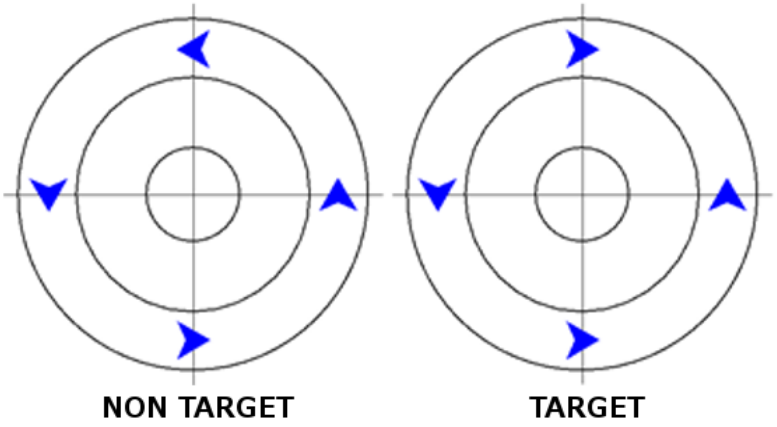


Figure 5.7 – *Examples of stimuli used in the experiment. Four blue arrows representing UAVs were drawn on the computer screen. In non-target stimuli (left), the four UAVs flew in the same direction. In target stimuli (right), one of the UAVs flew in an opposite direction.*

signal (no risk of collision), observers were given 1000ms to produce the appropriate *go* response (i.e. to press the spacebar of the computer keyboard). Stimuli consisted of three concentric circles, whose diameters respectively were 9.53cm, 6.35cm and 2.54cm. UAVs were figured by blue arrows flying counter-clockwise in the case of a non-critical target : in the case of a critical signal, one of them was flying clockwise (see Fig.5.7).

5.2.2.2 Procedure

Twenty-four participants, all students or interns at ISAE-SUPAERO gave their consent to participate in the experiment. Data from two subjects were removed due to technical problems. Final group included 4 women and 18 men, 19 right-handed and 3 left-handed. They all were ranging in age between 20 and 23 years (mean age : 21.23), and had normal or corrected-to-normal vision.

Before the experiment, the subjects were provided with instructions concerning the SART protocol and went through a training block, during which 150 stimuli were displayed, 11% of them being targets. During the training session a computerized female voice provided feedback regarding correct detections, misses, and false alarms. This training took place before the beginning of the task, in order to familiarize the subject with the interface. The experiment consisted of two blocks, lasting approximately 10 minutes each, composed of 228 stimuli, 25 (11%) of them being targets. At the end of the experiment, subjects were asked to give their impressions about the task and whether or not they had experienced task unrelated thoughts while performing it.

5.2.2.3 Data Acquisition

During the experiment, hemodynamic data from the PFC were recorded using a fNIR100 device (Biopac Inc.). The device consists of four light-emitting diodes (LED) sources of $730nm$ and $850nm$ (LED current: $12mA$), and ten detectors. The source and detectors are separated by 2.5 cm, resulting in 16 optodes uniformly placed on a rectangular grid on the forehead. Data were collected with a sampling frequency of 2 Hz. A baseline of 10 s at the beginning of the experiment was used to calibrate the device. A thick tennis band was used to cover all the sensor and forehead in order to isolate it from the environment. Hemodynamic data from one subject had to be removed because of a technical problem during the acquisition. Furthermore, the contact between the sensor and the skin for optode 1 in the DLPFC was bad for many of the subjects, so that we decided to reject it from the analysis.

Simultaneously, Eye Tracking data including gaze position and pupillometry were recorded using a SMI RED500®device fixed on the computer screen (placed approximately 70cm from the participant), which consisted of two constantly emitting infrared sources and an infrared camera. The intensity of the two light sources was attenuated by using a circular variable neutral density filter, in order to reduce its impact on the fNIRS signal without loosing Eye tracking data. Data were sampled at 60Hz.

5.2.3 Data processing

Both hemodynamic and eye tracking data were processed using EEGLAB, similarly to the previous experiment. Saccades were extracted from Eye tracking data using a velocity-based detection method (see Duchowski (2007)), and were defined as a gaze angular speed superior to $30^\circ/s$. The saccade rate was computed from a sliding window of 2 seconds.

Continuous data were first high-pass filtered using IIR filter with a cut-off frequency of 0.02Hz (for fNIRS data) and 2Hz (for Eye tracking data), to remove the trends in the signal. Then, artifacts were removed from the data by using a moving average FIR filter of 2 seconds (of order 4 for NIRS, and order 120 for Eye tracking). Finally, epochs were extracted from continuous data from $-20s$ to $+10s$ around *SART Errors* trials (i.e. when the subject inaccurately pressed the spacebar when a target stimulus appeared) and **SART No Error** trials (i.e. when the subject did not press the spacebar when a target stimulus appeared). The statistical significance of the differences between the grand-average of the two conditions was tested using permutation statistics implemented in EEGLAB, with a significance threshold of 0.05 and cluster correction for multiple comparisons (Maris and Oostenveld 2007).

5.2.4 Results

5.2.4.1 Task performance

Subjects committed on average 15.6 errors (SD=5.24) for SART target stimuli during the whole experiment (i.e. 31.2% of target trials were SART errors). Simultaneously, subjects also committed a mean of 12 (SD=5.9) errors during non-target trials (by inaccurately identifiying non target stimuli as target ones), corresponding to an error rate of 2.9% for non target trial. The mean response time was 655.3 ms (SD=66.1) for non-target trials and 644.7 ms (SD=77.8) for target trials corresponding to SART Errors. A repeated measures ANOVA revealed that this difference was not significant. At the end of the experiment, 17 subjects out of 22 stated that they were conscious of having experienced mind wandering during the task.

5.2.4.2 Hemodynamics

Three significant changes were observable in the measured variations of the HbO_2 concentration ($p < 0.05$). First, from 12s to 8s before the apparition of the stimulus, a significant activation was observed at voxels 7 and 9, part of the mPFC, for SART Errors vs. SART No Errors (see Fig.5.8). Then, from 5s before the stimulus to its apparition, we observed a significant decrease of the HbO_2 concentration in SART Error vs SART No Error conditions at all optodes in the medial inferior frontal gyrus (optodes 4, 6, 8, 10, 12, see Fig.5.9). Finally, from 3s to 8s after the arrival of the stimulus, we observed a significant increase in the same area for SART Error vs SART No Error conditions (optodes 4, 6, 8, 10, see Fig.5.10). The time course of HbO_2 concentration at optode 7 in the mPFC (where the first activation is visible) and at optode 8 in the vmPFC (where the two last significant changes are visible) is given on Fig. 5.11 and Fig. 5.12. No significant changes were observed concerning HHb level.

5.2.4.3 Eye Tracking

The evolution of pupil diameter around target trials is given in Fig. 5.13. A significant increase in pupil diameter was observed prior to *SART Error* compared to *SART No Error* trials, from approximately 13s to 8s before stimulus apparition. Immediately after stimulus apparition, a significant increase in pupil diameter in *SART Error* conditions was also observed.

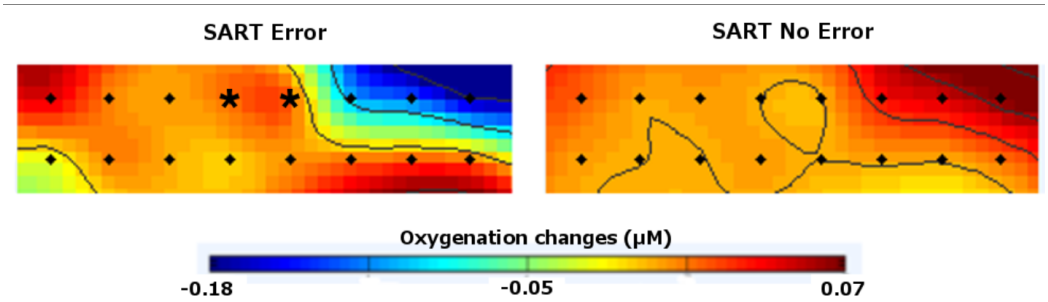


Figure 5.8 – *Topography of HbO_2 concentration over the prefrontal cortex during SART Error (on the left) and SART No Error (on the right) conditions, averaged $[-12s; -8s]$ before the apparition of the target stimulus across all subjects. The color code represents the level of HbO_2 concentration changes relative to baseline (in µM). Optodes exhibiting significant differences (all in the mPFC) are marked with a * (significance level = 0.05).*

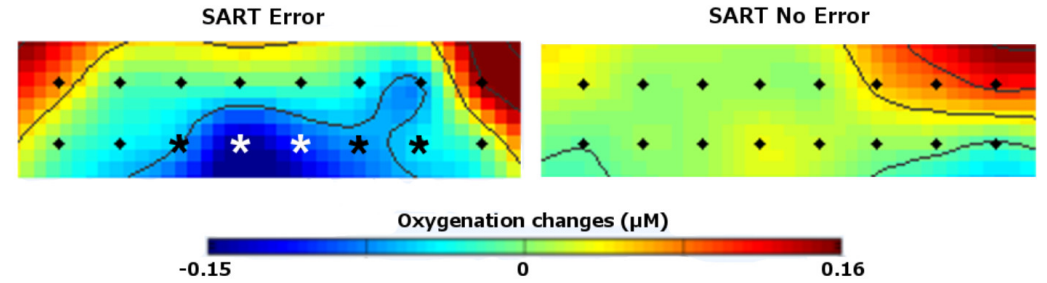


Figure 5.9 – *Topography of HbO_2 concentration over the prefrontal cortex during SART Error (on the left) and SART No Error (on the right) conditions, averaged $[-5s; 0s]$ before the apparition of the target stimulus across all subjects. The color code represents the level of HbO_2 concentration changes relative to baseline (in µM). Optodes exhibiting significant differences (all in the mPFC) are marked with a * (significance level = 0.05).*

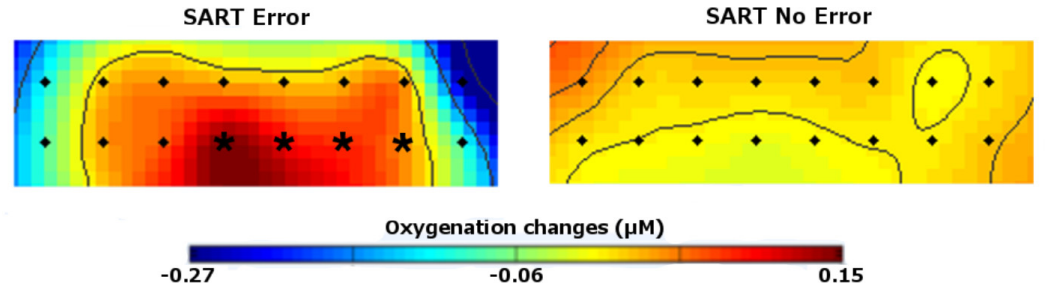


Figure 5.10 – *Topography of HbO_2 concentration over the prefrontal cortex during SART Error (on the left) and SART No Error (on the right) conditions, averaged $[+3s; +8s]$ after the apparition of the target stimulus across all subjects. The color code represents the level of HbO_2 concentration changes relative to baseline (in µM). Optodes exhibiting significant differences (all in the mPFC) are marked with a * (significance level = 0.05).*

The evolution of saccade rate around target trials is given in Fig. 5.14. Notably, a significant increase in this rate was observed immediately following *SART Errors*.

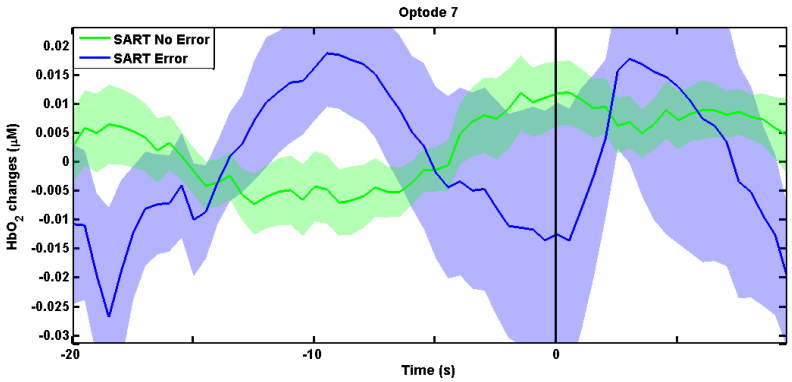


Figure 5.11 – *Variation of HbO_2 concentration (averaged trials across subjects) on optode 7 (in the mPFC) preceding SART Error (in blue) and SART No Error (in green) trials. Shaded areas represent the standard error of the mean for each condition.*

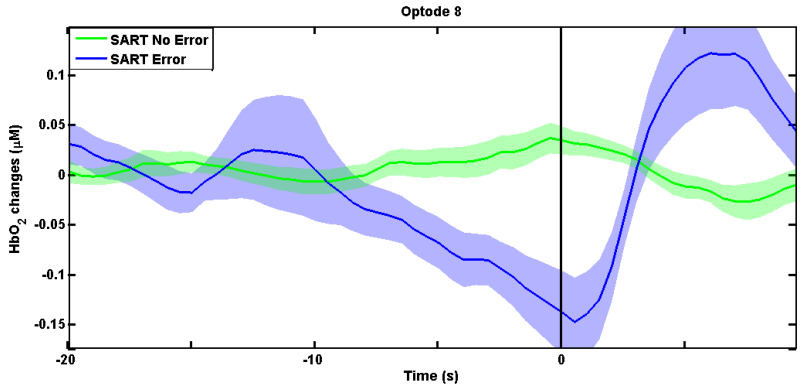


Figure 5.12 – *Variation of HbO_2 concentration (averaged trials across subjects) on optode 8 (in the vmPFC) preceding SART Error (in blue) and SART No Error (in green) trials. Shaded areas represent the standard error of the mean for each condition.*

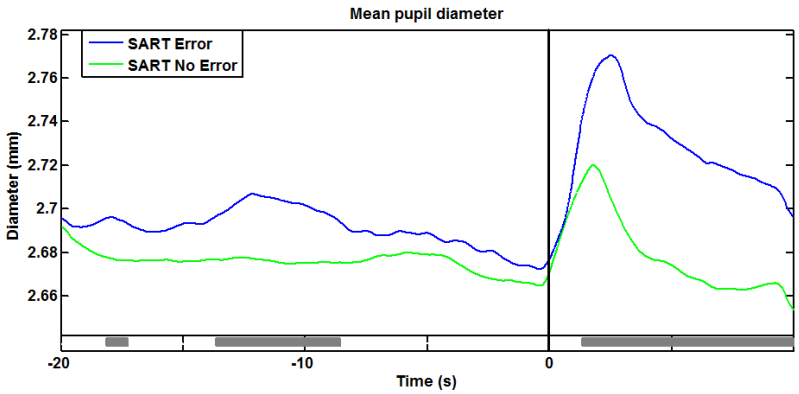


Figure 5.13 – *Variation of pupil diameter (averaged trials across subjects) around SART Error (in blue) and SART No Error (in green) trials. Significant differences between the two conditions ($\alpha = 0.05$) are indicated by the gray line at the bottom of the figure.*

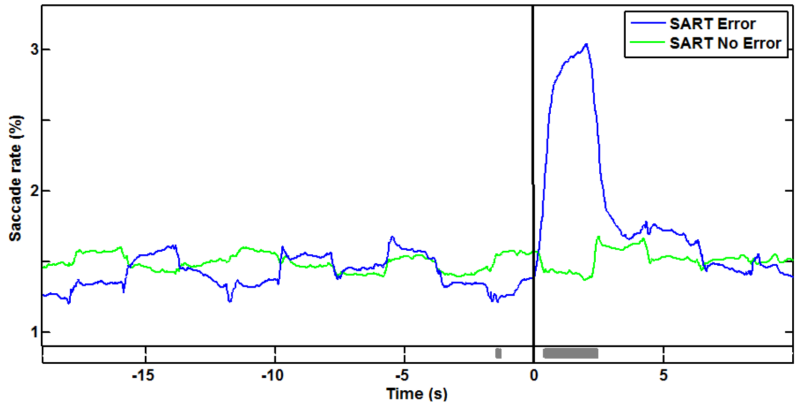


Figure 5.14 – *Variation of ocular saccade rate (averaged trials across subjects) around SART Error (in blue) and SART No Error (in green) trials. Saccade rate was computed using a velocity criterion. Significant differences between the two conditions ($\alpha = 0.05$) are indicated by the gray line at the bottom of the figure.*

5.2.5 Discussion of the current results and perspectives

The modified version of the SART protocol used in this study was efficient at inducing MW episodes (as witnessed by the high retrospective MW meta-awareness rate) and more generally executive failures (as witnessed by the high rate of SART Errors committed). Moreover, the error rate for target trials (31.2%) was similar to the one obtained in the previous experiment when using a classical SART protocol, while the performance at the identification of non target trials remained very high (98.1%). Altogether, these results suggest that the task was simple enough to generate a tendency of the participants to automate their behavior and commit executive failures.

The activation of the mPFC in the DMN preceding SART Errors was visible (see Fig. 5.8), similarly to the previous experiment, therefore reinforcing the hypothesis of its role in MW, as previously observed in various imaging studies (Christoff et al. 2009, Hasenkamp et al. 2012, Durantin et al. 2015). However, contrary to the previous experiment, these activations were followed by a significant decrease of the level of HbO_2 measured at the optodes in the medial inferior frontal gyrus (see Fig. 5.9), sustained until the arrival of the stimulus. A possible explanation of this fact is the role played by this area in the maintenance of sustained attention in a predictable task sequence (Koechlin et al. 2000) and the relative complexity of the task compared to a classical SART protocol. On this basis, the drop of the level of HbO_2 in this region would be the consequence of the disengagement from the task experienced by the subject. Another potential correlate of the disengagement experienced by the subjects during SART Errors is the high level of pupil diameter preceding SART Errors (see Fig.5.13). At first sight, this result seems inconsistent with previous studies showing that pupil diameter increases with cognitive load. However, previous studies of mind wandering including attendance to a visual stimulus have already linked the occurrence of MW with high pupil diameter (Franklin et al. 2013, Grandchamp et al. 2014), in line with the locus cœruleus theory. Accordingly,

the disengagement of the operator would be associated with tonic activations of the locus cœruleus, resulting in higher baseline levels of pupil diameter (but lower event related pupillary responses) (Gilzenrat et al. 2010, Smallwood et al. 2012b). Interestingly, there could be a link between the activations of the locus cœruleus leading the noradrenergic system responsible for vigilance and the activity measured in the prefrontal cortex (Devoto et al. 2005) related to mind wandering. However, further investigation of the signals by correlation or causality analysis is required to state on this point. Nevertheless, these results suggest a fundamental role of both the DMN and the locus cœruleus system in the generation and maintenance of a MW state at the origin of executive failures during the task.

The last point of discussion concerns the significant variations observed posterior to the mean response time (that is to say after the error took place), including activations of the vmPFC (see Fig.5.10) and increases in the ocular saccadic movements rate (see Fig. 5.14). It is thus unclear whether these hemodynamics and ocular metrics modifications could be a correlate of mind wandering or of the participant being aware of his own mistake (as the rate of retrospective meta-awareness of mind wandering was high). However, together with the variations observed prior to SART Errors, these metrics could be used to train a formal classifier. Especially, at it was noted in the previous experiment that fNIRS alone did not allow good prediction of SART Errors, the addition of ocular metrics could improve the performance obtained. This work is left for further investigation.

Finally, the same limitation as the previous experiment remains, concerning the link between SART Errors and MW, as one may argue that SART Errors refer to motor response inhibition failure rather that an attentional drift. However, in this experiment, it is noticeable that SART Errors were not associated with impulsivity in the subject's responses. Indeed, there was no difference between the response time for SART Errors and the response time for non-target trials. This fact, together with the presence of errors both for target and non-target stimuli (even if these latter errors are not numerous), suggests that the subjects were not impulsively responding to stimuli. Notwithstanding, it is also possible that some of the mistakes were attributed to wrong identification of the stimuli independent from the fact that subjects tended to automate their behavior. Thus, MW cannot account for the totality of SART Errors, although it is one of the causes of these executive failures.

The results of this study confirm the relevancy of fNIRS combined with Eye Tracking to detect DMN activations and disengagement from the task associated with MW. In particular, the dynamics associated with the switch to a MW state and the deactivation of supervisory areas following this switch were detected by using fNIRS measurements. The study also suggests the implication of the locus cœruleus noradrenergic system in the occurrence of MW and disengagement when the operator faces low demanding tasks,

characterizing the loss of attentional focus and the consideration of alternative strategies (Hartmann and Fischer 2014). Taken together, these devices may provide relevant information to dynamically assess the engagement of an operator and to predict the executive failures under low demanding conditions, in line with the neuroergonomics approach.

6 Signal Processing and Brain Computer Interface Methods

- *We applied the MACD filter used in economic market analysis to the detection of engagement related hemodynamic changes.*
- *The MACD filter led to accurate on-task vs off-task state estimation, with no need for calibration.*
- *We developed two Kalman filters, taking into account the characteristics of each participant, to improve EEG and fNIRS signals in real-time.*

The previous chapters (see 7.4 and 7.5) demonstrated the useability of brain measurement techniques, and especially fNIRS, to exhibit the physiological variations consequent of the executive failures in the presence of extreme workload (disengagement, mind wandering). In line with the neuroergonomics approach (see chapter 7.2), this chapter presents different methods developed to enhance the usability of brain measurements for BCI design (see chapter 7.3). The contribution on this point was oriented on three aspects, that will be developed in the chapter : first, the utilization of a **Moving Average Convergence Divergence** (MACD) filter for artifact rejection and idle mode detection with fNIRS ; then, the development of a **Kalman filter** to improve signal usability for

fNIRS and EEG ; finally, the development of a passive fNIRS -based BCI to monitor the level of mental workload experienced by a pilot in ecological situations.

6.1 Moving Average Convergence Divergence (MACD) filter

6.1.1 Introduction

We saw in chapters 7.2 the potential that fNIRS offers to detect the neural correlates of mental workload and engagement. In chapter 7.3, we described the growing use of fNIRS in BCI design, and the challenges still limiting its performance. In particular, we saw that signal processing (i.e., feature extraction and translation) represents a main challenge and a fundamental requirement for BCI design (McFarland et al. 2006), as the features extracted have to be available in real-time. It is especially true concerning fNIRS, whose signal processing techniques still have to be developed to reduce noise and to improve usability of the signal.

In this study, we propose a **Moving Average Convergence Divergence** (MACD) low order digital filter, as a tunable tool for real-time bandpass filtering of fNIRS signal. This filter, already employed in economic market analyses, performs a first derivative estimation of the signal and can be used to perform online detection of stimuli onsets (Appel 2005) without use of a preliminary learning phase. Notably, this type of filter was already used by Utsugi et al. (2007a) for noise reduction in previous fNIRS BCI design. However, the filter still not has been compared to classical types of filter, and still has not been used to perform peak activity onset detection on fNIRS data. We performed MACD filtering and onset detection on real fNIRS data collected during a digit sequence memorization task, and analyzed the influence of filter parameters on magnitude response and detection accuracy. The results of the study were presented in Durantin et al. (2014c).

6.1.2 MACD filter principle

The MACD filter is used in economic market analysis (Appel 2005), and is based on the principle of **Exponential Moving Averages** (EMA), described in equation 6.1.

$$y = EMA_N(x) \Leftrightarrow y_n = \frac{2}{N+1}x_n + \frac{N-1}{N+1}y_{n-1} \tag{6.1}$$

EMA filtering computes the output as a weighted average of past values of the input, using a first order Infinite Impulse Response (IIR) filter. The average weights defined by the parameter N decrease exponentially over time, so that the output value is computed

giving more weight to more recent values. As a consequence, a large value of N produces an output sensitive to slow variations of the input, whereas a small value of N results in an output sensitive to fast variations.

Then, the MACD filter is obtained by subtracting a long-term EMA filter from a short-term one (cf. equation 6.2), thus obtaining an output corresponding to instantaneous variations of the input signal.

$$MACD_{N_{short},N_{long}}(x) = EMA_{N_{short}}(x) - EMA_{N_{long}}(x) \tag{6.2}$$

A prediction technique used by economists is to associate stable increases of the stock market signal with zero-crossovers of a histogram line computed from MACD. The histogram line is usually obtained by subtracting an EMA filtering with $N = 10$ of MACD data (called "*signal line*") (Appel 2005) from the unfiltered MACD data, as described in equation 6.3. Thus, a zero-crossover of the histogram line corresponds to the moments when MACD is increasing and crosses the signal line. An example of stock market analysis performed using MACD technique is presented in Fig.6.1.

$$h(x) = MACD(x) - EMA_{10}(MACD(x)) \tag{6.3}$$

An interesting property of the MACD is its capacity to detect stable increases or decreases of a signal, making it a good candidate for hemodynamic response onset detection. In economic market analysis, the parameters used for MACD filtering are commonly fixed to $N_{short} = 12$ and $N_{long} = 26$ (empirical values). However, if we want to apply this filter for physiological signals (here, fNIRS), we have to carry out an optimization procedure to determine the adapted parameters, and to test its performance on experimental data.

6.1.3 Experimental Methods

Nine participants (Mean age = 24; SD = 3.6 ; eight males, eight right handed) participated in the study. The volunteers performed a digit sequence memorization task, while NIRS measurements of the prefrontal cortex were recorded.

Each trial of the experiment consisted in the presentation, on a computer screen, of a series of 5 to 10 randomly chosen digits that the subjects had to memorize. The size of the series defined a level of difficulty. The digits appeared in white at the center of a black screen for $600ms$, then the screen turned black for $300ms$ until the next digit presentation. After the last digit presentation, a fixation cross appeared. The subjects

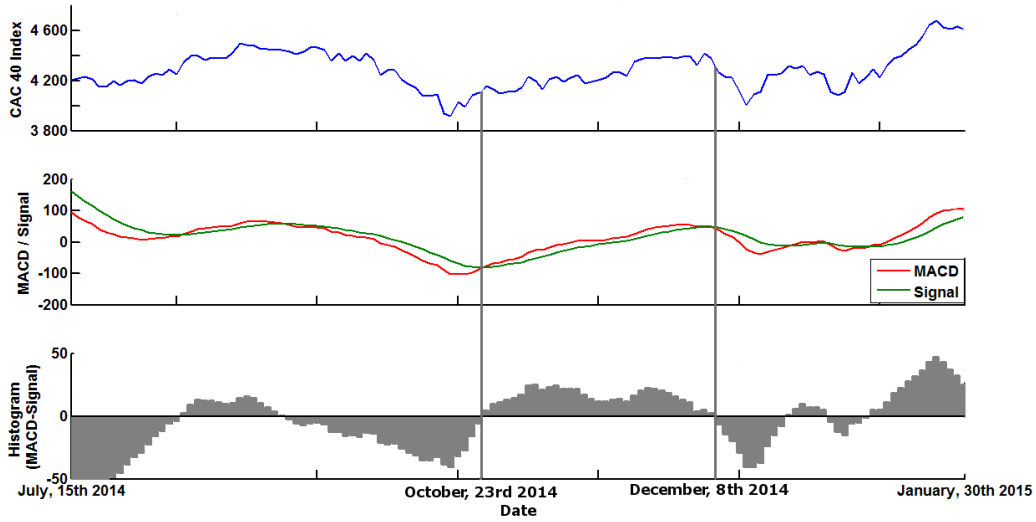


Figure 6.1 – *Example of MACD applied on the* ***CAC40 stock market index*** *(top) from July,15th 2014 to January, 30th 2015. We compute the MACD line (middle, red) by two EMA (N=12 days and N=26 days), and the signal line (middle, green) by an EMA applied on MACD data (N=9 days). The crossovers between those two lines (whose difference is represented by the histogram on the bottom graph) predict an acceleration of the stock in the direction of the crossover, which translate into a recommendation to buy or to sell. On the graph, two recommendations provided by the MACD are given as examples : a recommendation to buy on October, 23rd 2014 (trend to increase) and a recommendation to sell on December, 8th 2014 (trend to decrease).*

then had 8*s* to type the memorized sequence on the keyboard. Between two consecutive trials, the subjects looked passively at the fixation cross for 6 to 9*s* (the inter-trial interval was chosen randomly to avoid task periodicity). Fig.6.2 summarizes the time sequence of a trial. The experiment consisted of 24 trials (four trials for each of the six levels of difficulty), presented in a random order.

6.1.3.1 Data Acquisition

During the experiment, hemodynamics data of the prefrontal cortex were recorded using a fNIR100 (Biopac®) device with 16 optodes regularly placed on the forehead, and a sampling frequency of 2Hz. We calibrated the device using a baseline of ten seconds at the beginning of the experiment, during a rest period. The relative concentration in oxygenated hemoglobin [HbO_2] measured across optodes was averaged and used as a prediction signal.

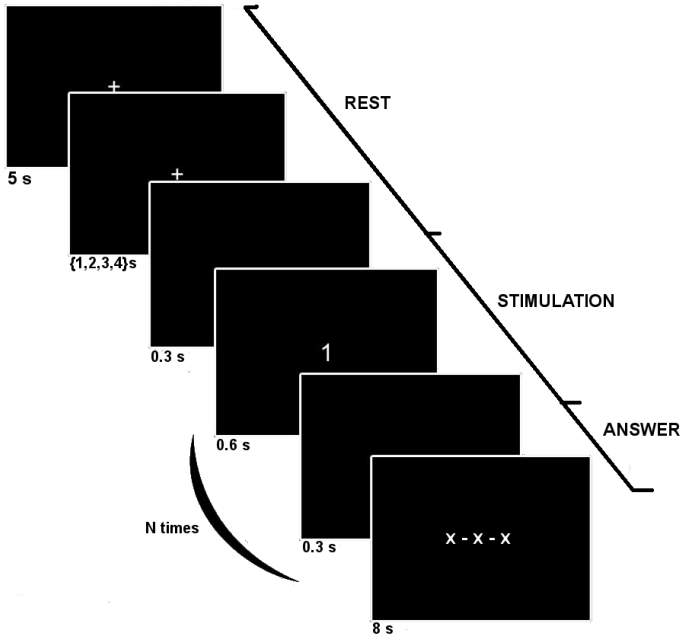


Figure 6.2 – *Time course of one trial of the experiment. The experiment consisted of a total of 24 trials (four for each level of difficulty N = digit sequence size from 5 to 10), presented in a random order.*

6.1.3.2 MACD filtering module

In this experiment, data from each participant were filtered using MACD filters with parameters (N_{short}, N_{long}) in the interval $([1;20], [N_{short}+1; N_{short}+20])$, in order to remove low frequency and high frequency components. An example of filtered signal is shown on Fig.6.3. In order to estimate MACD filtering power, we analyzed its magnitude and phase responses, compared with three other types of filters : Finite Impulse Response (FIR), Infinite Impulse Response (IIR) Butterworth, and Elliptic, all of order 2. The bandwidth of these filters was equivalent to the MACD one ([0.02 ; 0.3]). We also computed the correlation between the filtered signal and the first derivative of the raw signal.

6.1.4 Cerebral hemodynamics

We filtered raw data using one specific MACD filter, whose parameters were chosen accordingly to hemodynamic response characteristics ($N_{short} = 12 datapoints = 6s$ for the peak response,$N_{long} = 30 datapoints = 15s$ for the return to baseline) (Jasdzewski et al. 2003). For each trial, the mean HbO_2 concentration over the 16 seconds following a stimulus onset was computed. We then averaged data across trials of the same size, and tested the effect of the sequence size using Friedman ANOVA.

6.1.4.1 Onset prediction

We applied the MACD-based onset prediction technique described in part 6.1.2, using a histogram line obtained by subtracting to MACD data a signal line computed with $N = 10$. Thus, a zero-crossover of the histogram line corresponds to the moments when MACD is increasing and crosses the signal line. An example of histogram is presented in Fig.6.3. We used this technique to predict trial onsets as $\{t|h(t) = 0, h(t-1) < 0, h(t+1) > 0\}$.

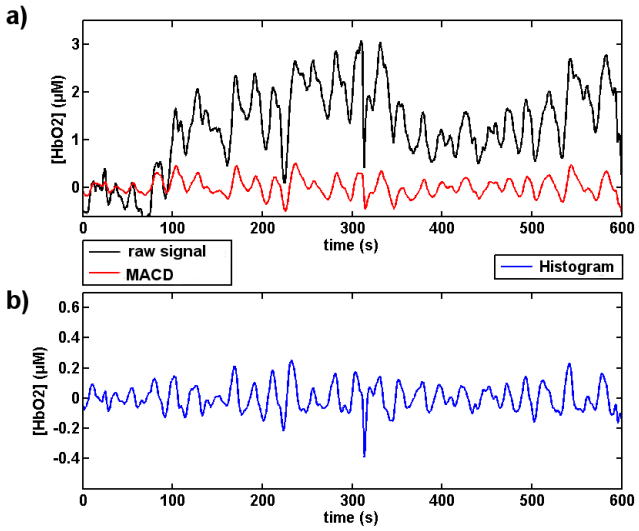


Figure 6.3 – *Filtering module output for participant 1. a) Raw (in black) and MACD-filtered (in red) signals, obtained with MACD parameters* $N_{short} = 12$ *and* $N_{long} = 30$. *b) Histogram obtained with* $N_{signal} = 10$.

An onset detection was considered a *True Positive* when the real onset (at the beginning of the stimulation period) occured in the range $[t_{detection} - 7sec; t_{detection} + 3sec]$, to account for hemodynamic response delays and stimulation duration(Aguirre et al. 1998, Jasdzewski et al. 2003). Otherwise, the onset detection was considered as a *False Positive*. An example of onset detection is shown on Fig.6.4.

We computed the Correct Detection Rate (CD Rate) as the percentage of real onsets effectively leading to an onset detection. As our objective was to jointly maximize the CD Rate and to minimize the False Positive Rate (FP Rate), we computed the corresponding ratio (i.e. Performance Ratio) to be maximized.

We performed the same analysis on onset predictions obtained by applying a threshold ($[HbO_2] = 0.2\mu M$) on data filtered by IIR Butterworth filter (which is a widely used filter in signal processing) of order 2, to compare the performance of MACD onset detection

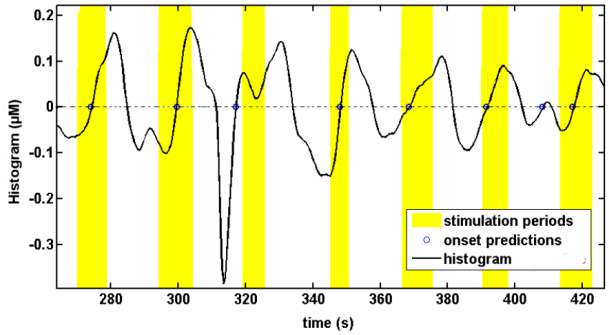


Figure 6.4 – *Onset detections obtained from seven trials of participant 1, using MACD filtering with* $N_{short} = 12$ *and* $N_{long} = 30$*. The stimulation period length, in yellow, depends on the length of the digit sequence size.*

method. In this method, onset prediction was associated with crossover between filtered fNIRS signal and a threshold.

6.1.5 Results

6.1.5.1 Filtering power

The magnitude response of MACD, in comparison with FIR, Butterworth, and Elliptic filters, is presented in Fig.6.5. MACD has same bandpass behaviour as other filter types, although all frequency components are attenuated to at least 8dB, leading to a smaller signal amplitude than the one which would be obtained with "classical" filtering. However, this attenuation does not impede noise rejection, especially concerning low frequencies ($< 0.02Hz$). The phase response of MACD filter is comparable to IIR structure. Notably, it has a quasi-linear phase in its bandwidth (an important property for physiological signal processing).

MACD implements bandpass filtering, whose bandwidth and cutoff frequencies depend on the difference between parameters N_{long} and N_{short}, as a higher gap between those parameters brings a narrower bandpass. Their evolution is given on Fig.6.6. Furthermore, the MACD filtered signal is highly correlated with the first derivative of raw signal ($r_{Pearson} = 0.63$; $SD = 0.06$), thus MACD can be considered as a first derivative estimate.

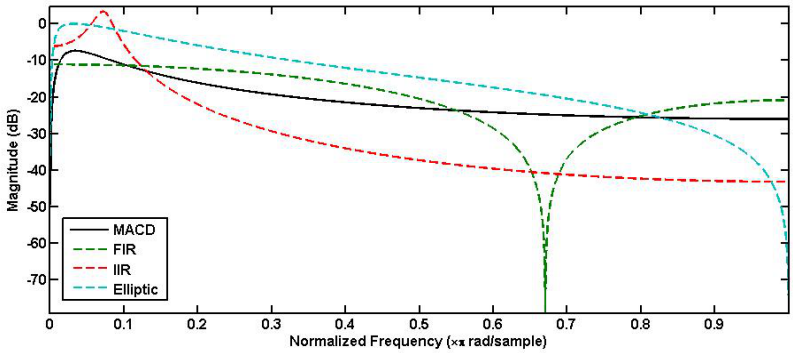


Figure 6.5 – *Magnitude response of MACD filter (in black) with* $N_{short} = 12$ *and* $N_{long} = 30$*, compared with FIR filter (in green), IIR Butterworth filter (in red), or Elliptic filter (in blue).*

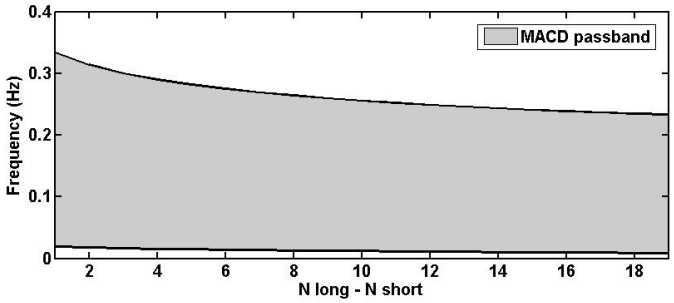


Figure 6.6 – *MACD bandpass in function of the difference* $N_{long} - N_{short}$*.*

6.1.6 Cerebral hemodynamics

Friedman ANOVA run on data filtered by the specific MACD filter ($N_{short} = 12$; $N_{long} = 30$) revealed a significant main effect of sequence size on HbO_2 ($p = 0.001$). The results, shown on Fig.6.7, illustrate the increase of HbO_2 with greater sequence size. Post-hoc Wilcoxon sign-rank test showed significant differences between size 8 and sizes 6 and 9 ; between size 9 and sizes 5, 6, 7 and 8 ; and between size 10 and size 6.

6.1.6.1 Onset prediction and MACD parameters

We observed the variations of the mean CD Rate across subjects, with different values of MACD filtering module parameters. We also computed the ratio between the CD Rate and the FP Rate across subjects. The results are shown on Fig.6.8. A decrease in the value of $N_{long} - N_{short}$ triggers an increase in the CD Rate and also leads to a reduction in the Performance Ratio, as shown in Fig.6.8. However, the Performance Ratio reaches local maximum values for MACD tuning parameters on a curve $\mathcal{M}_k = \{(N_{short}, N_{long}) | N_{long} = \frac{2}{3}N_{short} + 10k + 2\}$ (graphically obtained from Fig.6.8). Table 6.1 sums up the average

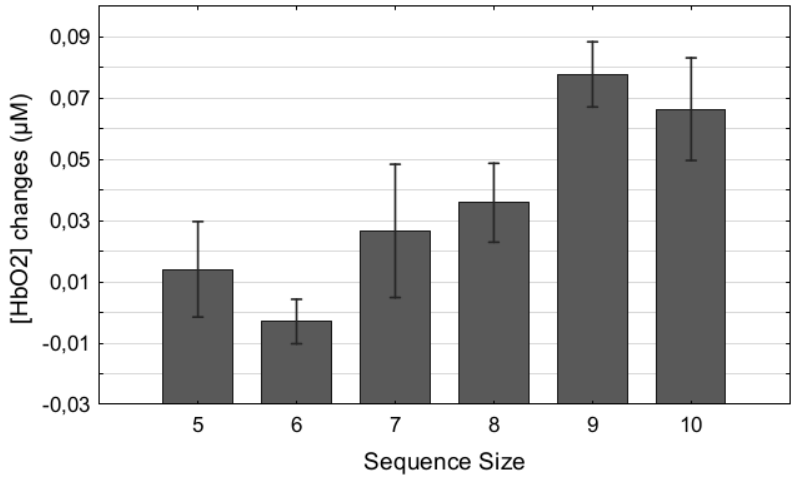


Figure 6.7 – *Mean HbO_2 changes in function of sequence size. The error bars represent the standard error of the mean.*

onset prediction performances measured on these curves, showing the trade off between the Detection Rate and the Performance Ratio.

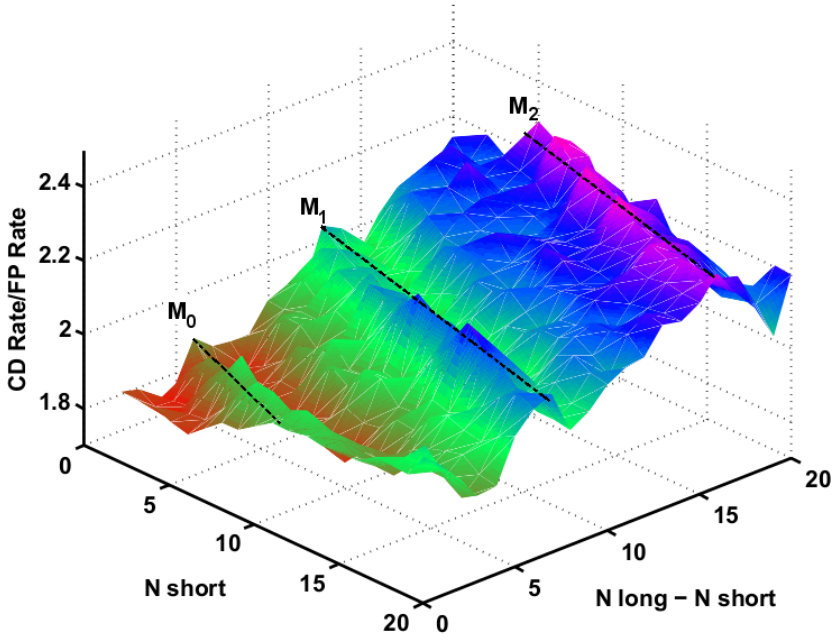


Figure 6.8 – *Performance Ratio (Mean Correct Detection Rate divided by False Positive Rate) across subjects, in function of MACD parameters N_{short} and N_{long}. Local maxima curves $\mathcal{M}_i$ are indicated in dashed lines.*

The 95% confidence intervals were computed using a bias corrected and accelerated percentile method with *bootci* function of Matlab©. Results show that the CD Rate is significantly better when using MACD rather than IIR. In particular, MACD parameters

Table 6.1 – *Mean onset prediction performance over local maxima curves $\mathcal{M}_1$,$\mathcal{M}_2$ and $\mathcal{M}_3$ (and 95% confidence intervals), compared with a thresholding method using IIR filter*

	CD Rate (%)	FP Rate (%)	Performance Ratio (arb. units)
$\mathcal{M}_0$	0.99 [0.98 ; 1]	0.54 [0.52 ; 0.57]	1.83 [1.75 ; 1.92]
$\mathcal{M}_1$	0.88 [0.83 ; 0.91]	0.42 [0.35 ; 0.47]	2.30 [1.93 ; 3.34]
$\mathcal{M}_2$	0.84 [0.79 ; 0.89]	0.39 [0.32 ; 0.44]	2.39 [1.99 ; 3.25]
IIR	0.51 [0.28 ; 0.65]	0.42 [0.30 ; 0.64]	1.76 [1.01 ; 2.38]

on curves $\mathcal{M}_1$ and $\mathcal{M}_2$ give significantly better CD Rate, FP Rate and Performance Ratio than curve $\mathcal{M}_0$. Notably, the confidence intervals of the performance obtained using MACD technique are lower than the ones obtained with IIR.

6.1.7 Discussion and conclusion

The objective of this paper was to test the usability of the MACD filter as an online noise reduction and analysis technique for fNIRS signal. On this purpose, we conducted an experiment on nine subjects involving a memory task. Overall results revealed the efficiency of this filter, and in particular, MACD has the same filtering behavior as IIR, FIR, and Elliptic filters of the same order (Fig.6.5), and especially better reject low frequency components ($< 0.02Hz$). These results confirm the relevance of this method for noise reduction, as it has already been done in previous BCI design (Utsugi et al. 2007b). In particular, the low order of MACD filter enables utilization with few latency even when the sampling frequency is low.

Our statistical results based on data filtered with MACD revealed an effect of the task difficulty on HbO_2 level. This result is consistent with previous studies with fNIRS (Hoshi et al. 2003), and suggests that fNIRS exhibited significant changes in response to stimuli presentation. The real-time detection of stimuli onsets using MACD shows good and homogeneous results (Table 6.1), compared to classical method based on IIR filtering. This good performance is similar to previous studies investigating the suitability of first derivative estimate of hemodynamics data for detection of event-related hemodynamic activity (Friston et al. 1998, Tomita and Mitsukura 2013). In particular, the lower variance across subjects when using MACD rather than IIR filtering suggests that this method could bring good results on a greater number of subjects, therefore reducing the BCI illiteracy problem (see chapter 7.3). Interestingly, the optimal parameters for real-time detection using MACD (Fig.6.8) are coherent with physiological time characteristics of

hemodynamic response, i.e. time-to-peak and return to baseline time (Aguirre et al. 1998, Jasdzewski et al. 2003). This result suggests that the tuning of the MACD filter could be optimized independently for each participant to yield better results.

Nevertheless, a challenge remains in decreasing the false positive rate of onset detection. One possibility is to test the influence of the window length used for histogram line calculation (set as 10 in this study, which is mainly used (Appel 2005)). This possible investigation is a direction for future work. Additionally, a reduction in the false positive rate could be obtained by using a multimodal approach (Fazli et al. 2012b), i.e. by combining fNIRS with other devices such as EEG (Pfurtscheller et al. 2010b).

Notably, the MACD technique could be applied to every physiological signal that has an event-related peak activity, such as fMRI BOLD response (Friston et al. 1998) or Electrocardiography (ECG) QRS wave detection (Xiaomeng 2011). Eventually, the use of MACD as a real-time onset detection technique could open up new perspectives of real-time improvement of event-related response without a priori information on the onset, based on the use of impulse response models (Buxton et al. 2004a) and Kalman filtering (Georgiadis et al. 2005b). Although further investigation is still needed to test the detection technique in more ecological situations (especially involving multitasking), MACD appears to be an easy to implement technique, suitable for real-time noise reduction and onset detection without learning, supporting the potential usage of fNIRS for BCI design. In particular, this technique would be useful for idle modes detection (see chapter 7.3), i.e. when the user does not request any command.

6.2 Application of the MACD filter to a BCI in a flight simulator

In addition to investigate the potential of fNIRS to detect the correlates of mental workload and engagement, and to develop signal processing techniques for this technique, this thesis project also included the application of these techniques to BCI design. In collaboration with Thibault Gateau, Sebastien Scannella and Francois Lancelot from ISAE laboratory, and in interaction with project MAIA ("Modelisation de l'Attention pour une Interaction Adaptative", i.e. Attention modeling for adaptive interaction), we implemented a passive BCI in a flight simulator to demonstrate the potential of fNIRS. My contribution to this project, in addition to task design and data collection, mainly concerned the implementation of MACD filter (see section 6.2.2.4) to detect on-task periods. However the experiment also included a formal classifier to detect high workload levels. The results of this study were published in Gateau et al. (2015).

6.2.1 Objectives of the study

The aim of this study was to apply fNIRS and MACD (which were previously investigated) to design an on-line inference system dedicated to:

- estimating the pilot's state (performing or not performing a WM task);
- assessing the WM load level.

To this end, we designed a simplified but plausible pilot-ATC interaction task, using prerecorded messages. The implementation of this inference system was challenging as the participants were placed in an immersive flight simulator (*i.e.* user interfaces, panoramic external "view") that induces additional cognitive activity (*i.e.* flight trajectory monitoring), as well as motion artifacts (*i.e.* programming autopilot device) therefore adding noise to the fNIRS signal.

To meet these goals, we measured changes in the oxygenation of the prefrontal cortex including the dorsolateral prefrontal cortex (DLPFC) which is known to be involved in WM (as we saw in Durantin et al. (2014a) and chapter 7.4). Real-time pilot's state estimation was performed using MACD as proposed in economic market analysis (see section 6.2.2.4) and did not require machine learning techniques. We implemented an on-line single trial classifier (Tai and Chau 2009, Schudlo and Chau 2014) to discriminate low WM load achieved trials versus high WM load achieved trials.

6.2.2 Methods

6.2.2.1 Participants

Nineteen visual flight rules (VFR) pilots (6 women; mean group age: 27.4 $\pm$ 6.4; mean flight hours 145 $\pm$45) completed the experiment. Pilots had normal or corrected-to-normal vision, normal hearing, and no psychiatric disorders. They all had medical clearance to fly. After providing written informed consent, they were instructed to complete task training. Typical total duration of a subject's session (informed consent approval, practice task, and real task) was about two hours. This work was approved by the Inserm Committee of Ethics Evaluation (Comité d'Evaluation Ethique de l'Inserm).

6.2.2.2 Equipment

fNIRS Equipment

During each experiment, we recorded hemodynamics of the prefrontal cortex using the functional near-infrared spectrometer fNIR100 (Biopac®) composed of 16 channels (see Fig.6.9). On this continuous-wave system, the optode separation was about 25 mm and two wavelengths were used, 730 nm and 850 nm.

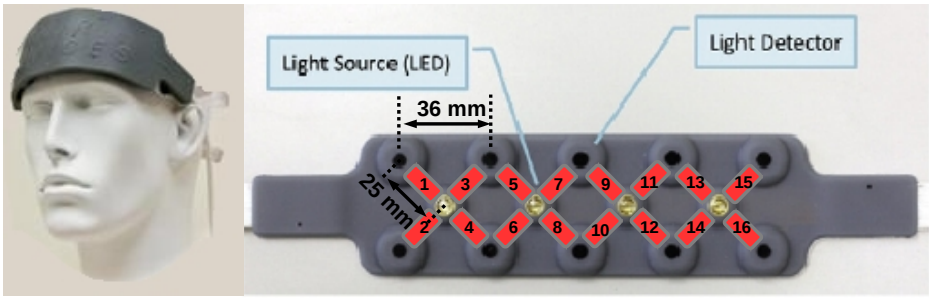


Figure 6.9 – **fNIR100®headband and associated channels numbering.** *Only the four closest detectors to an emitter constituted channels. The emitter-detector distance is 25 mm. Channels are represented in red with their associated number. The original image comes from the fNIRSOFT®manual and has been slightly modified.*

Each channel of the device records hemodynamics at a frequency of 2Hz in term of oxygenation level variations in comparison to a baseline. Changes in the concentrations of oxygenated (HbO_2) and deoxygenated hemoglobin (hHb) can be calculated from changes in detected light intensity using the modified Beer-Lambert Law (see section 2.3.4.2)

Cognitive Optical Brain Imaging (COBI) Studio®software (Ayaz and Onaral 2005) was used to collect data. The data stream was available on-line from a TCP/IP interface. Before recording, signals for each channel was carefully checked for saturation with COBI Studio which provides signal quality visual representation.

COBI studio was also used to check signal quality and to adjust consequently the headband on the participant's forehead. Channels 8 and 10, located above the nasal sinus were systematically removed because of saturation.

Flight Simulator

We used the ISAE (Institut Supérieur de l'Aéronautique et de l'Espace - French Aeronautical University in Toulouse, France) flight simulator to conduct the experiment in an

ecological situation. It simulates a twin-engine aircraft flight model and the user interface is composed of a Primary Flight Display (PFD), a Navigation Display, and an upper Electronic Central Aircraft Monitoring Display page. The pilot has a Flight Control Unit (FCU) to interact with the autopilot (see Fig.6.10).

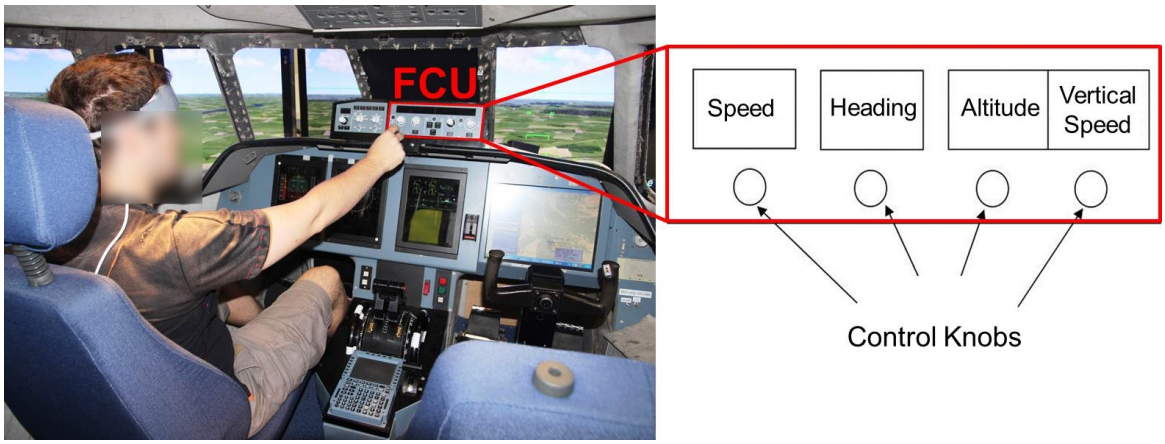


Figure 6.10 – **Pilot's interaction with the FCU.** *The participants controlled the flight simulator from the pilot's seat. The red rectangle corresponds to the FCU used to set the autopilot with four control knobs, according to ATC clearances (speed, heading, altitude, and vertical speed selection).*

6.2.2.3 Protocol

Task Description

Similar to true flying conditions, pilots heard ATC messages (pre-recorded for this experiment) and were asked to dial the corresponding flight parameters in the autoflight system using the four knobs (*i.e.* speed, heading, altitude, and vertical speed) of the FCU. The ATC messages were delivered at 78 dB through a Sennheiser®headset. Two levels of difficulty were defined based on the flight parameters that the participant had to set during the experiment:

- *Low WM load*: only one major digit per trial was used to set each flight parameter (e.g: 15 for "speed 150, heading 150, altitude 1500, vertical speed +1500").
- *High WM load*: each flight parameter value was different from the previous one and composed of different digits to increase the complexity (e.g: "speed 164, heading 235, altitude 8700, vertical speed -1600").

The task consisted of 20 repetitions of each difficulty for a total of 40 trials. The task difficulty order was randomly distributed with two constraints:

- the first 20 trials contained 10 trials of high difficulty, and 10 trials of low difficulty (which was necessary for machine learning purposes, see 6.2.2.3);
- the difficulty cannot be the same for more than two successive trials.

Each ATC message started with the airplane call sign (i.e. "Supaero 32"), immediately followed by a sequence of flight parameters and ended with the message "over". Pilots were instructed to set the parameters only after they heard the "over" message (see Fig.6.11). Thereafter, pilots had to dial the parameters on the autopilot interface during a 18 *s* response window. A practice session was conducted prior to the experiment runs to familiarize them with the experiment protocol and the interface.

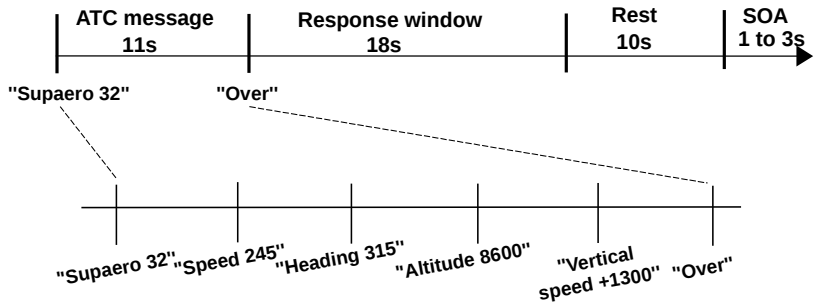


Figure 6.11 – **ATC span task trial design.**

Experimental Components' Architecture

We implemented a WM load estimator that integrated different components (see Fig.6.12):

- a **simulated ATC** which broadcasts a list of chosen messages to the pilot;
- the ISAE **flight simulator** which allows a pilot to be in an ecological flight condition;
- a **fNIR100 sensor** which measures the prefrontal oxygenation on 16 channels;
- a **MACD filter** for artifact removal;
- a **synchronization module** that also formats filtered data for the classification process: filtered fNIRS output must be synchronized with the pilot's state, according to the instant of the arrival of that incoming message and according to the pilot's response window;

- a **state estimator** which evaluates pilot's instantaneous current state in real-time, which can be off task or on task. The pilot is considered on task during ATC message reception, and off task in the other periods.
- a **classifier** which evaluates in real-time whether the last ATC instruction was a *high WM load* trial or a *low WM load* trial.

Task monitoring, data acquisition and computation were conducted on the same computer (core i5-3210M, 2.50GHz, 4GB RAM).

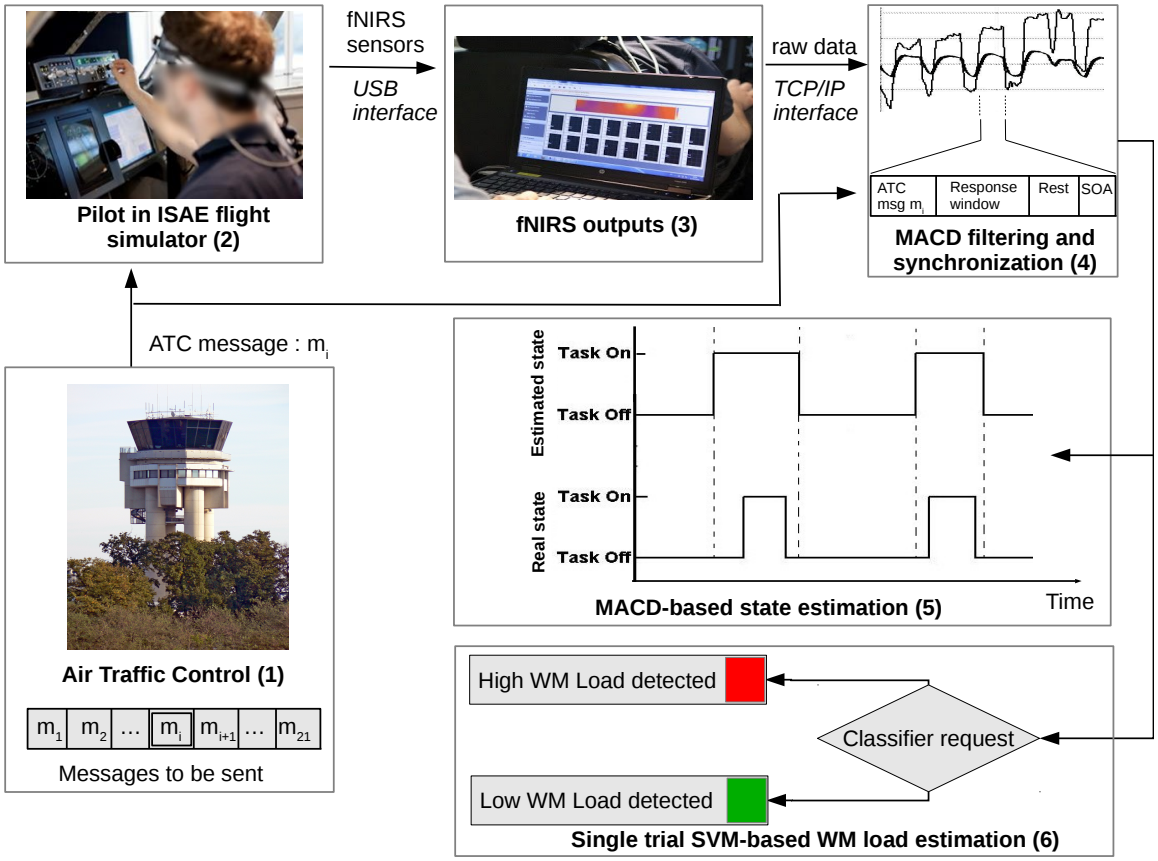


Figure 6.12 – **Illustration of the fNIRS based inference system.** *Pre-recorded ATC messages were sent to the pilot (1). The pilot's prefrontal activity was measured with a fNIRS device (2). Output measures (3) were MACD-filtered and synchronized with the temporal design of the trial (4). During the entire session, the MACD-based state estimator detected whether the pilot's state was on task or off task (5). When all of the required data were available for the trial, a request was sent to the pilot's classifier to assess the WM load of the trial (6).*

Experimental Time Course

For machine learning purposes, the experiment was split into three successive phases (see Fig.6.13):

- Phase D – data gathering phase: 20 instructions with two levels of difficulty were successively presented to the pilot in a random order. During phase D, the correctness of the pilot's response was also checked for further pilot performance analysis. Entered FCU parameters were available through the ISAE flight simulator software. The fNIRS's data were processed and recorded for each trial's response window. The levels of difficulty of the message were also recorded.
- Phase L – classifier training phase: the classifier training process was activated, based on the data gathered during phase D. This phase was not perceived by the pilot and allowed further classification actions. At the end of this phase, the pilot's classifier – the pilot's specific classification model, correctly trained – was available for classification requests.
- Phase T – classifier testing phase: 20 instructions with random levels of difficulty (*high WM load* or *low WM load*) were successively presented. The aim of the classification process was to discriminate the difficulty of the trial, as soon as possible. After each response window of trials, the classifier returned WM load estimation of the trial.

The transition from phase D to phase T was transparent to the participants.

6.2.2.4 Data Analysis

MACD Filter

Raw fNIRS data were real-time filtered using a MACD filter, commonly used in economic market analysis Appel (2005). This filter, based on the difference between a short-term EMA and a long-term EMA, implements a second order band-pass filtering to eliminate low-frequency (<0.02Hz) and high-frequency (>0.33Hz) components from the raw fNIRS signal Utsugi et al. (2007b). This low order filter has a quasi linear phase in its bandwidth and is particularly suited for real-time applications. For the experiment, we proceeded to an on-line filtering of HbO_2 and HHb levels on 14 of the 16 channels (Channels 8 and 10 were excluded due to several artifact acquisition issues), as described in Equation 6.1, where N represents the number of time points defining the EMA window (see section 6.2.2.4). We chose a 6 s short-term EMA ($N_{short} = 12$) and a 13 s long-term EMA ($N_{long} =$

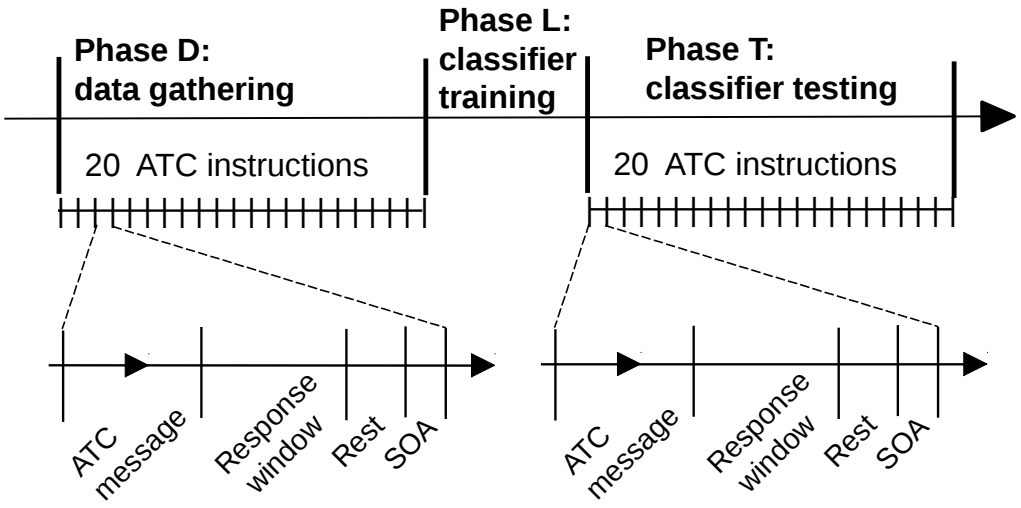


Figure 6.13 – **The experiment was split into three successive phases.** *Data gathering (phase D) and classifier testing (phase T) consisted of 20 ATC instructions each. The pilot's classifier was trained between these two phases (phase L). The time scale of the figure is illustrative.*

26), according to previous work on MACD filtering, to get the desired bandwidth (see section 6.2.2.4).

MACD-based State Estimation

We performed MACD analysis to estimate the participant's instantaneous mental state, on task versus off task, in real-time. In economic market analysis, Appel (2005) states that a sustainable increase in the signal can be predicted when the MACD line crosses the signal from below. On the contrary, a sustainable decrease in the signal can be predicted when the MACD line crosses the signal line from above. This method can help estimate task-onsets and task-offsets based on the fNIRS signal (Durantin et al. 2014c). We computed in real-time a state estimation chronogram by associating the moments when the MACD line crossed the signal line from below with stimuli onsets. To do this, we averaged MACD-filtered fNIRS data over the 14 channels. A *signal line* was computed using a 5 s EMA ($N = 10$) of this data, as described in Equation 6.3:

Similarly, we defined task-offsets when the MACD line crossed the signal line from above (for an example, see Figure 6.14). We then compared each time point of the state estimation chronogram to the actual task-onsets chronogram in order to estimate the accuracy of this method. We labeled the state estimation at time t as :

- *Correct Estimation*, if the estimated state and the actual task state matched;
- *False Positive*, if the state was estimated as on task, and actual state was off task;

- *False Negative*, if the state was estimated as off task, and actual state was on task;

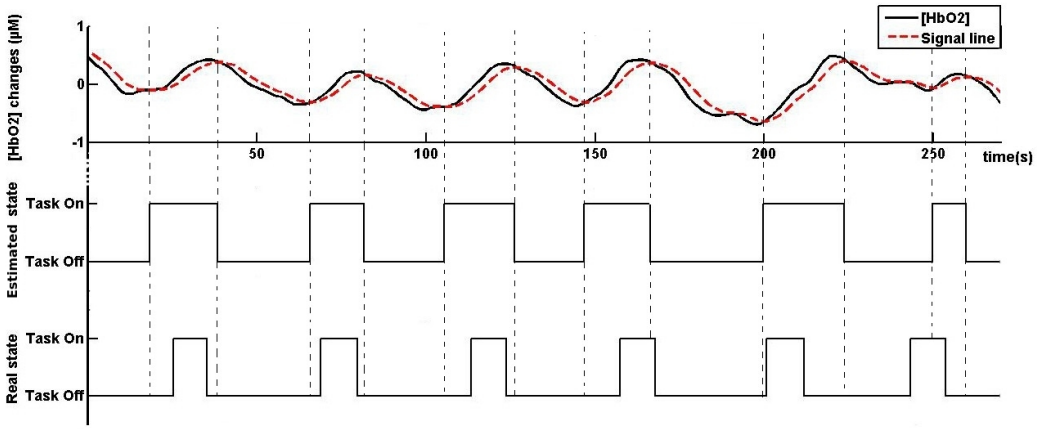


Figure 6.14 – **Example of real-time state estimation (performed on pilot 16).** *The upper graph shows MACD-filtered fNIRS signal and the signal line computed from the latter (dashed line). The two lower graphs show the participant's state estimated from crossovers between MACD and signal lines and the operator's actual state, respectively.*

Single Trial SVM-based WM Load Estimation

The classification's goal was to discriminate on-line whether the last trial was a *high WM load* trial or a *low WM load* trial. For each pilot, we used the first 20 trials to train the pilot's classifier (phase D and L, see section 6.2.2.3). From trial 21 to 40, we used the pilot's classifier to discriminate trial difficulty, without any further training. The classification was made based on a large variety of features representing various response characteristics (mean amplitude, kurtosis, skewness) for both HbO_2 and HHb, and various response windows (see Gateau et al. (2015) for details). The features used for each of the participants were optimized during the learning phase, and the classification was done using a linear Support Vector Machine (SVM).

Behavioral and Offline fNIRS Data

We performed classical off-line behavioral analysis to ensure that we correctly manipulated WM load (*i.e.* increased error rates in high WM load condition) and that the participants' performance was identical across the two blocks (*i.e.* identical error rate during the first 20 trials and the last 20 trials). A two-way Analysis of Variance (ANOVA) was carried out on the correct response rate between subject factors Phase (Learning vs. Test) and Position (Speed vs. Heading vs. Altitude vs. Vertical speed). An off-line analysis on the neurophysiological data to verify the consistency of prefrontal activation with ex-

isting neuroimaging literature was performed. To do so, we computed the frontal HbO_2 and HHb peak response (peak value within 30s post-trial onset minus 2s average pre-trial onset) for each trial and each pilot using the MACD-filtered data. We then performed a three-way ANOVA using within subject factors Oxygenation (HbO_2 vs. HHb), WM Load (High vs. Low) and Voxels (1 vs. 2... vs. 16), excluding voxels 8 and 10 due to several acquisition artifact issues. Tukey's HSD post-hoc tests were used to evaluate all behavioral and hemodynamic interaction effects.

6.2.3 Results

6.2.3.1 Behavioral and Physiological Results

The participants committed a mean of 13.2 errors ($SD = 4.7$) during the entire experiment, all occurring during the high load trials. All the subjects completed the low WM load trials correctly. There was no significant effect of the phase (learning or testing) on the number of committed errors.

The ANOVA over the fNIRS data revealed a main effect of the oxygenation ($F(1,18) = 95.2$; $p < 0.001$; partial $\eta^2 = 0.90$) with higher HbO_2 than HHb levels and a main effect of the load ($F(1,18) = 7.3$; $p < 0.05$; partial $\eta^2 = 0.29$) corresponding to higher peak response within the high load condition. In addition, a significant interaction effect between load and oxygenation was found ($F(1,18) = 28.7$; $p < 0.001$; partial η^2=0.61) showing that the load effect was only present for HbO_2 ($p < 0.001$). Finally, a second order interaction effect revealed that the load effect was not homogeneous across voxels ($F(13,221) = 2.87$; $p < 0.001$; partial η^2=0.14). Post-hoc tests revealed a maximum load effect within the right DLPFC (i.e, voxel 15) forHbO_2.

6.2.3.2 MACD-based State Estimation Results

Accuracy

The real-time estimated state was compared to the actual state of the subject during the experiment (on task or off task). The results show that the estimation matched 61.74% of the time ($SD = 4.27\%$), which was significantly better than chance ($F(1,18) = 145.52$; $p < 0.01$) (see Fig.6.14). We obtained a 58.24% mean specificity (SD=3.80%), and a 71.88% mean sensitivity(SD=9.34%).

Latencies

The required time for data filtering (maximum per sample $< 0.4\ ms$) is negligible regarding fNIRS time resolution (2Hz). Hence, pilot's state estimation (off task or on task) is available in real-time.

Off-line Analysis

The results of estimated onsets and offsets latencies compared to stimuli onsets and offsets are summarized in Figure 6.15. On average, the onset of state estimation significantly occurred $1.97s$ before the actual state onset ($SE = 0.34s$). The estimated offset occurred $2.43s$ after the stimulus offset ($SE = 0.52s$).

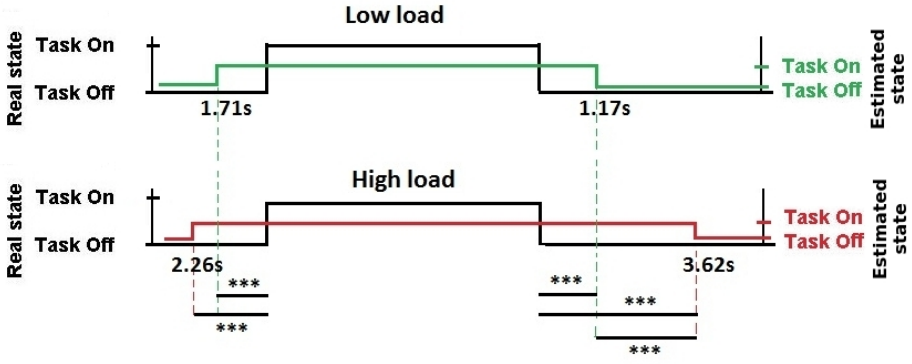


Figure 6.15 – **Off-line estimated onset and offset latencies compared to the stimuli onset, in low WM load and high WM load conditions.** *Average for 20 trials per difficulty, on 19 pilots' results. ***: $p<0.001$.*

The difficulty of the current trial had no effect on the onset estimation latency ($p = 0.15$). Concerning the offset estimation latency, the offset estimations occurred significantly later for high load trials ($Mean = 3.62s$; $SE = 0.66s$) than for low WM load ones ($Mean = 1.17s$; $SE = 0.59s$) ($F(1, 18) = 12.3$; $p < 0.01$).

6.2.3.3 Single trial SVM-based WM Load Estimation Results

During the testing phase, a mean of 80.8% ($SD = 10.6\%$) of the trials were accurately classified (discriminated into on-line *low WM load* trials and *high WM load* trials). We obtained a 72.11% mean specificity ($SD = 19.89\%$), and a 89.47% mean sensitivity ($SD = 15.72\%$). The details for classification accuracy and processing latencies are available in the article (Gateau et al. 2015).

6.2.4 Discussion

The objective of this study was to implement on-line tools to infer pilots' cognitive activity (Parasuraman and Wilson 2008). We focused on the monitoring of WM as this executive function is highly solicited when operating aircraft (Causse et al. 2011a;b). The design of such an inference system was challenging as, until now, only three studies involving on-line fNIRS-based state inference systems in an ecological context (Girouard et al. 2013, Solovey et al. 2012, Afergan et al. 2014) have been conducted. However, these studies did not include realistic simulators but simplified PC-based simulations. In order to test our inference system, an experimental protocol was designed, during which the pilots had to interact with ATC instructions of two levels of difficulty. The behavioral results confirmed that these levels were contrasted, as participants performed less well during the higher difficulty level. This result is coherent with Taylor et al. (2005) which has shown that pilots' WM decline when four different instructions have to be stored and recalled. The neurophysiological results also confirmed that the task difficulty statistically modulated oxygenation level in the prefrontal cortex (PFC). Moreover, the fNIRS data revealed particular activations of the right and left dorsolateral PFC (DLPFC; BA 9 and 46) that are seen as mediating monitoring, i.e., executive control in the Baddeley's model of WM (see chapter 7.1). Indeed, these results, in accordance with the findings of our previous experiments, confirmed that fNIRS is a suitable device for monitoring WM load level.

One novelty of the approach, which represented my main contribution to this study, was the use of a MACD filter as a systematic state estimator to detect the pilot's state (doing or not doing a WM task) (Durantin et al. 2014c). The results were promising as the MACD-based state estimation matched 62% of time with the real duration of on task and off task activity, i.e. receiving (memorizing) or not the ATC instruction (see Fig.6.14), with a good true positive rate (72%). The differences observed were explainable by the presence of 42% false positive rate, when the pilot's state was estimated as on task and the stimulus's state was off task. Our off-line analysis revealed that, on average, the on task estimated periods started before and ended after the stimulation periods (see Fig.6.15). Furthermore, the delay between stimulus onset and state estimation offset was significantly higher during the high WM load trials than during the low WM load ones, suggesting a potential different way of dissociating workload levels. According to our definitions of the real and the estimated states, it is of a great importance to consider that the estimated pilot's WM load could not perfectly fit to the real state as ATC messages are physical stimuli and pilot's state is the ATC brain-related activity that necessarily extends beyond physical stimuli, as it correspond to the operator's engagement in the task. To this extent, the latencies observed could be imputable to the anticipation of stimuli onset due to task rhythmicity (for estimated periods starting before the onset) and the maintenance in WM of ATC instructions (for estimated periods ending after the stimulus), two

functions that have neural substrates in the PFC (Miller and Cohen 2001). Therefore, the interesting potential of this approach is to detect on-task and off-task periods as internalized by the participants (i.e. when the participant actually engages in the trial, see chapter 7.1) and not based on the properties of the task. Altogether, these results are consistent with previous studies proving the potential of fNIRS for idle mode detection (Coyle et al. 2004). In fact, fNIRS offers an insight into the brain's reaction to stimuli, giving information on perceived workload that would not be available through behavioral measures (e.g. stimulation periods). The results also confirmed that MACD is an effective method for real-time task onsets detection (Durantin et al. 2014c), requiring no *a priori* information on task onsets, few computational resources, and no calibration. This result is key for ecological tasks, when onsets timing are unknown. For example, in realistic situations when events happen randomly, the information concerning task onsets could be retrieved using such a system. This method would provide a simple and systematic way to trigger classification algorithms for workload level assessment, in addition to a off task/on task mode detection.

6.3 Kalman filtering for signal improvement

We saw in the previous experiments that using innovative signal processing techniques (such as MACD filtering) allowed better extraction of the features (see sections 6.2.2.4 and 6.2). However, those results could be improved, especially in terms of false positive rates. An idea to improve feature extraction would be to use a priori knowledge on the temporal dynamics by implementing **estimation techniques**. Kalman (1960) developed the theory of estimation techniques to infer the state of a system from both measurements done on it and modeling of its temporal dynamics (see appendix A for details on the principle). Therefore, this technique could be used to improve the signal quality, for example by taking into account the dynamics of the BOLD response (Diamond et al. 2005) for fMRI (and similarly of the hemodynamic response for fNIRS) or of the ERP for EEG (Georgiadis et al. 2005a).

6.3.1 Kalman filter for EEG

The objective of the current study led in collaboration with CiNet laboratory (Osaka, Japan) was to assess the relevance of multiple processing techniques (including Kalman filtering) to process EEG data in real-world conditions. The signal processing was done during my stay in Osaka (Japan) at fall 2014, in the neuroergonomics laboratory of professor Daniel Callan. My contribution concerned mainly the development of a Kalman filter

to retrieve the ERP components from EEG data previously acquired by Pr. Callan, and the implementation of a formal classifier.

As we saw in chapter 7.2, retrieving the temporal dynamics (i.e. ERP) from EEG data on a single-trial basis still represents a great challenge for data processing. However, we also saw in chapter 7.3 that such components can be useful for BCI development. Georgiadis et al. (2005a) proposed a Kalman filtering approach using a mathematical model of auditory ERP as a way to improve single-trial retrieving of these components. In this study, we tested the relevance of this approach as a way to improve signal classification in realistic situations. The results of this study were published in Callan et al. (2015).

6.3.1.1 Methods

Task

The same subject was used for all studies. The subject was male, right handed, 45 years old, with normal hearing, and is one of the authors. He has 5 years of flying experience with more than 250 h total time and 200 h in biplanes of the same make and model. The subject gave informed consent for experimental procedures approved by the ethics committee of the National Institute of Information and Communications Technology in accordance with the principles expressed in the Declaration of Helsinki.

The experimental task consisted of passively listening to a chirp sound (0.1 s duration) in different environmental settings while recording brain activity using the Cognionics 64 channel dry wireless EEG system. The chirp sound was presented 200 times spaced randomly by at least 0.6–2.5 s of silence. Within these periods of silence 200 independent 0.5 s segments were randomly extracted. The different environmental settings and conditions are as follows: Electric Motion Platform (Platform Off, Platform On) and open cockpit Biplane (Engine Off, Engine On).

The CKAS V7 6 degree of freedom Motion System with a custom built cockpit utilizing dome projection for video was used (see Fig.6.16). In the motors off condition the subject was sitting still in the chair mounted on the platform. In the motors on condition the subject was engaging in aerobatic flight simulation (X-Plane Laminar Research) through a Redbull Air Race course (see Callan et al. (2011)). Just as in a real airplane, the subject controlled the ailerons and elevator by stick with the right hand, the throttle lever with the left hand, and the rudders with foot pedals. The motors of the motion platform moved in relation to the accelerations of the aircraft in the flight simulator. No sound was presented from the flight simulator. In both the platform off and platform on conditions the audio stimuli were presented using the Clarity Aloft Pro aviation headset using the same sound

level. The experiment took approximately 10 min for each condition. The platform on condition was conducted first followed by the platform off condition.



Figure 6.16 – *Motion Platform Flight Simulator composed of the CKAS V7 6 degree of freedom Motion System (CKAS Mechatronics, Melbourne, Australia) with a custom built cockpit utilizing dome projection for video. The flight controls consist of 1. A control stick to manipulate the elevator (pitch) and ailerons (roll), 2. Rudder pedals to manipulate the rudder (yaw), and 3. A Throttle to manipulate thrust. Single-trial auditory events using dry-wireless EEG were evaluated while the subject was flying through a simulated Redbull Air Race course.*

The biplane used in the experiment was an open cockpit two passenger Starduster SA300 (see Fig.6.17). In the engine off condition the subject was sitting in the front seat of the biplane on the tarmac with avionics off. In the engine on condition the subject was piloting the plane during cruise flight from the front seat with avionics on. In the engine on condition there was extreme vibration, acoustic noise, and wind. In both the engine off and engine on conditions the Cognionics EEG system was worn underneath a leather flight cap. The audio stimuli were presented using the Bose A20 active noise canceling aviation headset using the same sound level. The experiment took approximately 10 min for each condition. The engine off condition was conducted first followed by the engine on condition.

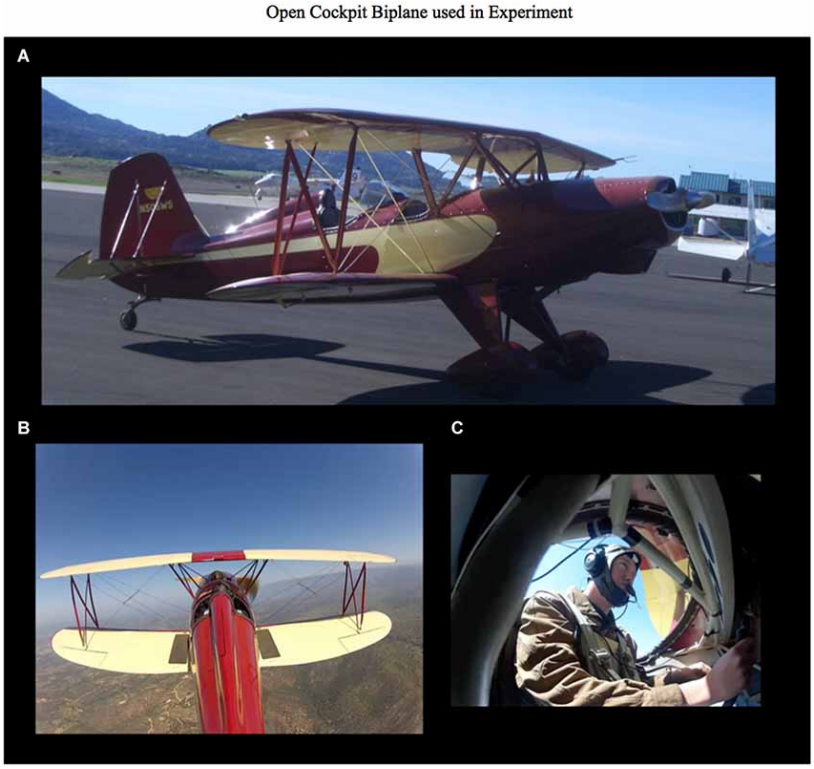


Figure 6.17 – *Open Cockpit Starduster SA300 Biplane used for in flight and ground testing of single-trial auditory events using dry-wireless EEG. (A) View off the biplane on the ground. (B) View of the biplane in the air. (C) Picture of the subject wearing the Cognionics 64 channel dry-wireless EEG under the leather flight helmet.*

EEG recording

EEG was recorded using the Cognionics HD-72 dry wireless EEG headset (Cognionics, Inc., San Diego). The EEG headset consists of a mechanically flexible spine to provide structure and ease of handling. Each segment of the spine contains a row of electrodes. The 64 electrodes provide full scalp coverage. An internal active shield, covering all sensor positions, spans the entire headset to minimize external noise pickup and artifacts. Reference and ground are placed on the mastoids with by two standard ECG adhesive electrodes (see Fig. 6.18).

Data processing

The EEG data were preprocessed using artifact-cleaning procedures available in EEGLAB (Delorme and Makeig 2004). The data were first filtered using a FIR band pass filter between 2 and 30 Hz with a filter order of 496. The data were separated into a training set of 75% of the stimuli (150 chirp stimuli and 150 silent stimuli) and a testing set of

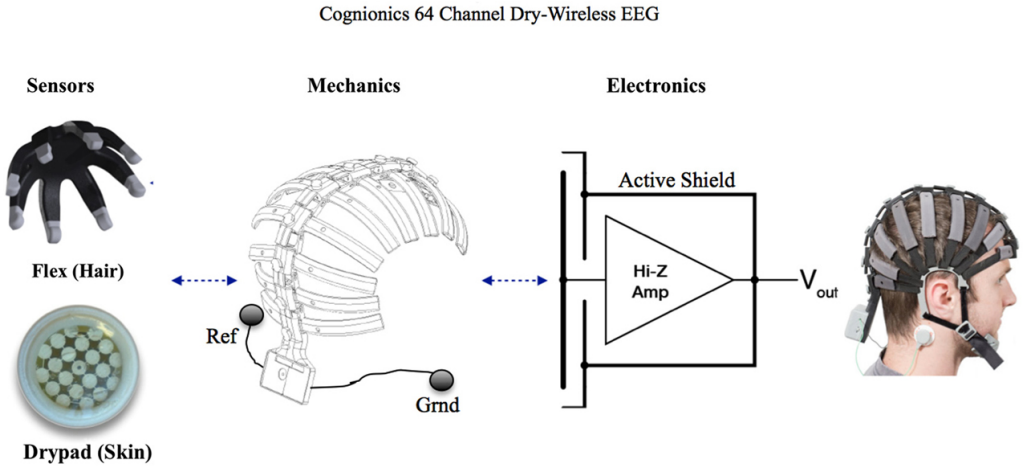


Figure 6.18 – *Cognionics 64 Channel Dry-Wireless EEG Headset. Two types of dry sensors are used: The flex sensors are placed over hair and the Drypad sensors are laced over bearskin, such as the forehead. In our setup, we utilized 7 Drypad sensors on the forehead (first band) and 54 flex sensors across the rest of the array. Each band of the headset is user adjustable to enable proper fit and sensor pressure over different head sizes and shapes. An internal flexible conductive layer forms an active shield, which spans the entire headset, to minimize external noise pickup even with high-impedance dry electrodes. Ground (Grnd) and reference (Ref) electrodes are placed under the ear as shown in the figure.*

25% of the stimuli (50 chirp stimuli and 50 silent stimuli) for the platform off and on conditions and the biplane engine off and on conditions. The rational for choosing 75% of the trials for training and 25% for testing is that we believed this division would provide enough training data to generalize to a novel testing set while maintaining enough trials in the testing set to be a reasonable representation of the population. Although we did not attempt to train and test on a 50% split of the data we believed that the amount of training data may not have been sufficient for generalization to the novel test set. The artifact cleaning procedures were applied only to the training sets. The results of which were then applied to the corresponding training sets. The exact same cleaning procedures were then applied to the corresponding testing sets, with the same parameters.

Channels were considered to be bad and were removed from the training data based on the following: Maximum amplitude in the channel is greater than 100 microvolts; Channel flatline duration is greater than 5 s; Channel is correlated at less than 0.8 to its robust estimate based on other channels; Channel has more line noise relative to its signal than 4 standard deviations from the channel population mean. These same channels were then removed from the corresponding test sets.

The next step was to clean the multi-channel training data with the artifact subspace reconstruction (ASR) method using the defaults given in the clean artifacts software within EEGLAB. ASR allows for the removal of non-stationary high-variance signals from EEG and reconstructs missing data using a spatial mixing matrix (see Mullen et al. (2013) for

details). Calibration data from clean segments of the training data were used to determine the ASR filters separately for each of the four conditions. The ASR filter was then applied to the corresponding test data.

Following cleaning of the data by ASR, ICA using the extended infomax algorithm in EEGLAB was applied in order to separate brain activity related to the auditory evoked potentials from other brain and artifact related components. The weights of the ICA were determined only from the training data separately for each of the four conditions (platform off, platform on, engine off, engine on). These weights were then applied to the corresponding test data to determine the ICA activation waveforms. See Fig. 6.19 for an example of the continuous data before ASR, after ASR, and after ICA.

Kalman filtering was applied to the ASR filtered data both before and after ICA, to produce an estimate of the single-trial auditory evoked response. Kalman filtering is a linear quadratic estimation method, which relies on both the measurements and a model of the Event-Related Potential (ERP) dynamics to product an estimate in real-time. This method has been successfully applied to single-trial event related potentials estimation. In our case, the parameters used as a model for the Kalman filter design were the N100 and P300 waves, generic features of the auditory evoked potential (see Fig. 6.20). Both waves were modeled by a third-order impulse response with peaks respectively at 100 ms and 300 ms. The state noise was considered as white noise (Georgiadis et al. 2005a), and the measurement noise covariance to state noise covariance ratio (Q/R, see appendix A for details) was fixed to 0.001, so that the Kalman estimate would put confidence in the measurements (see Alazard (2005)).

Trial classification

The Matlab Least Squares Probabilistic Classification (LSPC) toolbox (Sugiyama 2010) was used to determine how well single trial audio presented stimuli could be identified in the EEG signal from periods of audio silence. LSPC uses a linear combination of kernel function to model the class-posterior probability. Regularized least-squares fitting of the true class-posterior probability is used to learn its parameters. The use of least-squares fitting to determine a linear model allows for a global solution to be made analytically providing a considerable speedup in computational time. The default parameters were used in training of the LSPC models.

The features used to train the classifier consisted of the samples from onset to 500 ms after stimulus onset (150 samples) for both chirp and silent stimuli. In the case of ICA data the components selected to be included in the LSPC model were determined by visual inspection of the mean of the activation waveform of the chirp training data stimuli

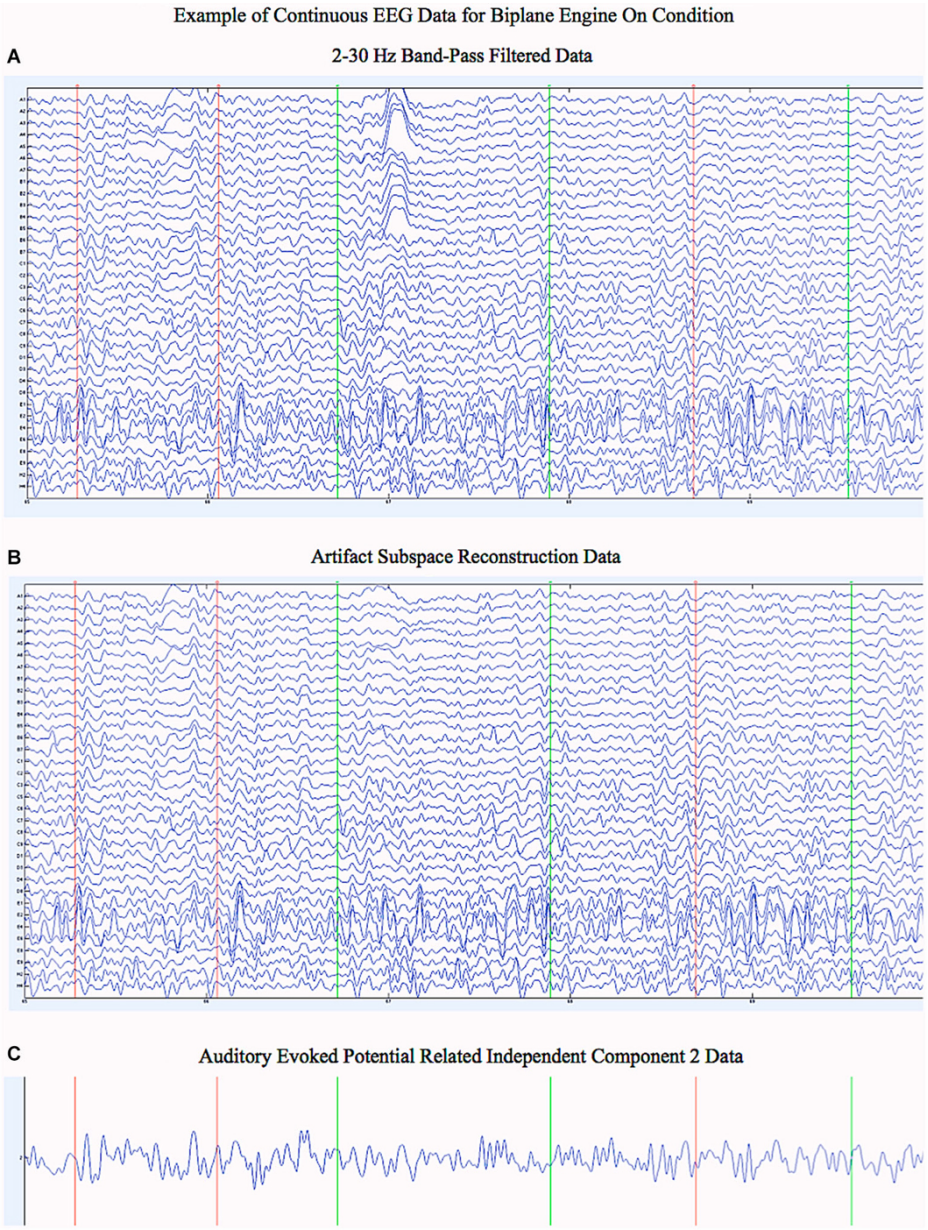


Figure 6.19 – *Example of Continuous EEG Data for the Bipocuslane Engine On Condition. (A) Five seconds of continuous data from the test set that has been band pass filtered from 2–30 Hz for the 34 electrodes included. (B) Five seconds of continuous data from the test set that has undergone artifact subspace reconstruction as well as band pass filtered from 2–30 Hz for the 34 electrodes included. Notice how ASR has cleaned up the artifact present in many of the channels (compare with (A) above). (C) Five seconds of continuous data from the test set for auditory evoked potential related independent component number 2. The ICA was carried out over the data in (B). Green lines denote the onset of audio stimuli and Red lines denote the onset of silent trials.*

that showed a characteristic auditory evoked potential. For the pre-ICA data all of the channels were included in the LSPC model. The large number of features relative to the number of trials used to train the classifiers has the potential for over-fitting, resulting in

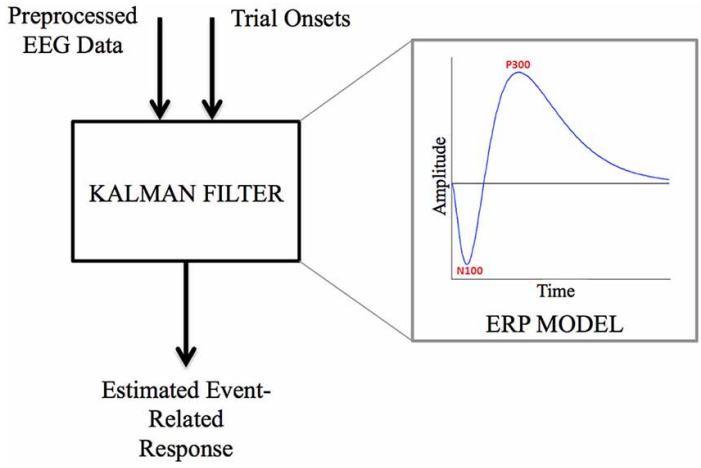


Figure 6.20 – ***Kalman filter functional model***. *Kalman filter is designed from an Event Related Potential (ERP) dynamical model (on the right), and uses both preprocessed EEG data and trial onsets to estimate the instantaneous event-related response.*

poor generalization to the novel test data. We utilized inner cross-validation procedures for model selection in part to assess and protect against potential over-fitting.

The following inner cross-validation procedures were used for model selection and testing: Randomized ten fold cross-validation was used on the training data and the trained model with the best performance was selected for evaluating performance on the test data. One fold of the training data consisted of 135 chirp and 135 silent stimuli and one fold upon which the model was evaluated for selection consisted of 15 chirp and 15 silent stimuli. The final test data consisted of 50 chirp and 50 silent stimuli. This procedure was conducted 100 times to determine the distribution of performance of the model given random aspects of training (e.g., the trials selected for training and their order into the model).

Several conditions were evaluated to assess the efficacy of ICA and Kalman filtering in improving classification performance on the data. These conditions included the four environmental conditions (Platform Off, Platform On, biplane Engine Off, biplane Engine On) and four preprocessing conditions (pre-ICA constituting the baseline no processing condition, ICA, pre-ICA Kalman filtered, ICA Kalman filtered) for a total of 16 analyses. The statistical significance between the various conditions was determined by repeating the nested cross-validation process ten times for each condition. The nested cross-validation procedure is commonly used to estimate the stability of the model across conditions. At each step, we computed the classification accuracy, false positive rate and false negative rate. The differences between the classification results obtained from each of the preprocessing conditions were then compared using Wilcoxon sign-rank test.

6.3.1.2 Results

The number of channels not rejected out of the 64 by the preprocessing steps for the various conditions are as follows: Platform Off (46), Platform On (20), Biplane Engine Off (30), Biplane Engine On (34). The independent components showing auditory evoked potentials used to train the LSPC models were the following for the various conditions: Platform Off (2, 3), Platform On (1, 3), Biplane Engine Off (2, 3), Biplane Engine On (2, 16). The components are sorted in descending order of mean projected variance. Therefore, the lower the component number the larger is the projected variance.

The results of the single trial classification performance for audio chirp verses silent stimuli for all environmental and analysis method conditions were significantly above chance ($p < 0.05$) based on permutation testing of 100 random shuffling of the labels compared to the mean performance of the respective models trained with correct labels. The accuracy and d' (discriminability index, taking into account both the false positive and false negative ratios) for the platform off and platform on conditions for the four analysis conditions (Pre-ICA, ICA Pre-ICA + Kalman Filter, and ICA + Kalman Filter) are presented on Fig. 6.21 and 6.22. Tables presenting all the statistical results and exact values are given in appendix B.

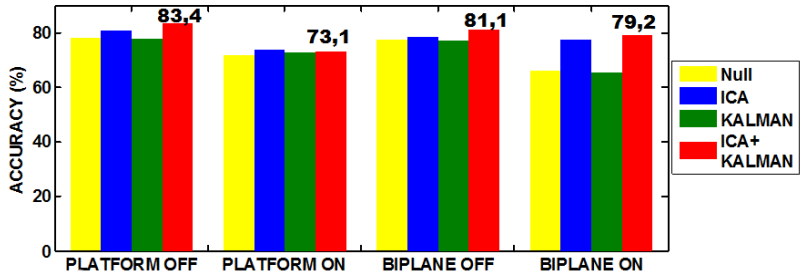


Figure 6.21 – ***Accuracy results for all conditions in function of the signal processing techniques used (in percent).*** *Bars represent the mean accuracy obtained for Pre-ICA (Yellow), ICA only (Blue), Kalman only (Green) and Kalman+ICA (Red) conditions. The value of the accuracy is given for the Kalman+ICA condition.*

6.3.1.3 Discussion

The primary objective of this experiment was to show as a proof of concept that a dry-wireless EEG system could be used in extremely noisy real-world environments while the operator is carrying out complex perceptual motor tasks and still show good classification performance of a perceptual event. Both the motion platform based flight simulation and the open cockpit biplane environments produced considerable artifacts challenging the ability for accurate classification of a perceptual event from EEG data. In the platform on

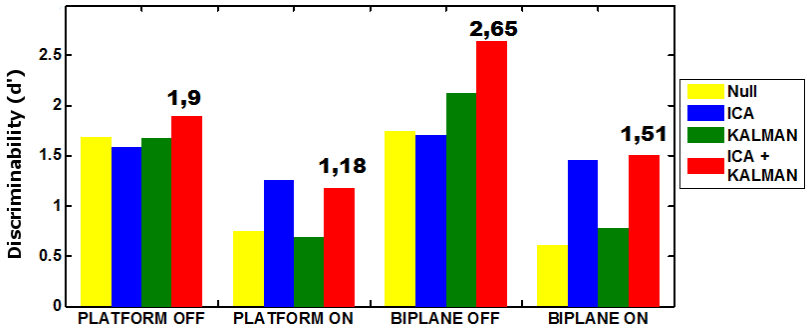


Figure 6.22 – ***Discriminability results for all conditions in function of the signal processing techniques used***. *Bars represent the discriminability obtained for Pre-ICA (Yellow), ICA only (Blue), Kalman only (Green) and Kalman+ICA (Red) conditions. The exact value is given for the Kalman+ICA condition.*

condition the subject was flying through a Redbull Air Race course making abrupt banks to go through the cones in the correct orientation (Callan et al., 2013). The movement of the platform and the head and body to these banks produces large artifacts in the EEG trace. In the open cockpit biplane there is considerable vibration, wind, acoustic noise, and physiological noise resulting from control of the plane and movement of the body that produces large artifacts in the EEG trace.

Despite these challenging environments our study showed that far above chance classification performance can be achieved for all environmental conditions (*Platform Off*, *Platform On*, *Biplane Engine Off*, *Biplane Engine On*) and for all analysis conditions (*Pre-ICA*, *Pre-ICA Kalman filtered*, *ICA*, and *ICA Kalman filtered*). The classification performance was the best when applying both ICA and Kalman filtering to the data. This was especially true in the Biplane Engine On condition in which classification performance improved from 66.1% in the baseline pre-ICA no processing condition to 79.2% in the ICA Kalman filtering condition . This classification performance was comparable to that of the *Engine Off* condition (81.1%) in which the biplane was sitting on the tarmac. It is unclear why the *Platform On* condition did not show the same beneficial effects of ICA Kalman filtering vs. the baseline no processing condition (73.1% vs. 71.6% respectively), as did the *Biplane Engine On* condition. One possibility is that the abrupt movement of the body and head during the flight task produced non-stationary artifacts that could not be separated by the EEGLAB artifact cleaning methods or by ICA and Kalman filtering. The significantly greater accuracy for the Off over the On conditions is to be expected not only because of the much greater artifacts present in the *On* conditions but also because auditory responses are decreased depending on attention (see chapter 7.2). Therefore, it is likely that the attention needed for piloting in the *On* conditions reduced attention to the auditory stimuli.

Interestingly, the use of Kalman filtering was most effective when conducted on ICA processed data. Kalman filtering on pre-ICA data did not show improvement in perfor-

mance compared to that of ICA processed data (see tables in appendix B). One possible explanation for these results is that ICA is able to separate the brain activity related to the auditory event from other brain activity and artifacts. Only the ERP components are enhanced by the Kalman filter that utilizes generic information about the auditory evoked response as a dynamical model. In the pre-ICA data the brain activity for various processes as well as artifacts are mixed within all the channels and estimation of the auditory events by the Kalman filter model may be more difficult for these signals.

Because the results are only of a single subject it is not clear how well they generalize to other individuals. It is likely that different individuals with different levels of flight experience may show more/or less movement and physiological artifacts that may decrease/increase classification performance. Despite these limitations the results do stand as a demonstration that it is possible to collect data in real-world environments that have considerable noise, both environmental and physiological, using dry-wireless EEG and obtain classification performance sufficient for BCI applications. In particular, it is important to observe that the Kalman filter developed in this study was efficient at improving ERP single-trial extraction even for data collected in real-world situations, leading to a better classification performance.

6.3.2 Kalman filter for fNIRS

The use of a Kalman filter including a physiological model for signal improvement has been previously applied to EEG (Georgiadis et al. 2005a, Callan et al. 2015) or fMRI (Diamond et al. 2005). However, concerning fNIRS, this technique has been limited to the estimation of model parameters (Abdelnour and Huppert 2009) or the correction of motion artifacts (Izzetoglu et al. 2010), therefore not requiring the use of a physiological model of hemodynamic response to stimulation. The objective of this study was to design a Kalman filter including a physiological model of hemodynamic response (Boynton et al. 1996) suitable for fIRS. By applying this filter to fNIRS data collected during both controlled and ecological experiments, we also aimed at testing the improvements such a filter can bring to fNIRS signal.

One of the greatest challenges regarding Kalman filter design is the tuning of its parameters, i.e. to evaluate the level of measurement noise (R) and state noise (Q) in the model (Diamond et al. 2005). As discussed in appendix A, the value of the ratio Q/R greatly influences the behavior of the Kalman filter. Indeed, a Kalman filter with a low Q/R will put confidence in the dynamical model, whereas a Kalman filter with a high Q/R will put confidence in the measurements. In practice, the value of this ratio often has to be chosen empirically (e.g. Abdelnour and Huppert (2009), Callan et al. (2015)), as there exists no efficient way to measure it. Consequently, the dynamics of the Kalman

filter may not be adapted to the data needed to be improved. In this study, we propose a two step approach for Kalman design and tuning. In a first controlled experiment, we applied multiple Kalman filters with different tunings in order to find an optimal value for the Q/R ratio. The resulting optimal Kalman filter was then applied to data from an ecological experiment to assess its stability.

6.3.2.1 First step : optimization of filter parameters

Procedure

Nine healthy participants from ISAE (Mean age 21.6, SD=1.5 ; 1 woman) participated in the experiment. The experiment procedure was similar to the one used in the study described in section 6.2.2.4, and consisted in the memorization of digit sequences. The only difference from the protocol used in 6.2.2.4 was the use of only three levels of difficulty, defined by the size of the digit sequence (5, 7 or 9). Nine trials were presented for each difficulty level. During the experiment, the level of HbO_2 and HHb was measured over the prefrontal cortex using a fNIR100 device (Biopac®).

Data were processed using Matlab®. Two different types of Kalman filters were applied to the data (their development is described in appendix C), the **nominal Kalman filter** (in which we assumed that the stimulus perception bias is null on average) and the **augmented Kalman filter** (without this assumption). The inputs for both filters were the stimuli onsets and the raw fNIRS data. For each filter, the value of the Q/R ratio ranged from 10^{-5} to 10^{5}, in order to look for the optimal results. Simultaneously, we also applied the MACD filter (see section 6.2.2.4) to raw data in order to compare Kalman results with classical linear filtering.

Preliminary, the potential good values for Q/R (i.e. those leading to improvement in the signal) were isolated by computing an **Effect Size Index** (E, illustrated on Fig. 6.23. For each filter (MACD, nominal Kalman or augmented Kalman with a given Q/R value) and each digit sequence size N, we computed the mean (μ_N) and standard deviation (σ_N) of the level of HbO_2 or HHb measured at each optode. We used these values to compute the 68% confidence interval as $[\mu_N - \sigma_N; \mu_N - \sigma_N]$. The **Effect Size Index** (ESI) was defined as the gap between the confidence intervals of each condition (negative if the confidence intervals are overlapping), i.e. :

$$ESI = ((\mu_7 - \sigma_7) - (\mu_5 + \sigma_5)) + ((\mu_9 - \sigma_9) - (\mu_7 + \sigma_7))$$

We then proceeded to visual inspection to find the best values for Q/R ratio, by finding

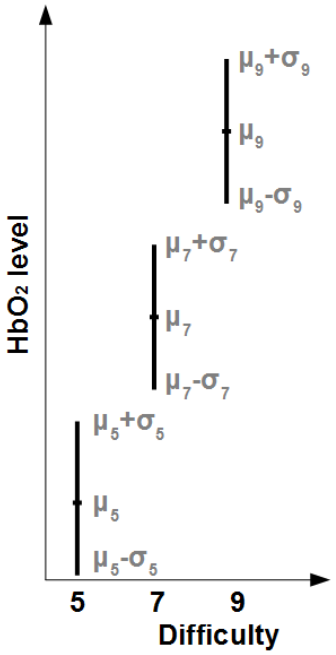


Figure 6.23 – ***Illustration of the computation of the Effect Size Index (ESI) for*** HbO_2**.** *We computed the mean (μ) and standard deviation (σ) of the level of HbO_2 across subjects for each difficulty. For a difficulty N, the 68% confidence interval was computed as $[\mu_N - \sigma_N; \mu_N - \sigma_N]$. The corresponding effect strength index was computed as the sum of the gaps between the confidence intervals (negative if the confidence intervals are overlapping), i.e. :* $ESI = ((\mu_7 - \sigma_7) - (\mu_5 + \sigma_5)) + ((\mu_9 - \sigma_9) - (\mu_7 + \sigma_7))$.

the parameters leading to higher ESI values. Each set of data was finally tested using a 16x3 repeated measures ANOVA, with two factors (16 voxels, 3 levels of difficulty), performed using STATISTICA®. The strength of the statistical effect of difficulty level, evaluated using the $partial\eta^2$, was used to compare the results of the different filters.

Results

As shown on Fig. 6.24, the optimal results were obtained for HbO_2 at optode 2 located in the DLPFC, when using the nominal Kalman filter with $Q/R = 3.98$ or the augmented Kalman filter with $Q/R = 0.50$. Table 6.2 summarizes the effect sizes obtained for each of the signal processing techniques tested (those effects showed an increase in the level of HbO_2 with growing sequence sizes).

Filter type	Q/R **ratio**	**Effect size (**$partial\eta^2$**)**
MACD		0.21
Nominal Kalman	3.98	0.32
Augmented Kalman	0.50	0.34

Table 6.2 – *Effect sizes obtained for the effect of difficulty over all the subjects for the level of HbO_2 measured at optode 2, depending on the type of filter used for signal processing.*

The frequency and phase responses (Bode diagram) of the nominal Kalman filter ($Q/R = 3.98$) and of the augmented Kalman filter ($Q/R = 0.50$) are given on Fig. 6.25.

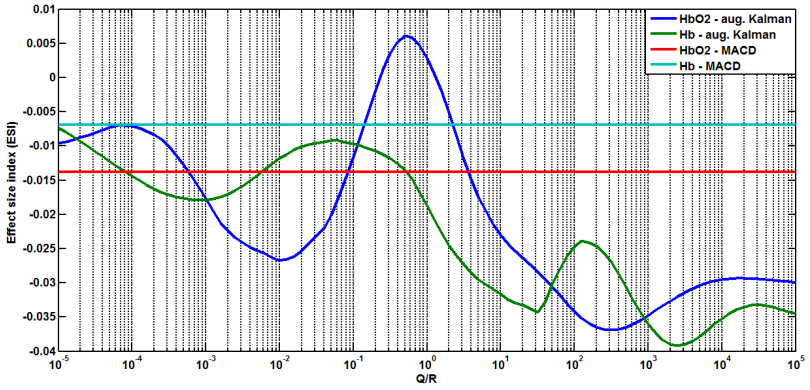


Figure 6.24 – *Estimated Effect Size Index (ESI) for the augmented Kalman filter (for HHb –in green– and HbO_2 –in dark blue–) in function of the value of Q/R, compared to the ESI of the MACD filter (in red for HbO_2, in light blue for HHb). The index is estimated using the data from optode 2.*

As the two filter exhibit very similar Bode diagrams (and therefore very similar filtering properties), we retained only the augmented Kalman filter for testing on new data.

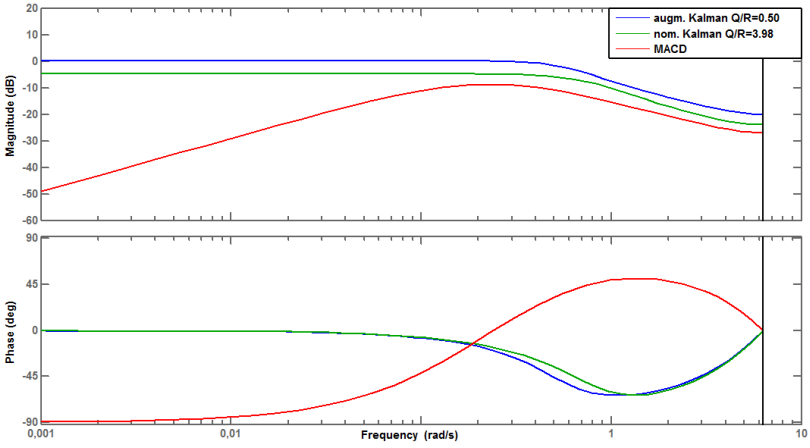


Figure 6.25 – *Bode diagrams (frequency response, phase response) of the MACD filter, the nominal and augmented Kalman filters.*

6.3.2.2 Second step : Optimal filters testing

Procedure

Data used for testing were extracted from the data of the BCI design experiment (Gateau et al. 2015), and included 18 healthy subjects. The task consisted in the memorization of ATC instructions in a flight simulator (see section 6.2 for full description). ATC instructions comprised two levels of difficulty : high or low. We chose this experiment

as it recruited WM, similarly to the experiment used for the Kalman filter optimization procedure.

The raw HbO_2 data of optode 2 were filtered using three types of filter. First, we used the MACD filter and the augmented Kalman filter retained from the optimization phase ($Q/R = 0.50$). We also added a classical IIR Butterworth passband filter ($0.02 < f < 0.1$) (Ayaz et al. 2012b), in order to compare the results to classical filtering. The statistical effect sizes were evaluated using repeated measures ANOVA performed with STATISTICA®. The performance of the different filters were compared in terms of $partial\eta^2$.

Results

The partial η^2 obtained for each type of filters are given in table 6.3 (at optode 2 and across all the optodes). The results show that the use of MACD elicits a better statistical effect size than the classical IIR filter. Similarly, the use of Kalman filter yields better results than both MACD and IIR filters. This result is true not only when filtering data from optode 2, but also when filtering data from every optode.

	Optode 2	**All optodes**
Filter type	**Effect size ($partial\eta^2$)**	**Effect size ($partial\eta^2$)**
IIR	0.3623	0.3627
MACD	0.4356	0.3315
Augmented Kalman ($Q/R = 0.50$)	0.5504	0.5164

Table 6.3 – *Effect sizes obtained for the effect of difficulty over all the subjects for the level of HbO_2 measured at optode 2 or at all optodes, depending on the type of filter used for signal processing.*

The effect of trial difficulty on the level of HbO_2 measured at optode 2 is shown on Fig.6.26 (normalized by the mean value for low load trials, for clarity). On this figure, we observe that the advantage of both MACD and Kalman filter over classical IIR filter is the reduction of the standard error of the mean, meaning that the value measured varies less across trials and subjects. In particular, Kalman filter exhibits both a large effect size and a small standard error, compared to other filters. The performance of the MACD filter is limited by its low static gain, leading to reduced effect sizes.

6.3.2.3 Discussion

The objective of the study was to design a Kalman filter to improve fNIRS signal. In particular, the main challenge concerned the tuning of the parameters Q and R (Diamond et al. 2005). Based on a simple model of the hemodynamic response to neuronal

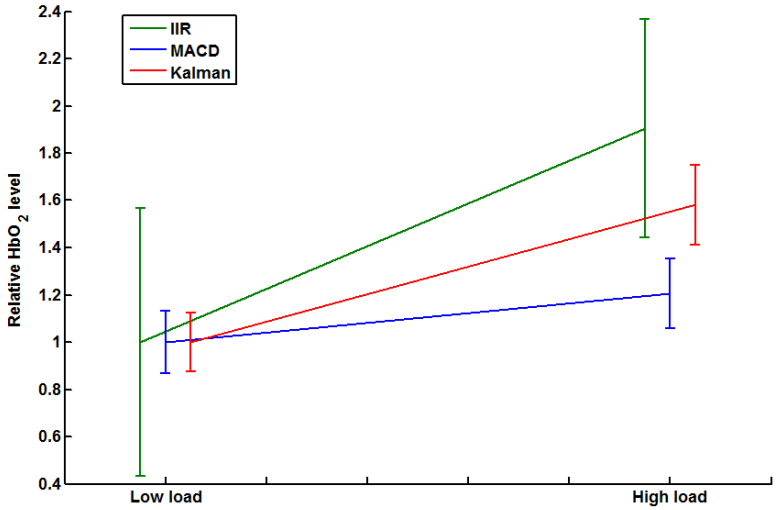


Figure 6.26 – *Comparison of the effect of trial difficulty (Low load or high load, see section 6.2) on the level of HbO_2 measured at optode 2, in function of the type of filter used for signal processing (classical IIR in green, MACD in blue, Kalman in red). For each filter, the level of HbO_2 is normalized by its mean value during low load trials, for better readability. Error bars represent the standard error of the mean.*

stimulation (Boynton et al. 1996), we designed a Kalman filter model taking into account both the measurement noise and the stimulus perception bias that can occur in periods of disengagement or when the level of attention varies (see appendix C). During an optimization process, we showed that it was possible to find values for the parameters which leads to better statistical results. Finally, we applied the optimal results found during the optimization process on new data from an ecological experiment, and showed that the optimal Kalman filter tuning can be applied generically. In particular, this filter led to higher effect sizes when looking at the effect of task difficulty in both tasks, compared to classical filters. Additionally, the use of a dynamic physiological model by the Kalman filter implies less variability across trials and subjects, therefore explaining the greater stability of the results obtained with this filter (see Fig.6.26). These results suggest that this filter would improve the discriminability between the two conditions.

Interestingly, the optimal results for this experiment were found at optode 2 located in the DLPFC. This result is concordant with previous fNIRS studies that have found that this region is sensitive to WM solicitation (Ayaz et al. 2012b, Durantin et al. 2014a). The improvement of the fNIRS signal collected in this region sensitive to WM suggests that this filter could be applied to any experiment recruiting the same region. However, further investigation is still needed to assess if this filter could be used with the same model and tuning in experiments recruiting different brain areas.

Some modifications of the model could lead to better usability and results of the Kalman filter. For instance, the use of a stimulus onset prediction technique such as

MACD filter (Durantin et al. 2014c, Gateau et al. 2015) could replace the stimulus onsets input of the Kalman filter, therefore reducing the complexity of the filter. In addition, it would be interesting to compare the results of the current Kalman model relying on a simple modeling of the hemodynamic response to more complex (but more accurate) models (e.g. Buxton et al. (2004b)). Finally, using an adaptive Q/R gain or realizing an optimization process for each subject instead of using a generic filter could also yield better results.

Altogether, the promising results of the study stand in favor of the use of Kalman filtering as a signal improvement technique for fNIRS signals. In particular, the improved signal would be available in real-time, as most signal improvement techniques require calibration or can only be used offline.

IV

Conclusion

General discussion

7

The objective of this PhD project was to investigate the effects of attentional dynamics on human performance. Our approach was based on a neuroergonomics methodology that used brain imaging techniques and psycho-physiological measurements in different experiment set-ups. Signal processing techniques were also implemented to allow the use of such measurement techniques under realistic settings for on-line purposes.

7.1 Attentional dynamics

7.1.1 The dynamics of engagement and disengagement

The studies conducted during this PhD project, aimed at characterizing failures in the operation of complex systems using psycho-physiological and cerebral measurements, and point towards the existence of specific strategies during task execution. On one hand, the study characterizing mental overload (see chapter 7.4 and Durantin et al. (2014a)) exhibited specific resource allocation patterns with less recruitment of the DLPFC in the case of mental overload, associated with decreases in performance. This disengagement pattern was also present in the LF/HF ratio, confirming a decrease in ANS activity during the most difficult condition. On the other hand, the studies on mind wandering (see chapter 7.5 and Durantin et al. (2015)) demonstrated the progressive decrease of attention directed at the task in the case of low mental demands, which was associated with activations of brain areas of the DMN (in particular the mPFC) and task errors. In addition, the second experiment exhibited a larger baseline level of pupil dilation during these moments. In all these experiments, the behavioral description would be the same for both levels of demands (i.e. disengagement from the task leading to poorer performance), but the neural activations measured were not the same, as we observed a decrease of the activity of the lateral areas of the PFC in one case (mental overload), and an increase of the activity of the medial areas of the PFC in the other (mind wandering).

These results can be analyzed in a consistent manner when considering the **goal-**

oriented nature of human cognition, describing the fact that human action is always driven by goals and motivations (see chapter 7.1, Leontiev (2014), and Bedny and Karwowski (2004)). From an evolutionary perspective, it is likely that the emergence of goal-directed behavior has been one of the driving forces guiding the evolution of *Homo* (Koechlin 2014), and explains the existence of mechanisms of resource allocation specific to humans, visible on both cognitive and physical levels [1]. Accordingly, the decrease in the lateral PFC activity observed in the mental overload experiment is to be put in relation with the role of this area in goal maintenance in working memory (Rossi et al. 2007) and the fact that it is part of the Task Positive Network (Johnson and Zatorre 2006, Harrivel et al. 2013). This activity could correspond to an attentional focus on a single task (i.e. piloting the plane, regardless of the instructions) and therefore a partial disengagement from the task, in accordance with the goals of the participants. The decrease of DLPFC activity could then be associated with an inability for multitasking as participants were only able to focus on handling the aircraft, knowing that the lateral parts of the PFC are responsible for the evaluation of alternative goals and strategies (Koechlin et al. 2000, Donoso et al. 2014). On the same basis, we observed the activation of the medial areas of the PFC (part of the Task Negative Network) before mind wandering episodes in the experiments presented in chapter 7.5. The role of these areas in the switch between strategies in adaptive behavior (Koechlin et al. 2000, Donoso et al. 2014) could also be an evidence of the goal-driven cognition. Indeed, the participants would disengage from the task due to low interest, and switch their attention to their own interests. The mPFC activation would be a result of this mode transition. This hypothesis is also consistent with the supposed functional role of mind wandering in autobiographical planning (Baird et al. 2011) and with the implication of the locus cœruleus network of vigilance (suggested by the Eye Tracking results). In both experiments, the observed activations on the PFC would therefore be the consequence of the progressive transition between the "**engaged**" and the "**disengaged**" modes.

The results of these studies are particularly interesting for the Neuroergonomics approach when it comes to the monitoring of the operator. Indeed, they are consistent with the **transient hypofrontality hypothesis** (Dietrich 2003), which suggests that human failures are caused mainly by a state-induced transient modification of the functioning of the prefrontal cortex. Accordingly, the detection of these transient modifications by brain imaging techniques (and particularly fNIRS) could be a relevant precursor of the degradation of cognitive performance and could be used as a trigger for adaptive human-system interaction. Moreover, the important role of the notion of engagement in the explana-

[1] Endurance running, for instance, is predominant in humans as it is directed towards a goal, although it is an energetically costly mechanism. On the contrary, it is a very uncommon behavior in other mammals and thus suggests a specific evolution of *Homo Sapiens* towards goal directed activity (Bramble and Lieberman 2004, Raichlen et al. 2012)

tion of the behavior points to the existence of specific goal-directed strategies of resource allocation.

7.1.2 A triad of concepts to understand human behavior

Our experiments exhibited the importance of the operator's engagement when it comes to the explanation of the causes of resource allocation. A comprehensive way to formalize it is through the concept of **intentional set**. This concept describes the factors influencing the way the brain chooses to focus on a different set of stimuli[2] and the way it shapes its reactions to a given set of stimuli[3], in a manner which is consistent with its goals (Rushworth et al. 2002). Taken together, **task demands**, **resources**, and the **intentional set** explain the dynamics of resource allocation (i.e. the relation between the intentional set and cognitive resources), as well as the dynamics of engagement or disengagement from a task (i.e. the relation between task demands and the intentional set). We earlier discussed (§1.2) that the concept of mental workload brings an answer to the question **What?** (i.e. what types of resources are allocated to face task demands and in which quantity), but not to the questions **How?** (i.e. how is resource allocation influenced by the mental state) and **When?**. The addition of the intentional set to the pair of concepts corresponding to the mental workload theory (i.e. task demands and resources) allows the definition of a triad of concepts to explain human behavior (see Fig. 7.1). The triad of concepts presented in this model implies that mental workload corresponds to the question **What?**, whereas operator engagement corresponds to two questions (**How?** and **When?**). Thus, engagement appears to be a more complex state to define, as it aggregates both the causes and the dynamics of resource allocation. Nevertheless, when taken together, these concepts can explain the transition between operational states and executive failures, as shown in Fig.1.5 (chapter 7.1).

On this basis, we conclude that understanding human behavior requires the assessment of the state of the individual and the task according to three concepts : the state of his/her resources, his/her intentional set, and the state of the task. Consequently, the advantage of the Neuroergonomics approach is to study human brain function under realistic settings (under which the goals and motivations of the operator are directly put in relation with the actions performed, as in everyday life), rather than in laboratory tasks (usually adopting a procedural view of human cognition).

[2] Also known as "Attentional Set" (Béguin and Clot 2004).
[3] Also known as "Process set" (Duncan 1980).

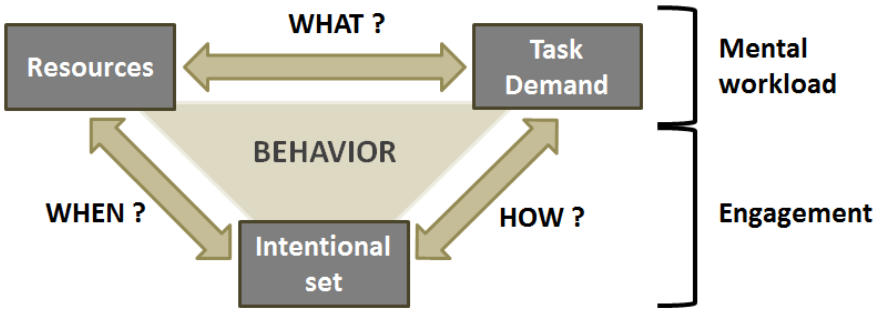


Figure 7.1 – *The triad of concepts underlying human behavior. Behavior involves interactions between task demands, resources, and the intentional set. The relation between task demands and resources is characteristic of the mental workload theory and describes the mobilization of cognitive resources to face task demands (What?). The addition of the intentional set in this approach explains the motivations and substrates underlying resource allocation (How?) as well as the dynamics of engagement and disengagement when facing task demands (When?).*

7.2 Neuroergonomics : use of brain imaging signals in realistic conditions

During this PhD project, we applied a neuroergonomics methodology to study attentional processes from basic to realistic settings. The results of these different studies confirm the relevancy of this approach as our findings were consistent across experiments. Indeed, most of the studies presented in this dissertation exhibited similar activations during working memory tasks, although the experimental conditions were not the same. The DLPFC activations were measured by fNIRS in the laboratory task (digit sequence memorization) presented in sections 6.2.2.4 and 6.3.2, but also in the micro-world task involving working memory (mental overload experiment, see chapter 7.4) or in the flight simulator involving ATC instructions memorization (see §6.2). This result supports the relevance of the Neuroergonomics approach, and justifies that the understanding of human behavior in ecological situations can benefit from its study in laboratory conditions. Therefore, in the interest of designing solutions to mitigate human errors in operational situation, studies should consider the gradual approach presented in chapter 7.2, ranging from very controlled but less realistic conditions to badly controlled but realistic conditions. In addition, the results from the signal processing part of this contribution also suggest that the approach of Neuroergonomics and BCI design can benefit from practices of other domains. The application of the MACD filter used in economic market analysis and of the Kalman filter mainly used in systems control and estimation (especially in aircraft and spacecraft design) significantly improved the quality and usability of the signal. Therefore, both the signal analysis performed in laboratory studies and the technologies developed for BCI use in everyday life can be improved by considering the methods from other fields, and especially those in which signal processing abilities have been dominant from their be-

ginning, such as system control. In most of the other scientific fields, the sensors usually measure variables that are physically well known (e.g. acceleration, velocity) but the use of the sensors in degraded environment requires the use of advanced signal processing techniques. In Neuroergonomics, their use is even more necessary, when we consider that in addition to the fact that the measures are done in a degraded environment, the variables that are estimated are often poorly defined or not well-known (e.g. mental workload, attentional state).

First, we applied the **MACD filter** on fNIRS signal collected over the PFC to detect event-related hemodynamic changes in real-time (Durantin et al. 2014c), in order to predict stimulus onsets during WM tasks. The promising results obtained (88% of the onsets were correctly detected, see §6.2.2.4) added to the fact that this technique does not require any calibration makes it a good candidate for fNIRS signal processing in real-time. When applied on fNIRS signal during simulated flight (see §6.2), this technique allowed a good classification (62%) of stimulation periods, as well as the successful development of a passive BCI to assess the level of mental workload (with 80.8% of accuracy). These results confirmed that MACD filtering allows the use of fNIRS signal in real-time. Interestingly, the particularly high false positive rate obtained in the controlled experiment (42%, see §6.2.2.4) as well as in the ecological one (42%, see §6.2) suggested that this filter was sensible to user **engagement** in the task rather to stimuli onsets. The latencies obtained for the detection of stimulation periods in the flight simulator (see §6.2, Fig. 6.15), showing on-task estimated times extending largely around the stimulation periods, adds support to this hypothesis. Therefore, the study of the MACD filter confirmed that this technique can be applied to obtain an online estimation of the engagement of the operator in the task. The studies also support the use of fNIRS for idle mode detection (Coyle et al. 2004, Tomita et al. 2014), which is one of the major challenges towards the design of passive and active asynchronous BCI (see §3.3). The unnecessity of training phase when using this technique also makes it a good candidate to improve the Information Transfer Rate of fNIRS-based BCI.

Second, we developed two models of the **Kalman filter** applicable for fNIRS and EEG data. In line with the Neuroergonomics approach and with what was previously discussed in this discussion (see §7.1), this technique allows for the processing of brain signals by taking into account both the characteristics of the task (stimuli onsets) and of the individual (by implementing a physiological model of the individual response to stimulation). On EEG signals, the technique led to a better retrieval of the ERP components on a single-trial basis (see §6.3.1). The information contained in the ERP components would then be accessible without multiple repetitions (as is often needed in EEG BCI ; see §3.2), therefore improving both the Information Transfer Rate and the accuracy of EEG-based BCI. When applied on fNIRS signal (see §6.3.2), Kalman filtering induced a better discriminability between the levels of oxygenation induced by various levels of mental workload,

both in controlled and in ecological situations. Interestingly, it also improved the stability of the results across trials of the experiment and across subjects. Taken together, the results of the Kalman filtering studies suggest that this technique would be a good candidate for brain signal improvement, in order to get a more accurate estimation of the level of mental workload and engagement of the operator.

7.3 Perspectives

7.3.1 Extending the conceptual framework of behavior : inattentional deafness/blindness

The understanding of the neural substrates underlying executive failures and their measure with brain imaging techniques represents the biggest challenge towards the understanding of one's attentional and process sets, and therefore of the behavior. During this PhD project, we investigated some of the possible executive failures, linked mainly to disengagement mechanisms (mind wandering, overload). From these studies and from the existing scientific literature, we proposed a conceptual framework for the understanding of human behavior (see §7.1), based on the notion of mental workload and engagement, explained by a triad of concepts (task demands, resources, and intentional set). To extend this framework, it would be interesting to assess the correlates of the other potential causes of disengagement (fatigue, stress ; see chapter 7.1). In particular, high levels of engagement also leads to executive failures such as attentional tunneling, eventually leading to **inattentional deafness or blindness**. In this context, the study of a larger range of mental states could add complementary evidence to support or to dispute our conceptual framework. It would also extend the knowledge about the neural basis underlying all types of executive failures and therefore facilitate their detection.

The extension of brain imaging assessment of the operator's state during more types of executive failures is planned for the continuity of the work presented in this PhD project. In collaboration with the Neuroergonomics laboratory of Dr. Daniel Callan at Osaka University, an experiment was conducted to study the neural correlates of inattentional deafness and blindness in fMRI. In the experiment, we used a task similar to the one used in Callan et al. (2011), i.e. a simulated aerobatics flight, where participants were asked to pass through gates. We played both auditory and visual alarms during the experiment. The high task demands placed on the operators during the experiment resulted in alarm misses. The participants also performed a control task in which they were asked to fly straight, without passing through the gates. Although the fMRI data have not been analyzed yet for this study, the behavioral results from the 22 participants exhibit a mean inattentional deafness rate (missed auditory alarms) or 36% (SD:6%) and a mean inatten-

tional blindness rate (missed visual alarms) of 31% (SD:3.8%). This task will allow us to study the neural correlates of alarm misses during simulated flight. In particular, we will be able to identify the neural networks activated during periods of alarm misses, in order to conclude on the existence of a specific intentional set during those executive failures, in accordance with our conceptual framework.

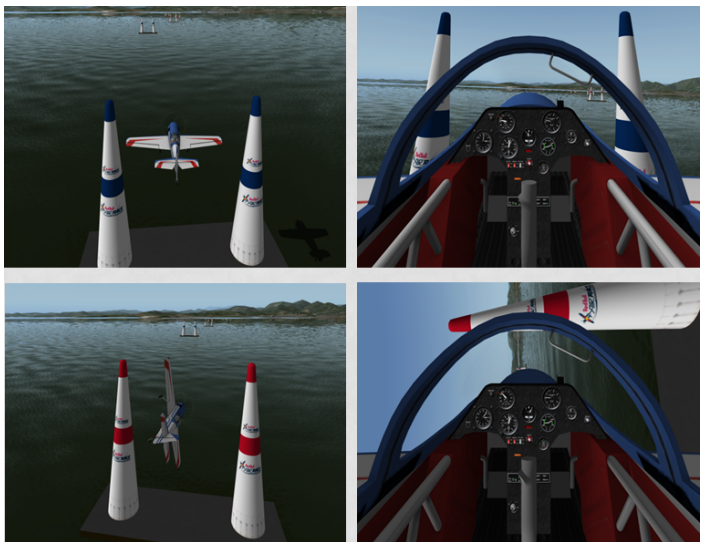


Figure 7.2 – *Screen capture of the task used in the fMRI experiment to study inattentional deafness and blindness (cf. Callan et al. (2011)). The participants controlled the plane with a joystick and performed a complex and engaging piloting task : they had to fly through gates with the correct angle (horizontally for the blue gates, vertically for the red ones). Auditory alarms as well as visual alarms (displayed on the instrument panel) were played during the task.*

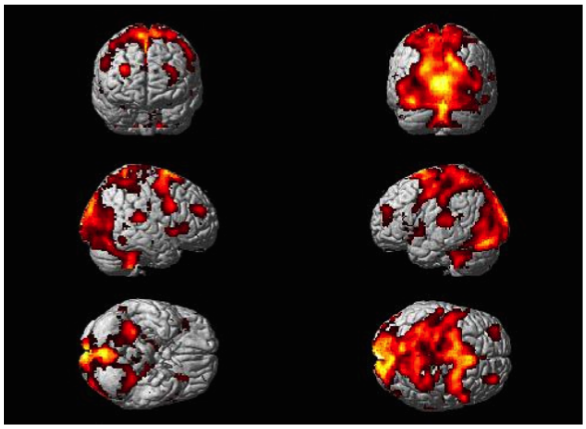


Figure 7.3 – *Preliminary results of the inattentional blindness/deafness experiment in fMRI from one test subject. The activation map shows the effect of the task alarm response vs the control task alarm response, showing the implication of bilateral motor cortex as well as the visual cortex and executive regions.*

In order to fulfill the objectives of Neuroergonomics, another prospective study is the experimentation of the executive failures in real flight conditions. To this end, future investigation will include real flight measurements of brain activity to refine our understanding of the attentional dynamics. Such an experiment is currently in progress (see Fig.7.4), consisting of a 1 hour flight in a light aircraft (*TB20*) with different levels of dif-

ficulty (normal navigation vs simulated emergency landing). Auditory alarms are played during the flight. The participant is asked to push a button located on the stick when he/she hears the alarms. EEG measurements are collected to assess the cognitive state of the participants. So far, 6 participants have participated in the experiment, and they missed between 5 and 25 alarms during the whole flight, suggesting that this experiment will allow us to investigate the neural correlates of inattentional deafness in real flight conditions.



Figure 7.4 – *Inattentional deafness experiment during real flight. Auditory alarms were triggered during different flight conditions. A portable dry EEG was used to measure pilots' brain activity.*

7.3.2 Modeling human behavior

The conceptual framework supported by the results of our studies and presented in section 7.1.2 could be used to describe a computational model of the performance, taking into account task demands, operator resources, and the operator intentional set to explain the dynamics of engagement and disengagement. The model proposed in Fig.7.5 describes a preliminary implementation of these contributions with a systemic approach. We included the contributions of the intentional set (composed of the attentional and process sets) to behavior. The interdependence of the components of the intentional set, constantly updated according to task execution, is represented by the large gray arrows. The second set of arrows, in black, represents the functional interconnections of these blocks, defining the inputs and outputs of each of them. Taken together, these elements could be used to define a dynamical model of executive control (i.e. the mechanisms of resource allocation, see chapter 7.1). In particular, one has to notice that the model proposed by Fig.7.5 relies on two feedback loops : one centered on the individual and one centered on the task itself. This property is in line with the theory of work previously proposed in Leontiev (2014) or Dejours (1995), or with the objectives of Neuroergonomics (see chapter 7.2).

When relying on this functional model to predict a level of resources allocated for

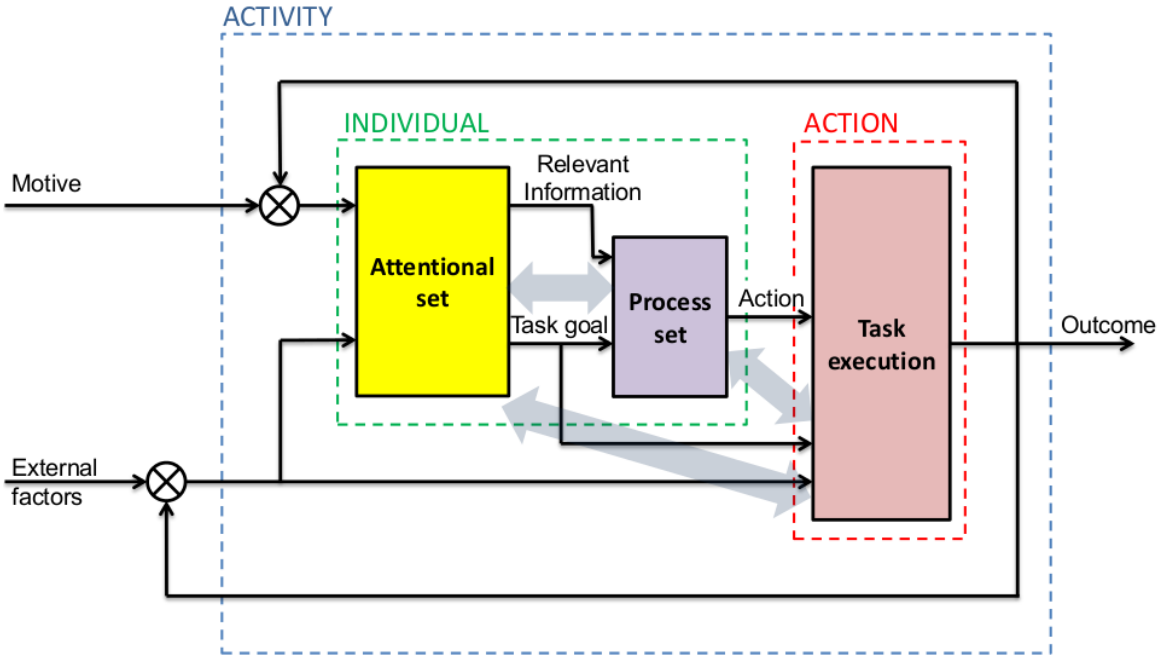


Figure 7.5 – *Systemic model of activity, involving attentional set, process set and task execution, three interconnected functions (represented by the large blue arrows). The black arrows represent the inputs and outputs of each of these three functions. The model relies on two feedback loops : the first one (top) centered on the individual, the second one (bottom) centered on the task itself.*

a specific task (see Fig.7.6 for the computational model), such an implementation could allow us to explain thoroughly the relationship between task demands and resources, by considering the goals of the operator. Interestingly, the knowledge brought by this computational model would allow us to predict which type of resource allocation will occur (see Fig.7.7), which was one of the major limitations of the concept of mental workload (see Fig. 1.3 in §1.2). In addition of being a major supporting evidence of the existence of the intentional set, the application of such a computational model to realistic situations would help us anticipate potential human errors, a critical challenge for the neuroergonomics approach.

7.3.3 Towards a BCI to assist the operator

During this PhD project, we developed and tested two signal processing techniques to improve brain signals : the MACD and the Kalman filter. Those filters managed to improve the estimation of the mental workload and of the engagement of the operator conducted from brain signals, but several challenges were faced concerning the tuning of the filters parameters and their practical use. The investigation of online **adaptive filters**, whose parameters would change dynamically over time, is left for future investigation. In particular, this approach could lead to better stimulus onset detection (in the case of the MACD

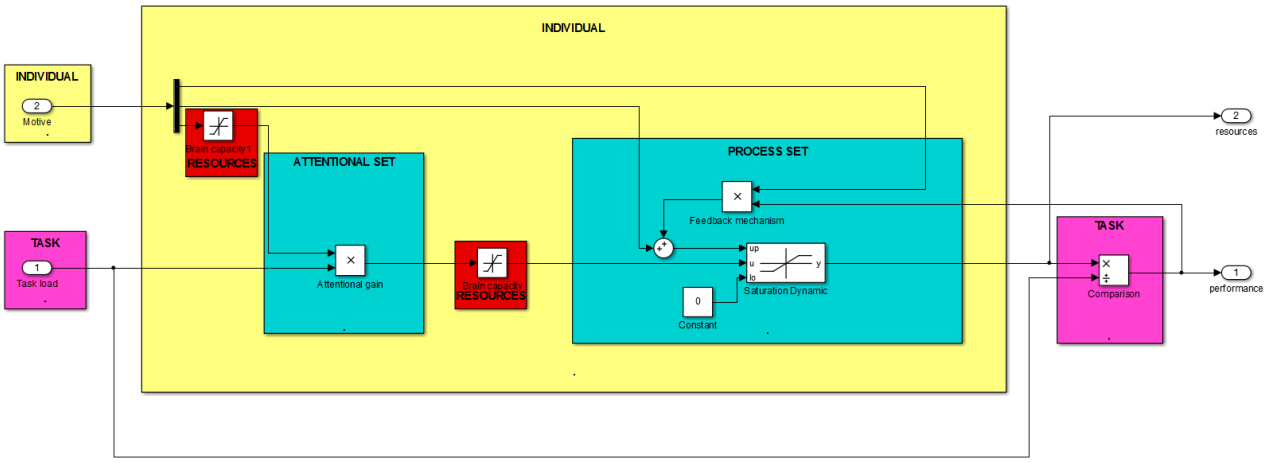


Figure 7.6 – *Matlab ®Simulink representation of the dynamical model of behavior proposed in Fig. 7.5. The magenta parts stand for the characteristics of the task, whereas the yellow parts stand for the characteristics of the individual. They include the parameters modeling the resources of the operator (in red) as well as his/her intentional set (in turquoise).*

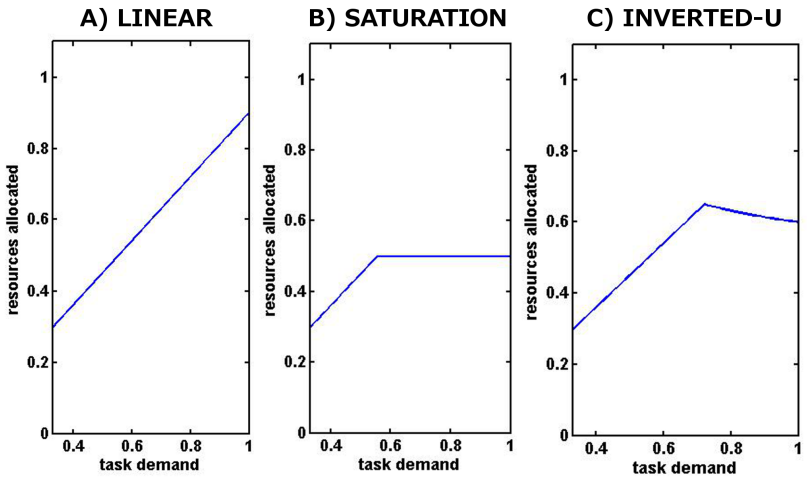


Figure 7.7 – ***Simulated level of resources allocated in function of task demands obtained for varying intentional set properties (determining the engagement of the participant) and brain capacity (preliminary results).***
A) When the task is within the intentional set and brain capacity, we obtain a linear positive relationship.
B) When the task is within the intentional set but above brain capacity, we obtain a saturation curve.
C) When the task is without the intentional set, we obtain an inverted-U curve.

filter) or reduced calibration time and improved results (in the case of the Kalman filter). It is suggested that such methods could be implemented in software platforms as tools to design BCIs that are usable for real-world conditions.

Another point left for further investigation is the **transferability** of the results to other subjects and other sessions (which would allow for a one-time calibration). Indeed, our results showed that both the MACD filter and the Kalman filter reduced the variability of

the results across subjects and trials, which suggests that they could be used as a mean to increase the usability of fNIRS-based BCI and a potential solution to the BCI Illiteracy problem. An experiment involving the test of a BCI on a large panel of subjects would be needed to state on this point. In particular, the usability of a fNIRS-based BCI could also be made more consistent across subjects with the consideration of physiologically more fundamental variables than the HbO_2 and HHb levels. Processing methods, based on for instance the inversion of the balloon model (Buxton et al. 1998), could be developed to estimate the levels of the Cerebral Blood Flow (CBF) and the Metabolic Rate of Oxygen ($CMRO_2$).

Finally, the simultaneous use of the MACD filter and the Kalman filter could be an efficient way to improve the signal in real-time, without any information concerning the task executed (see Fig.7.8). The estimation of engagement periods by the MACD filter coupled with the improvement of the estimation of the engagement-related hemodynamic changes by the Kalman filter could form the cornerstones of an idiosyncratic attentional state estimator.

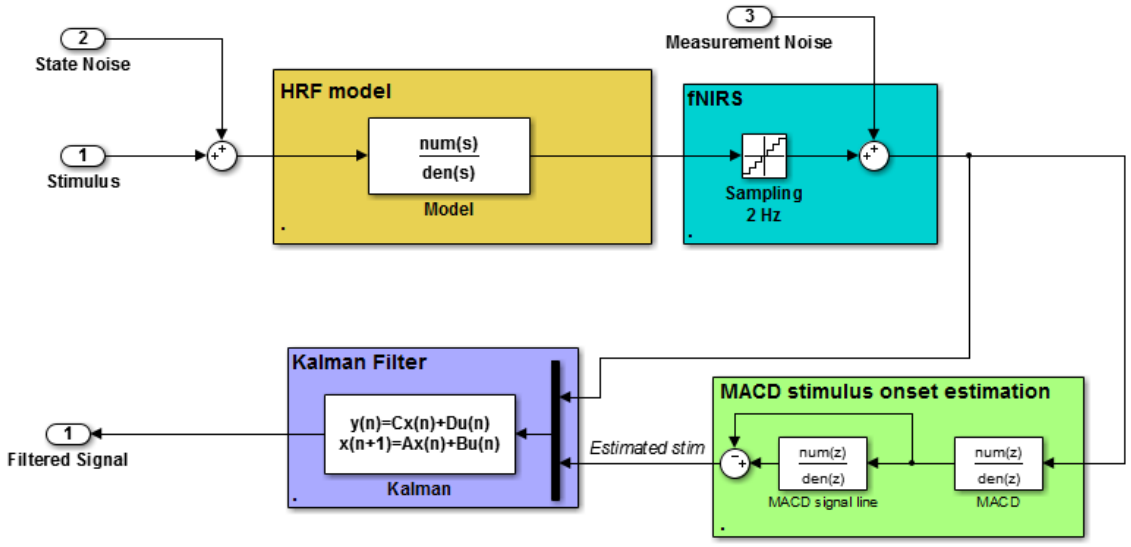


Figure 7.8 – *Simultaneous use of the Kalman filter and the MACD filter to improve signal in real-time with no information about the task. The MACD filter is used here to provide the onset estimation necessary for signal improvement. The Kalman filter is used to improve the engagement-related hemodynamic response.*

The identification of the neural structures underlying attentional dynamics developed on one hand and the improvement of brain signal with innovative processing techniques developed on the other hand pave the way to the **implementation of a multimodal BCI to monitor the attentional state** of an operator. According to the BCI design process described in section 3.2, the only missing step to reach this goal is the implementation of **counter-measures** and/or **neurofeedback**. With this approach one could design a closed-

loop system, in which the state of the operator would be taken into account to improve human-machine interaction.

In addition, further investigation is still needed regarding the robustness of the neural substrates and signal processing techniques investigated during this project to the three main challenges of BCI design : Information Transfer Rate, BCI Illiteracy, and Asynchronicity. Finally, if we want to build a BCI to assist an operator during his/her work, a significant challenge also concerns the integration of the brain imaging techniques in the workspace (e.g. the cockpit) with minimum intrusiveness. Technological development is still required to develop embarked devices (for instance an EEG or a fNIRS installed directly inside a helmet).

General conclusion

This PhD project was driven by the desire to understand the causes of human errors in operational situations. This problematic led to questioning the neural mechanisms responsible for the executive failures, and to proposing a larger conceptual framework to add the notion of engagement to the concept of mental workload. The studies presented in this dissertation put in relief the necessity to adopt a double rationality to understand and monitor human action : one centered on the task and one on the individual. Indeed, understanding the executive failures requires the knowledge of the characteristics of the task being performed by the operator, but also the knowledge of his/her current state at a given time, coupled with his/her goals and motivations. On the same basis, the development of technical procedures of signal processing and use should bear in mind the characteristics of both the task and the individual.

In line with the Neuroergonomics approach, we identified that brain imaging techniques (fNIRS in particular), in complement with other operator state assessment techniques (subjective measures, psycho-physiological measures), carry enough information to account for the observed behavior, as they inform on the causes of resource allocation and the attentional state of the individual. Thus, only the simultaneous consideration of the ergonomics and cognitive neuroscience approaches can lead to a proper explanation of human behavior.

Appendices

Kalman Filter

A

The Kalman filter is based on the estimation of the dynamics of a system, generically described as presented on Fig. A.1.

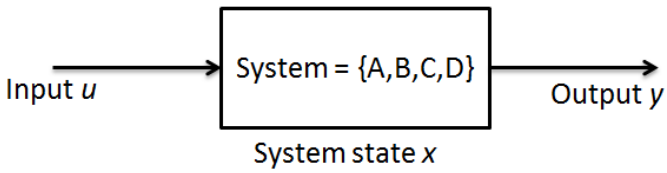


Figure A.1 – *State-space model. The system is defined by its inputs u, its outputs y, and by its state x. Its dynamics is defined by the four matrices A, B, C, D, according to equation A.1.*

The system is described by a **state-space model**, defined by four matrices A, B, C, D, describing the dynamics of the system according to equation A.1.

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases} \tag{A.1}$$

In this equation, u represents the input(s) of the system and y its output(s). x is a vector of variables describing the state of the system at a given time, called **state variables**. The number of state variables required to thoroughly describe the state of the system defines the **order** of the system. Thus, the first equation of A.1 describes the evolution of x (i.e. $\dot{x}$, the time derivative of x), and therefore the **temporal dynamics** of the system. The second equation links the state of the system to its measured output(s) y.

However, in practice, random processes alienate both the evolution of the system (whose state is influenced not only by the inputs and state of the system, but also by external factors) and the measures done to estimate its outputs. The consequence of these noises is that the state of the system x which is perceived when looking at $y_{measured}$ is erroneous. To take into account these random processes, the model proposed in A.1 is generally extended :

$$\begin{cases} \dot{x} = Ax + Bu + w \\ y = Cx + Du + v \end{cases} \tag{A.2}$$

This model includes two vectors of random noises, w and v. On one hand, w represents the unknown random signals that perturb the state equation (i.e. the temporal dynamics of the system), and is called **state noise**. On the other hand, v represents the random signals perturbing the measurements of the outputs of the system, and is called **measurement noise**. The principle of the **Kalman filter** relies on the knowledge of the inputs (u) and outputs (y) of the system to produce an estimate of the system's real state state (noted $\hat{x}$), taking into account the random noises affecting its dynamics. The state estimation can eventually be used to produce an estimate of the outputs independent from the noises w and v (noted $\hat{y}$). The performance of such a state estimator is generally evaluated by considering the estimation error $e = x - \hat{x}$. The definition of a Kalman filter requires three assumptions :

- The system must be **detectable**, which means that every mode of the system that cannot be observed when looking at its outputs must be stable.
- The two signals w (state noise) and v (measurement noise) are considered centered gaussian white noises, i.e. $w = \mathcal{N}(0, Q)$ and $v = \mathcal{N}(0, R)$.
- R must be invertible (i.e. we assume that there are as much sources of noise as measurements)

Provided these assumptions, Kalman proposed an estimator of the state of the system described in equation A.2. The dynamics of this **Kalman filter**, whose inputs are u and y, is defined as follows :

$$\begin{cases} \dot{\hat{x}} = (A - K_f C)\hat{x} + (B - K_f D)u + K_f y \\ \dot{P} = AP + PA^t - PC^t R^{-1} CP + Q \\ K_f = PC^t R^{-1} \\ \hat{y} = C\hat{x} + Du \end{cases} \tag{A.3}$$

The first equation of the system A.3 describes the **Kalman filter dynamics**, where K_f is the gain of the Kalman filter. It is obtained by considering the fact that the Kalman filter is a **non biased estimator** (i.e. the estimation error tends to zero, see Alazard (2005) for demonstration). The second equation is called the prediction equation, and gives the dynamics of P, which represents the **covariance of estimation error** (i.e. $Cov(E)$). It is obtained by considering the fact that the Kalman filter has a **minimal variance estimation**

error (i.e. $Cov(e)$ is minimal). Finally, the third equation links the evolution of the Kalman filter gain K_f to the evolution of P. Therefore, the parameters of a Kalman filter change over time (Kalman filter is a **non stationary system**).

When considering a steady-state Kalman filter, the main parameter influencing the produced estimation is the ratio Q/R (i.e. the ratio between the variances of state noises and measurement noises). Indeed, as Q represents the confidence we have in the state equation (i.e. in the dynamical model of the system), and R represent the confidence we have in the measurements. Therefore, a Kalman filter with $Q/R = 0$ would have absolute confidence in the model rather than in the measurements (as $Q = 0$ and/or $R = +\infty$), and therefore produce an estimation based only on the dynamical simulation of the model. On the contrary, a Kalman filter with $Q/R = +\infty$ would have absolute confidence in the measures (as $Q = +\infty$ and/or $R = 0$), and produce an estimation based solely on them (which would be equivalent to classical linear filtering). In practice, the choice of Q and R values defines different Kalman **tunings**, producing different estimates.

Kalman filter for EEG - Classification results

B

These tables present the performance of the LSPC classifier trained to classify sound vs. silent trials in all the conditions presented in Callan et al. (2015) (see section 6.3.1).

	Platform off			
	Accuracy (%)	FP (%)	FN (%)	d′
Pre-ICA	78.3 (0.33)	19.8 (0.20)	23.6 (0.65)	1.69 (0.01)
ICA	81.0 (0.49)	21.4 (0.99)	16.6 (0.73)	1.59 (0.07)
Pre-ICA + Kalman Filter	77.8 (0.55)	20.2 (1.05)	24.2 (1.94)	1.68 (0.07)
ICA + Kalman Filter	83.4 (0.27)	17.2 (0.53)	16.0 (0.60)	1.90 (0.04)
	Platform on			
	Accuracy (%)	FP (%)	FN (%)	d′
Pre- ICA	71.6 (0.56)	35.4 (1.52)	21.4 (0.60)	0.75 (0.08)
ICA	73.7 (0.52)	26.8 (1.91)	25.8 (1.75)	1.26 (0.12)
Pre-ICA + Kalman Filter	72.8 (0.53)	36.6 (1.58)	17.8 (1.94)	0.69 (0.08)
ICA + Kalman Filter	73.1 (0.8)	28.4 (2.61)	25.4 (1.79)	1.18 (0.16)

Table B.1 – ***Classification performance for the platform condition on the test data.*** *The standard error is given in parentheses below each mean value of the 10 nested cross-validation iterations. (FN= False Negative Rate ; FP= False Positive Rate)*

	Biplane engine off			
	Accuracy (%)	FP (%)	FN (%)	d′
Pre-ICA	77.4 (1.01)	20.2 (2.72)	25.0 (2.20)	1.75 (0.21)
ICA	78.5 (0.72)	20.0 (1.58)	23.0 (0.68)	1.71 (0.11)
Pre-ICA + Kalman Filter	77.0 (0.97)	14.6 (1.16)	31.4 (2.53)	2.13 (0.1)
ICA + Kalman Filter	81.1 (0.35)	10.0 (1.19)	27.8 (1.35)	2.65 (0.18)
	Biplane engine on			
	Accuracy (%)	FP (%)	FN (%)	d′
Pre- ICA	66.1 (0.75)	38.0 (1.23)	29.8 (1.31)	0.61 (0.06)
ICA	77.3 (0.45)	23.6 (1.26)	21.8 (0.81)	1.45 (0.08)
Pre-ICA + Kalman Filter	65.3 (0.96)	35.4 (2.78)	34.0 (2.40)	0.78 (0.17)
ICA + Kalman Filter	79.2 (0.39)	22.6 (1.12)	19.0 (1.16)	1.51 (0.07)

Table B.2 – ***Classification performance for the biplane condition on the test data.*** *The standard error is given in parentheses below each mean value of the 10 nested cross-validation iterations. (FN= False Negative Rate ; FP= False Positive Rate)*

ICA vs. Pre-ICA								
Overall	Platoff	Platon	Engoff	Engon	Platoff+ Kalman	Platon+ Kalman	Engoff+ Kalman	Engon+ Kalman
***		**		**		**		**
Off vs. On								
Overall	Platform	Engine	Platform Kalman	Engine Kalman	Platform ICA	Engine ICA	Platform ICA + K	Engine ICA + K
***	**	**	**	**	**		**	*
Kalman vs. No Kalman								
Overall	Platoff	Platon	Engoff	Engon	Platoff ICA	Platon ICA	Engoff ICA	Engon ICA
*					**		*	**

Table B.3 – ***Statistical significance of various contrasts using nested cross-validation to evaluate model accuracy.*** *Significance (one-tailed) is determined using the Wilcoxon sign-rank test. * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$.*

Kalman Filter design for fNIRS

C

The functional model used to design the Kalman filter for fNIRS signal was inspired by the Hemodynamic Response Function (HRF) proposed by Boynton et al. (1996). For a first kalman approach, this model is simple enough to allow quick computation. This model assumes a third order impulsional response to stimulation, as shown on Fig. C.1 :

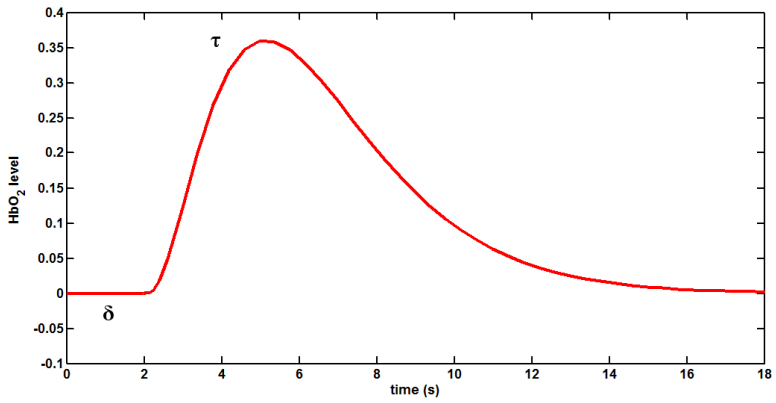


Figure C.1 – ***Boynton model of hemodynamic response for*** HbO_2**.** *The model order is three, and depends on two parameters : δ represents the delay between stimulation and HbO_2 increase (units for HbO_2 levels are arbitrary) ; τ influences the time-to-peak delay. Typical values are $\delta = 2s$ and $\tau = 1.5s$. The same model with reduced amplitude and opposite response direction can be used for HHb response.*

The model has the following transfer function :

$$HRF(p) = \frac{\tau^3 e^{-\delta p}}{(p+\tau)^3} \tag{C.1}$$

As shown in equation C.1 ad in Fig.C.1, the response shape depends on two parameters : δ represents the delay between stimulation and HbO_2 increase ; τ influences the time-to-peak delay. Typical values that were chosen here are $\delta = 2s$ and $\tau = 1.5s$. In accordance with the Kalman filter principles (see appendix A), we took into account the state noise w and measurement noise v affecting the process. As shown on Fig. C.2, we

chose to represent the state noise as a perturbation affecting the stimulus (i.e. the input of the model). This choice led us to consider that state noise represents a stimulus perception (or internalization) bias.

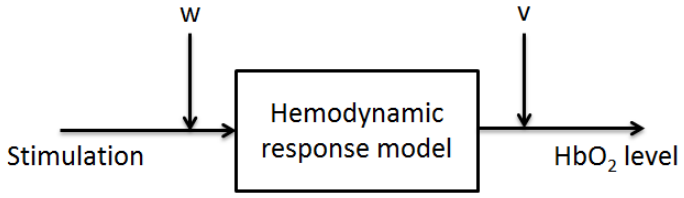


Figure C.2 – ***Proposed approach for Kalman filtering of fNIRS.*** *The state noise is considered as a perturbation of the input of the system.*

The perception bias perturbing the stimulus is noted b. In the nominal model, Kalman filter assumptions impose that $b = w$, where w is the state noise following a gaussian centered distribution. This model, in addition to the choice of a Q/R value, allowed us to design a **Kalman filter** for fNIRS signal improvement (see appendix A for details). The inputs of the Kalman filter are the stimuli onsets and the fNIRS raw signal, as shown on Fig. C.3.

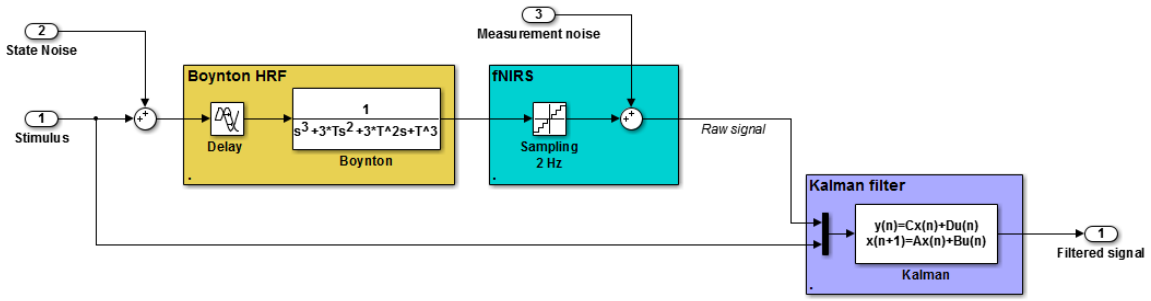


Figure C.3 – ***Signal processing chain including the Kalman filter.*** *The model includes the HRF model (see Fig.C.1) that links stimulation to HbO_2 changes, the fNIRS model adding the measurement noise, and the Kalman filter, whose inputs are the stimuli onset and the raw fNIRS signal.*

One of the main limitations of this approach is the fact that the stimulus perception bias has to be centered (i.e $b = 0$ on average), which can be erroneous when the subject sustainably disengage from the task and doesn't pay attention to the stimuli (see chapter 7.1). To take this element into account, we built a second model, which is an augmentation of the nominal model, and in which $\dot{b} = w$. Thus, as the first derivative of b follows the gaussian centered distribution, it is still possible to design a Kalman filter, without assuming that b is null on average. This augmented model, along with the value of Q/R, allowed the computation of the **augmented Kalman filter** for fNIRS signal processing. For both filters, the value of the ratio Q/R was fixed according to an optimization process (see section 6.3.2).

Résumé en français

Introduction

L'avion est aujourd'hui considéré comme le moyen de transport le plus sûr. Toutefois, les innovations technologiques qui ont conduit à l'amélioration de la sécurité des avions rendent de plus en plus évidente la présence d'une responsabilité humaine (estimée à 70-80%) dans les accidents aériens. Plus largement, les erreurs dans l'opération des systèmes complexes présentent souvent une part de responsabilité humaine. Les différents modèles proposés pour expliquer les causes de tels accidents mettent en avant l'importance de l'état attentionnel des opérateurs dans le risque d'erreur humaine. En effet, l'opération des systèmes complexes se fait bien souvent dans un environnement dynamique, où les décisions doivent être prises de manière rapide, conduisant souvent les opérateurs vers des erreurs. Malgré ce risque, les nombreuses actions positives de l'opérateur humain ainsi que la complexité des stratégies opérationnelles rendent impossible l'automatisation complète de ces systèmes. Il apparaît donc comme primordial de mieux comprendre les mécanismes psycho-physiologiques et cérébraux à l'origine de l'erreur humaine, et de trouver des moyens d'estimer l'état attentionnel de l'opérateur. Les neurosciences cognitives peuvent apporter une contribution de première importance à l'ergonomie de tels systèmes. En particulier, le domaine de la neuroergonomie postule que l'étude des mécanismes cérébraux liés à l'opération de tels systèmes permet de proposer à la fois des recommandations mieux adaptées aux situations opérationnelles complexes, ainsi que des solutions d'assistance à l'opérateur.

L'objectif de cette thèse, qui adopte une approche de neuroergonomie, est d'étudier la pertinence des mesures cérébrales dans la compréhension et l'estimation de l'état attentionnel, ainsi que la possibilité de leur utilisation en situations réelles pour assister l'opérateur. Le premier chapitre de ce projet est consacré à la présentation des concepts généraux permettant d'analyser le comportement et la performance humaine en conditions réelles. Les chapitres deux et trois présenteront ensuite les apports de la neuroergonomie à l'amélioration de l'interaction homme-machine, et porteront une attention particulière aux moyens de mesures utilisables pour évaluer le comportement de manière objective, ainsi que leur utilisation pratique dans le cadre des interfaces cerveau-machine. Enfin, les trois derniers chapitres présenteront les études menées dans le cadre de ce projet de thèse. Celles-ci portent en particulier sur l'évaluation de multiples moyens de mesures psycho-physiologiques et cérébraux à l'estimation de l'état attentionnel dans des conditions de demandes élevées (chapitre 4) ou faibles (chapitre 5). Le chapitre 6 présentera

enfin les différentes méthodes d'utilisation du signal cérébral en conditions réelles, en vue de la mise en place d'interfaces cerveau-machine. Nous discuterons en conclusion l'ensemble de ces résultats, afin de proposer un cadre théorique pour l'évaluation et la prédiction du comportement de l'opérateur, ainsi que des solutions pour la mise en place de moyens d'assistance à l'opérateur.

Etat de l'art

7.1 Chapitre 1 : Demandes de la tâche et performance

Dans ce chapitre, nous expliquons le lien qui existe entre les demandes d'une tâche et la performance observée. En particulier, nous développons les différents facteurs et mécanismes entrant en jeu dans l'allocation des ressources cognitives. Enfin, nous expliquons l'impact des variations des demandes d'une tâche sur l'allocation des ressources cognitives, en mettant en relief l'effet de niveaux de demande extrêmes.

Le chapitre présente tout d'abord le concept de **charge mentale**, et son lien avec la notion de ressources mobilisées pour effectuer une tâche. A travers des considérations sur les **mécanismes du contrôle cognitif** mis en relation avec la **performance** d'un individu, nous mettons en avant l'ambigüité du concept de charge mentale. Celui-ci présente en effet un bon potentiel descriptif, puisqu'il permet d'analyser une performance en terme de ressources mobilisées. Toutefois, la seule considération du concept de charge mentale ne permet pas d'expliquer ni de prédire la manière dont sont affectées les ressources.

La suite du chapitre met en avant la notion d'**engagement**, complémentaire du concept de charge mentale, pour expliquer la manière dont les ressources sont mobilisées, et les facteurs influençant cette mobilisation. Les substrats neuronaux et mécanismes cognitifs reliés à cette notion sont présentés. Cette notion d'engagement est particulièrement importante lorsque l'opérateur est soumis à des niveaux de demande très faibles, ou au contraire très élevés. La considération du comportement d'un individu montre en effet que d'une part, les hauts niveaux de demande peuvent conduire à des phénomènes de **tunnélisation attentionnelle**, amenant l'opérateur à négliger des informations critiques. De même, ils peuvent également conduire à des phénomènes de **persévération** ou au contraire de **désengagement** de l'opérateur face aux demandes de la tâche. D'autre part, les faibles niveaux de demande donnent parfois lieu à un désengagement de l'opérateur dans des états de rêverie (ou **mind wandering**).

Le chapitre conclut alors sur la nécessité de prendre en compte conjointement le niveau de charge mentale imposé à l'opérateur ainsi que son niveau d'engagement (à mettre en relation avec ses propres buts et motivations) pour évaluer sa performance.

7.2 Chapitre 2 : Neuroergonomie

Ce chapitre introduit le domaine de la **Neuroergonomie** (Parasuraman 2003), qui se définit usuellement comme l'application des neurosciences à l'ergonomie. Dans un premier temps, le chapitre présente les objectifs de ce domaine. En particulier, nous mettons en avant le bénéfice que peut apporter la connaissance des mécanismes neuronaux à la base de la dynamique attentionnelle à l'amélioration des performances d'un opérateur, en relation avec le chapitre 1. Nous insistons également sur l'approche de la neuroergonomie, basée sur un **gradient d'expérimentation** depuis des études menées en laboratoire, en général peu représentatives de la réalité mais dans lesquelles tous les paramètres sont contrôlés, vers des tâches plus écologiques, fidèles à la réalité, mais dans lesquelles les paramètres expérimentaux sont moins contrôlés. Cette approche expérimentale, couplée à l'utilisation de **mesures psycho-physiologiques et cérébrales** pour évaluer le niveau de charge mentale et d'engagement de l'opérateur, a le potentiel de fournir des outils pour améliorer les performances d'un opérateur.

La suite du chapitre se concentre sur les mesures utilisables en neuroergonomie pour évaluer le niveau de charge mentale et d'engagement. Cette partie insiste particulièrement sur les **critères** à prendre en compte pour juger de la qualité d'une mesure, et évalue un éventail large de mesures allant des mesures **subjectives** (questionnaires) aux mesures **objectives comportementales** (analyse du niveau de performance), **psycho-physiologiques** (Oculométrie, Electrocardiographie –ECG–,...) et **cérébrales** (avec un focus particulier sur l'Electroencéphalographie –EEG– et la Spectroscopie fonctionnelle en proche infrarouge –fNIRS–).

Pour obtenir l'évaluation la plus précise possible du niveau de charge mentale et d'engagement de l'opérateur, nous concluons sur l'**intérêt des mesures cérébrales**, qui renseignent sur l'état attentionnel de l'individu par l'évaluation du recrutement de diverses régions cérébrales. Ces mesures représentent un atout majeur de la Neuroergonomie. En outre, la **fNIRS**, dont l'utilisation est encore moins fréquente que l'EEG, semble offrir une bonne évaluation des niveaux de charge mentale et d'engagement, qu'elle soit utilisée seule ou en complément d'autres mesures. Elle fera donc l'objet d'une attention plus particulière dans ce projet de recherche.

7.3 Chapitre 3 : Interfaces Cerveau-Machine

Dans la continuité de l'approche neuroergonomique, ce chapitre présente les **Interfaces Cerveau Machine** (ICM), qui sont l'un des outils majeurs d'application de la neuroergonomie à la conception des systèmes. Définis tout d'abord d'une façon générale comme des systèmes dont l'une des entrées est un signal cérébral, les ICM sont ensuite décomposées en deux catégories selon leurs objectifs. Le chapitre présente ainsi l'opposition entre des **ICM actives** d'une part, dans lesquelles le signal cérébral est transformé en commande et qui visent donc à contrôler un système, et les **ICM passives** d'autre part, dont le but n'est pas de commander un système mais d'utiliser une estimation de l'état de l'opérateur afin d'adapter le fonctionnement du système à celui-ci (par l'utilisation de contre-mesures cognitives ou par l'adaptation du niveau d'automatisation, par exemple). Nous insistons en particulier sur l'intérêt des ICM passives dans le cadre de la neuroergonomie, afin d'assister l'opérateur pour diminuer son risque d'erreurs.

Nous introduisons ensuite les enjeux fondamentaux de la conception des ICM, et identifions deux étapes cruciales dans leur conception : l'étape de **traitement du signal**, ainsi que celle de **conception d'une loi de commande**. Dans un premier temps, nous présentons les enjeux du traitement du signal, visant à isoler l'activité cérébrale utile. Cette partie repose à la fois sur la description des bruits et perturbations présentes dans le signal cérébral, ainsi que sur la présentation des diverses techniques de traitement utilisables pour les rejeter. Dans un second temps, nous détaillons les moyens de transformer l'activité cérébrale d'intérêt en une commande pour le système, en s'attardant plus précisément sur les **méthodes d'apprentissage** permettant d'associer une activité cérébrale à un état mental.

Enfin, nous présentons les principaux verrous scientifiques à lever dans l'objectif de concevoir des ICM utilisables en situation opérationnelle. Ceux-ci concernent principalement l'utilisabilité des signaux pour en extraire une quantité maximale d'information (maximisation du **débit d'information**), le transfert de l'initiative de l'envoi de commandes vers l'opérateur plutôt que vers la machine (création d'**ICM asynchrones**) et la limitation du nombre d'utilisateurs pour qui l'ICM est non performante (problème de l'**illétrisme des ICM**). Nous concluons alors sur la nécessité de mettre en place des techniques innovantes de **traitement et d'analyse du signal** pour résoudre ces défis.

Contribution

Les chapitres précédents mettent en avant la nécessité d'évaluer le comportement d'un individu non seulement en terme de **charge mentale**, mais aussi en terme d'**engagement**. En particulier, nous avons vu que leur évaluation nécessite souvent une **approche neuroergonomique** graduelle, partant des tâches de laboratoire très contrôlées mais peu écologiques vers des tâches réalistes, peu contrôlées mais très écologiques. Dans ce cadre, nous avons vu que la mise en place de mesures cérébrales, et en particulier la fNIRS, peut apporter de précieuses informations sur l'état attentionnel de l'opérateur, malgré le fait que leur implémentation en situation réelle se heurte à de nombreux challenges, notamment en terme de traitement du signal.

La suite de ce résumé présente les contributions réalisées dans le cadre du projet de thèse. En adoptant une démarche de neuroergonomie, nous avons évalué l'état de l'opérateur soumis à des demandes extrêmes, à l'aide de multiples mesures psychophysiologiques et cérébrales. Ce travail nous a mené à nous focaliser sur les états de **surcharge mentale** et de **mind wandering**. Enfin, nous nous sommes penchés sur la mise en place de **méthodes de traitement du signal** en vue de l'utilisation en conditions réelles du signal cérébral.

7.4 Chapitre 4 : Surcharge mentale

Ce chapitre présente une étude dont les résultats sont publiés dans Durantin et al. (2014b) et Gagnon et al. (2012), et qui se concentre sur l'évaluation des conséquences comportementales, psycho-physiologiques et neuronales de la surcharge mentale.

Durant une tâche de pilotage simulé sur ordinateur, les douze participants de l'étude devaient appliquer des instructions et suivre un avion cible (choisi parmi cinq). Durant l'expérience, le niveau de charge mentale imposé au participant était manipulé à travers deux facteurs : le niveau de **difficulté** de la tâche (correspondant à la difficulté de pilotage de l'avion –facile ou difficile–), ainsi que le niveau de **mémoire** (les instructions devaient être mémorisées –ou non– avant leur application). Tout au long du protocole, l'état attentionel des participants était mesuré à l'aide d'un questionnaire subjectif d'évaluation de la charge de travail (le NASA Task Load Index), ainsi qu'à l'aide d'un ECG (pour l'évaluation de l'activité du système nerveux autonome) et d'une fNIRS (pour l'évaluation de l'activité du cortex préfrontal).

Les résultats subjectifs montrent que la charge mentale perçue augmente bien avec la manipulation des paramètres (la difficulté de pilotage ainsi que l'utilisation de la mémoire de travail rendent donc bien la tâche plus complexe). Cette augmentation de la charge mentale est liée à une diminution de la performance, et à une augmentation du niveau d'oxygénation mesurée dans le cortex préfrontal dorsolatéral (DLPFC) ainsi qu'à une augmentation du ratio LF/HF, indicateur de la variabilité du ryhtme cardiaque. Ces résultats sont en accord avec des études précédentes (Ayaz et al. 2012a, Hjortskov et al. 2004). En revanche, le niveau de difficulté le plus élevé (lorsque la difficulté de pilotage est élevée et que les instructions doivent être mémorisées) conduit à une diminution des niveaux d'oxygénation dans le DLPFC et du ratio LF/HF. Les courbes d'activation du DLPFC et du ratio LF/HF en fonction de la difficulté suivent ainsi une forme en **U inversé**. De même, les sujets ont reporté dans les questionnaires la difficulté trop grande de la condition la plus complexe. Cette difficulté ressentie se voit également au niveau de la performance, qui chute lors de la **surcharge mentale**. Ces résultats vont dans le sens de l'existence d'une **stratégie d'allocation des ressources** spécifique lorsque le sujet se trouve en situation de surcharge mentale, qui traduit un **désengagement** partiel face à la difficulté de la tâche.

Sur le plan méthodologique, les résultats de l'expérience confirment la nécessité de l'approche neuroergonomique pour observer la **dynamique de l'engagement** et du désengagement en conditions réelles. En particulier, les résultats observés ici montrent que le comportement d'un opérateur ne peut pas être décrit complètement par la seule considération du niveau de charge mentale qui lui est imposé.

7.5 Chapitre 5 : Mind wandering

Ce chapitre, en opposition avec le précédent, se concentre sur les conditions où la demande de la tâche est **très faible**, et étudie les conséquences comportementales de ce niveau de demande. En particulier, nous nous concentrons sur l'étude du **mind wandering**, qui traduit le désengagement attentionnel alors que l'opérateur commence à laisser ses pensées divaguer. Le chapitre présente les résultats de deux expériences : une première expérience utilise une tâche de laboratoire et vise à tester la capacité de la fNIRS à détecter les instants de mind wandering. La seconde tire parti des résultats de l'expérience et les applique à une situation plus écologique, en utilisant une double mesure de l'état attentionnel : la fNIRS et l'oculométrie. La première de ces deux études a été publiée dans Durantin et al. (2015).

Dans la première expérience, 23 sujets ont été confrontés à un protocole de type **Sustained Attention to Response Task** (SART). Dans ce protocole, les sujets sont assis face à un écran et doivent appuyer sur un bouton chaque fois qu'un chiffre apparaît à l'écran (toutes les 1.5 secondes), sauf si celui-ci est le chiffre "3". La répétitivité de la tâche rend les sujets prompts à automatiser leur comportement face à une telle tâche et à laisser leur attention divaguer (Manly et al. 1999). La conséquence de cela est l'apparition d'erreurs (les sujets appuient sur le bouton alors que le chiffre présenté est le chiffre 3). Durant l'expérience, l'activité du cortex préfrontal était mesurée à l'aide de la fNIRS. Lors de l'analyse, nous avons comparé les niveaux d'oxygénation mesurés au moment des stimuli cibles (le chiffre 3) lorsque le sujet commettait une erreur (associée au mind wandering) ou lorsqu'il n'en commettait pas (associée au focus sur la tâche). Ceux-ci révèlent que les instants de mind wandering étaient précédés d'une augmentation du niveau d'oxygénation mesuré au niveau des optodes médianes du cortex préfrontal (mPFC), traduisant une activation de ces zones. Ce résultat confirme l'implication de ces régions faisant partie du Default Mode Network (DMN) dans le passage vers un état de mind wandering (Christoff et al. 2009), et la capacité de la fNIRS à détecter leur activité. Enfin, nous avons tenté d'utiliser ce résultat afin d'entraîner un **classifieur formel** pour distinguer les instants de mind wandering des instants de concentration, avec des résultats relativement faibles (56% des stimuli étaient classifiés correctement), ce qui laisse supposer qu'une bonne détection des instants de mind wandering pourrait se baser sur une approche multimodale.

Dans la seconde expérience, le même protocole de type SART a été appliqué dans des conditions plus écologiques, inspirées de Shaw et al. (2013). En effet, dans cette expérience, les sujets endossaient le rôle de superviseurs de drônes, et les stimuli utilisés représentaient des drônes volant en formation circulaire, pouvant être potentiellement dangereuse (risque de collision entre les drônes) ou non. Les participants devaient valider le stimulus en appuyant sur le bouton lorsque celui-ci était non dangereux. Au cours du

protocole, l'activité cérébrale et comportementale des participants était mesurée à l'aide d'une fNIRS et d'un oculomètre. Les résultats montrent les mêmes activations du mPFC précédant les instants de mind wandering que dans le première expérience, confirmant le rôle du DMN dans la génération du mind wandering. En outre, cette activation était suivie par une désactivation des aires médianes du gyrus frontal inférieur, précédemment associées au maintien d'une attention soutenue (Koechlin et al. 2000) dans ce type de tâche. Ainsi, ce résultat supporte l'idée d'un désengagement de la tâche de SART consécutif à l'occurrence du mind wandering. Cette hypothèse est également cohérente avec la présence d'un niveau de base de diamètre pupillaire plus élevé durant les épisodes de mind wandering. Ce résultat, mis en relation avec la théorie du fonctionnement du locus cœruleus (Smallwood et al. 2012b), suggère le passage d'un état de concentration vers un état d'exploration.

Les deux études suggèrent que la fNIRS est adaptée à la détection du mind wandering, que ce soit en conditions de laboratoire ou en situations écologiques. En particulier, les résultats appuient la notion de désengagement associée à cet état et soutenue par l'activation d'un **Task Negative Network** (Johnson and Zatorre 2006), et suggèrent tout comme dans le chapitre précédent que l'étude du comportement et de la performance d'un individu ne doit pas se limiter à l'analyse d'un niveau de charge mentale.

7.6 Chapitre 6 : Traitement du signal et ICM

Dans ce chapitre, nous présentons les différentes considérations techniques qui ont été mises en place pendant la thèse en vue de l'application des mesures cérébrales en conditions réelles. Le chapitre s'articule autour de l'implémentation et de la mise en place de deux techniques de traitement du signal : le filtre **Moving Average Convergence Divergence** (MACD) inspiré de l'analyse économique de marché, et le **filtre de Kalman**, couramment utilisé pour le traitement du signal embarqué (en particulier dans le domaine aéronautique et spatial).

Une première étude détaillée dans le chapitre et publiée dans Durantin et al. (2014c) présente l'investigation du **filtre MACD** pour le traitement du signal fNIRS et son utilisation pour la prédiction des instants de stimulation neuronale. En effet, ce filtre issu de l'analyse économique de marché possède deux atouts. Tout d'abord, il possède les caractéristiques d'un filtre passe-bande, ce qui le rend utilisable pour le **nettoyage des signaux fNIRS**. Sa comparaison avec d'autres filtres utilisés classiquement dans le traitement hors-ligne des données (filtres IIR) montre qu'il est un bon candidat pour le rejet en temps réel des artefacts stationnaires. En outre, le filtre MACD peut être utilisé pour prédire des hausses stables dans le signal (c'est même sa fonction principale en économie). Ainsi, en considérant la forme de la réponse hémodynamique suite à une stimulation neuronale, il est suggéré que la méthode de prédiction des hausses stables basée sur le filtre MACD pourrait être appliquée aux signaux fNIRS pour **détecter les instants de stimulation**. Dans cette étude, le test est réalisé lors d'une tâche de mémoire (mémorisation de série de chiffres). Le filtre MACD s'avère efficace pour la détection des instants de stimulation, ce qui suggère qu'il pourrait être implémenté en situations réelles en vue de la mise en place d'un estimateur de l'état (engagé ou non engagé) de l'opérateur.

La deuxième partie du chapitre présente les résultats d'une étude menée en simulateur de vol, en vue de la réalisation d'une **ICM passive** pour **surveiller l'état du pilote** (engagé ou non engagé) ainsi que le niveau de sa charge mentale, à partir de la mesure fNIRS. Durant cette étude, les participants devaient mémoriser puis rentrer dans le pilote automatique des instructions données par la tour de contrôle (plus ou moins difficiles à mémoriser). Les résultats sont publiés dans (Gateau et al. 2015). L'estimation de l'état du pilote, en particulier, est faite à partir du filtre MACD, et montre des résultats prometteurs puisqu'une bonne adéquation existe entre les instants de stimulation (instructions de vol données par la tour de contrôle) et l'engagement du pilote estimé par le MACD. En outre, la classification des niveaux de difficulté (facile ou difficile) des instructions à partir des données fNIRS en temps réel correspond dans 80% des cas à la difficulté réelle. Ces résultats montrent que la fNIRS est un bon candidat pour la mise en place d'ICM en conditions opérationnelles. En particulier, l'utilisation du filtre MACD permet

d'améliorer l'utilisabilité du signal et de disposer en temps réel d'une estimation de l'état d'engagement de l'opérateur.

Les deux dernières expériences présentées dans le chapitre se concentrent sur une autre technique de traitement utilisable pour améliorer le signal cérébral en temps réel : le filtrage de Kalman. Dans un premier temps, nous concevons un **filtre adapté à l'EEG** basé sur la connaissance a priori de la dynamique temporelle du signal EEG suite à une stimulation auditive (Georgiadis et al. 2005a). Le filtre est appliqué pour améliorer le signal EEG collecté en situation de vol réel ou simulé (Callan et al. 2015). Cette étude démontre qu'il est possible d'obtenir une mesure EEG utilisable en temps réel malgré des conditions expérimentales très dégradées (vol réel dans un avion à cockpit ouvert). Dans un second temps, nous concevons un second **filtre de Kalman adapté aux signaux fNIRS**, et basé sur la modélisation de la réponse hémodynamique (Boynton et al. 1996). Nous détaillons la problématique liée à son réglage, et prouvons qu'il est possible d'obtenir un réglage "universel" permettant daméliorer l'utilisabilité du signal fNIRS dans des conditions de laboratoire (tâche de mémorisation de chiffres) ou écologiques (tâche de mémorisation d'instructions de la tour de contrôle en simulateur de vol). Quel que soit le type de mesure utilisée (fNIRS ou EEG), nous concluons donc que le filtre de Kalman représente un outil intéressant pour la neuroergonomie, dans la mesure où il améliore le signal collecté en situations réelles.

Conclusion

7.7 Discussion

L'objectif principal de ce projet de thèse était l'utilisation d'une démarche de neuroergonomie afin de fournir des recommandations et solutions pour l'évaluation de l'état attentionnel en situations écologiques.

Dans un premier temps, on constate que l'ensemble des résultats des expériences visant à explorer le comportement en situations de demandes extrêmes peut s'analyser de manière cohérente en considérant le fait que la **cognition humaine est toujours dirigée vers un but** (goal-oriented cognition). Vus sous cet angle, les résultats de l'expérience sur la surcharge mentale et de celles sur le mind wandering mettent en avant que l'opposition entre les buts de l'opérateur et les demandes de la tâche qu'il effectue peut aboutir à un désengagement au niveau comportemental, résultant de stratégies d'allocation des ressources mesurables au niveau du cortex préfrontal. Ces résultats mettent en avant le rôle prépondérant du cortex préfrontal dans la génération de stratégies en accord avec les buts propres de l'opérateur (Koechlin et al. 2000, Ridderinkhof et al. 2004). Du point de vue de la neuroergonomie, ce résultat est intéressant dans la mesure où il associe les erreurs humaines à des modifications transitoires du fonctionnement du cortex préfrontal (c'est la théorie de l'**hypofrontalité** de Dietrich (2003)). Cette hypothèse est en accord avec la notion de "**intentional set**" (que l'on peut traduire par "stratégies intentionnelles") (Rushworth et al. 2002), définissant la manière dont l'opérateur sculpte l'attention qu'il porte à l'environnement et ses réponses à celui-ci en fonction de ses propres buts et motivations. Considérés comme un ensemble, l'**intentional set**, les **ressources** d'un opérateur, et les **demandes** de la tâche qu'il doit effectuer, forment ainsi un tryptique de concepts qui permettent d'expliquer la charge mentale et l'engagement d'un opérateur, et ainsi sa performance.

En outre, la consistance des résultats obtenus entre les différentes expériences justifie l'approche de la neuroergonomie, et met en particulier en avant la valeur ajoutée que peuvent apporter les tâches de laboratoire à la compréhension du comportement en situations écologiques. Nous avons également vu qu'il était possible de mettre en place des techniques innovantes de traitement du signal (notamment en s'inspirant d'autres do-

maines des sciences) permettant de rendre plus facilement utilisable le signal cérébral en situations réelles. Ces deux avancées représentent deux pierres angulaires de la conception d'ICM pour assister l'opérateur en vue de détecter ces états attentionnels dégradés et d'adapter dynamiquement l'interaction homme-machine.

7.8 Perspectives

Afin d'améliorer notre compréhension du comportement humain en situations écologiques et de **valider** la pertinence de la notion de charge mentale et d'engagement pour son évaluation, d'autres études doivent être menées pour étudier les mécanismes à l'origine d'autres types d'erreurs. En particulier, le phénomène de **surdité ou de cécité aux alarmes** en conditions de charge élevée (Lavie et al. 2014) est intéressant du point de vue de la démarche neuroergonomique. Deux expérimentations, en collaboration avec le laboratoire du Pr. Callan à Osaka (Japon), sont en cours pour l'étudier. Une première, très contrôlée, se base sur une tâche de pilotage acrobatique sur ordinateur, réalisée dans un IRMf. Durant cette tâche, des alarmes auditives et visuelles sont présentées aux participants. L'analyse des données IRMf nous permettra alors d'identifier les corrélats cérébraux des états de surdité de cécité inattentionnelle. Une seconde expérience, en conditions réelles de vol, est en cours pour étudier ce phénomène : les participants devront répondre à des alarmes auditives pendant une heure de vol en avion léger. L'utilisation d'un EEG portable durant le vol permettra l'étude des corrélats de la surdité aux alarmes en conditions réelles.

La prise en compte à la fois des ressources de l'opérateur et des demandes imposées par la tâche, mais aussi des buts et motivations de l'opérateur (à travers la définition de l'intentional set), doit également permettre la mise en place d'un modèle computationnel pour prédire le comportement (et en particulier les niveaux de charge mentale et d'engagement expérimentés par l'opérateur). Plus particulièrement, des simulations numériques réalisées sur un tel modèle devraient pouvoir valider ou invalider la pertinence de cette approche pour expliquer le comportement de l'opérateur. Ces travaux peuvent également permettre d'enrichir la connaissance sur l'origine de la conscience.

Enfin, les différents corrélats de l'état attentionnel ainsi que les différentes méthodes de traitement du signal développées durant ce projet de recherche devraient permettre la mise en place d'une ICM passive en conditions écologiques (basée sur la fNIRS, par exemple) pour assister l'opérateur. Celle-ci devrait être capable d'anticiper les erreurs humaines et de décider d'une intervention à réaliser pour les éviter (mise en place de contre-mesures cognitives, ou adaptation du niveau d'automatisation de la tâche). Dans cette optique, nous pourrons envisager une approche multimodale, basée sur l'utilisation

de plusieurs mesures cérébrales (EEG et fNIRS, par exemple). En particulier, un verrou technologique majeur à lever est l'intégration des différents moyens de mesure et des différentes contre-mesures dans le poste de travail de l'opérateur (par exemple le cockpit).

Bibliography

Abdelnour, A. F. and Huppert, T. (2009). Real-time imaging of human brain function by near-infrared spectroscopy using an adaptive general linear model. *Neuroimage*, 46(1):133–143.

Afergan, D., Peck, E. M., Solovey, E. T., Jenkins, A., Hincks, S. W., Brown, E. T., Chang, R., and Jacob, R. J. (2014). Dynamic difficulty using brain metrics of workload. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3797–3806. ACM.

Aguirre, G., Zarahn, E., and D'esposito, M. (1998). The variability of human, bold hemodynamic responses. *Neuroimage*, 8(4):360–369.

Ahlstrom, U. and Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36(7):623–636.

Alazard, D. (2005). Introduction to kalman filtering. *SUPAERO*.

Allison, B., Faller, J., and Neuper, C. (2012). Bcis that use steady-state visual evoked potentials or slow cortical potentials. *Brain-Computer Interfaces: Principles and Practice*, pages 241–250.

Allison, B. Z. and Neuper, C. (2010). Could anyone use a bci? In *Brain-computer interfaces*, pages 35–54. Springer.

Allison, B. Z. and Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological psychology*, 77(3):277–283.

Amiri, S., Fazel-Rezai, R., and Asadpour, V. (2013). A review of hybrid brain-computer interface systems. *Advances in Human-Computer Interaction*, 2013:1.

Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American Psychologist*, 51(4):355.

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, 18(3):251–270.

Antonenko, P., Paas, F., Grabner, R., and van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, 22(4):425–438.

Appel, G. (2005). *Technical analysis: power tools for active investors*. FT Press.

Aqil, M., Hong, K.-S., Jeong, M.-Y., and Ge, S. S. (2012). Cortical brain imaging by adaptive filtering of nirs signals. *Neuroscience letters*, 514(1):35–41.

Arnsten, A. F. (2000). Stress impairs prefrontal cortical function in rats and monkeys: role of dopamine d1 and norepinephrine α-1 receptor mechanisms. *Progress in brain research*, 126:183–192.

Attwell, D. and Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in neurosciences*, 25(12):621–625.

Ayaz, H., Cakir, M., Izzetoglu, K., Curtin, A., Shewokis, P. A., Bunce, S. C., and Onaral, B. (2012a). Monitoring expertise development during simulated uav piloting tasks using optical brain imaging. In *Aerospace Conference, 2012 IEEE*, pages 1–11. IEEE.

Ayaz, H., Izzetoglu, M., Bunce, S., Heiman-Patterson, T., and Onaral, B. (2007). Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy. In *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, pages 342–345. IEEE.

Ayaz, H. and Onaral, B. (2005). Analytical software and stimulus-presentation platform to utilize, visualize and analyze near-infrared spectroscopy measures. Master's thesis, Master's Degree Thesis, Drexel University.

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., and Onaral, B. (2012b). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59(1):36–47.

Ayaz, H., Willems, B., Bunce, S., Shewokis, P. A., Izzetoglu, K., Hah, S., Deshmukh, A., and Onaral, B. (2011). Estimation of cognitive workload during simulated air traffic control using optical brain imaging sensors. In *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*, pages 549–558. Springer.

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

Baddeley, A. D. and Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, 8:47–89.

Baird, B., Smallwood, J., and Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Consciousness and cognition*, 20(4):1604–1611.

BALPA (2013). Half of pilots have fallen asleep while flying. `https://www.balpa.org/News-and-campaigns/News/HALF-OF-PILOTS-HAVE-FALLEN-ASLEEP-WHILE-FLYING.aspx`. Accessed: 2015-07-15.

Banich, M. T. (2009). Executive function the search for an integrated account. *Current Directions in Psychological Science*, 18(2):89–94.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., and Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia*, 35(10):1373–1380.

Bashashati, A., Fatourechi, M., Ward, R. K., and Birch, G. E. (2007). A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals. *Journal of Neural engineering*, 4(2):R32.

Bedny, G. Z. and Karwowski, W. (2004). Activity theory as a basis for the study of work. *Ergonomics*, 47(2):134–153.

Béguin, P. and Clot, Y. (2004). Situated action in the development of activity. *revue électronique*, page 50.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., and Craven, P. L. (2007). Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, 78(Supplement 1):B231–B244.

Beverina, F., Palmas, G., Silvoni, S., Piccione, F., and Giove, S. (2003). User adaptive bcis: Ssvep and p300 based interfaces. *PsychNology Journal*, 1(4):331–354.

Birbaumer, N., Elbert, T., Rockstroh, B., and Lutzenberger, W. (1981). Biofeedback of event-related slow potentials of the brain. *International Journal of Psychology*, 16(1-4):389–415.

Blankertz, B., Sanelli, C., Halder, S., Hammer, E., Kübler, A., Müller, K.-R., Curio, G., and Dickhaus, T. (2009). Predicting bci performance to study bci illiteracy. *BMC Neuroscience*, 10(Suppl 1):P84.

Blankertz, B., Sannelli, C., Halder, S., Hammer, E. M., Kübler, A., Müller, K.-R., Curio, G., and Dickhaus, T. (2010). Neurophysiological predictor of smr-based bci performance. *Neuroimage*, 51(4):1303–1309.

Block, R. A., Hancock, P. A., and Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta psychologica*, 134(3):330–343.

Borisoff, J. F., Mason, S. G., and Birch, G. E. (2006). Brain interface research for asynchronous control applications. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, 14(2):160–164.

Bos, D. P.-O., Reuderink, B., van de Laar, B., Gürkök, H., Mühl, C., Poel, M., Nijholt, A., and Heylen, D. (2010). Brain-computer interfacing and games. In *Brain-Computer Interfaces*, pages 149–178. Springer.

Boucsein, W. (2000). *Engineering psychophysiology: issues and applications*. CRC Press.

Boyer, M. W. (2014). *The impact of rapid transItions in mental workload in a supervisory control setting*. PhD thesis, Massachusetts Institute of Technology.

Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human v1. *The journal of neuroscience*, 16(13):4207–4221.

Braboszcz, C. and Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage*, 54(4):3040–3047.

Bramble, D. M. and Lieberman, D. E. (2004). Endurance running and the evolution of homo. *Nature*, 432(7015):345–352.

Brehmer, B. and Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2):171–184.

Bromiley, M. (2008). Have you ever made a mistake. *Royal College of Anaesthetists Bulletin*, 48:2442–5.

Brouwer, A.-M., Van Erp, J., Heylen, D., Jensen, O., and Poel, M. (2013). Effortless passive bcis for healthy users. In *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion*, pages 615–622. Springer.

Brünken, R., Plass, J. L., and Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science*, 32(1-2):115–132.

Bunce, S. C., Izzetoglu, K., Ayaz, H., Shewokis, P., Izzetoglu, M., Pourrezaei, K., and Onaral, B. (2011). Implementation of fnirs for monitoring levels of expertise and mental workload. In *Foundations of augmented cognition. Directing the future of adaptive systems*, pages 13–22. Springer.

Buxton, R. B., Uludağ, K., Dubowitz, D. J., and Liu, T. T. (2004a). Modeling the hemodynamic response to brain activation. *Neuroimage*, 23:S220–S233.

Buxton, R. B., Uludağ, K., Dubowitz, D. J., and Liu, T. T. (2004b). Modeling the hemodynamic response to brain activation. *Neuroimage*, 23:S220–S233.

Buxton, R. B., Wong, E. C., and Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic resonance in medicine*, 39(6):855–864.

Callan, D., Gamez, M., Cassel, D., Terzibas, C., Callan, A., Kawato, M., and Sato, M. (2011). Dynamic visuomotor transformation involved with remote flying of a plane utilizes the'mirror neuron'system. *PloS one*, 7(4):e33873–e33873.

Callan, D. E., Durantin, G., and Terzibas, C. (2015). Classification of single-trial auditory events using dry-wireless eeg during real and motion simulated flight. *Frontiers in systems neuroscience*, 9.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., Goldberg, T. E., and Weinberger, D. R. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional mri. *Cerebral Cortex*, 9(1):20–26.

Carpenter, P., Just, M., Keller, T., Eddy, W., and Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Cognitive Neuroscience, Journal of*, 11(1):9–24.

Casner, S. M. and Schooler, J. W. (2013). Thoughts in flight automation use and pilots' task-related and task-unrelated thought. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, page 0018720813501550.

Causse, M., Dehais, F., Arexis, M., and Pastor, J. (2011a). Cognitive aging and flight performances in general aviation pilots. *Aging, Neuropsychology, and Cognition*, 18(5):544–561.

Causse, M., Dehais, F., and Pastor, J. (2011b). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *The International Journal of Aviation Psychology*, 21(3):217–234.

Chen, S., Epps, J., and Chen, F. (2011). A comparison of four methods for cognitive load measurement. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, pages 76–79. ACM.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fmri reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences*, 106(21):8719–8724.

Coffey, E. B., Brouwer, A.-M., Wilschut, E. S., and van Erp, J. B. (2010). Brain–machine interfaces in space: using spontaneous rather than intentionally generated brain signals. *Acta Astronautica*, 67(1):1–11.

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

Correa, A. G., Laciar, E., Patiño, H., and Valentinuzzi, M. (2007). Artifact removal from eeg signals using adaptive filters in cascade. In *Journal of Physics: Conference Series*, volume 90, page 012081. IOP Publishing.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Coyle, S., Ward, T., Markham, C., and McDarby, G. (2004). On the suitability of near-infrared (nir) systems for next-generation brain–computer interfaces. *Physiological measurement*, 25(4):815.

Coyle, S. M., Ward, T. E., and Markham, C. M. (2007). Brain–computer interface using a simplified functional near-infrared spectroscopy system. *Journal of neural engineering*, 4(3):219.

Craik, K. J. (1948). Theory of the human operator in control systems. *British Journal of Psychology. General Section*, 38(3):142–148.

Crawford, J. R. and Garthwaite, P. H. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a bayesian approach. *Cognitive Neuropsychology*, 24(4):343–372.

Cui, X., Bray, S., and Reiss, A. L. (2010). Functional near infrared spectroscopy (nirs) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage*, 49(4):3039–3046.

Cummings, M. L. and Mitchell, P. J. (2008). Predicting controller capacity in supervisory control of multiple uavs. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 38(2):451–460.

Cutrell, E. and Tan, D. (2008). Bci for passive input in hci. In *Proceedings of CHI*, volume 8, pages 1–3. Citeseer.

Dal Seno, B., Matteucci, M., and Mainardi, L. (2010). Online detection of p300 and error potentials in a bci speller. *Computational intelligence and neuroscience*, 2010:11.

de Carvalho, J. L. A., da Rocha, A. F., de Oliveira Nascimento, F. A., Neto, J. S., and Junqueira Jr, L. F. (2002). Development of a matlab software for analysis of heart rate variability. In *Signal Processing, 2002 6th International Conference on*, volume 2, pages 1488–1491. IEEE.

de Greef, T., Lafeber, H., van Oostendorp, H., and Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. In *Foundations of augmented cognition. Neuroergonomics and operational neuroscience*, pages 219–228. Springer.

De Keyser, V. and Woods, D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In *Systems reliability assessment*, pages 231–251. Springer.

de Visser, E. and Parasuraman, R. (2011). Adaptive aiding of human-robot teaming effects of imperfect automation on performance, trust, and workload. *Journal of Cognitive Engineering and Decision Making*, 5(2):209–231.

Dehais, F., Causse, M., Régis, N., Menant, E., Labedan, P., Vachon, F., and Tremblay, S. (2012). Missing critical auditory alarms in aeronautics: Evidence for inattentional deafness? volume 56, pages 1639–1643. SAGE Publications.

Dehais, F., Causse, M., and Tremblay, S. (2011). Mitigation of conflicts with automation use of cognitive countermeasures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(5):448–460.

Dehais, F., Tessier, C., and Chaudron, L. (2003). GHOST: experimenting conflicts countermeasures in the pilot's activity. In *IJCAI*, volume 3, pages 164–168.

Dehais, F., Tessier, C., Christophe, L., and Reuzeau, F. (2010). The perseveration syndrome in the pilot's activity: guidelines and cognitive countermeasures. In *Human Error, Safety and Systems Development*, pages 68–80.

Dejours, C. (1995). Le facteur humain.

Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21.

Derosière, G., Mandrick, K., Dray, G., Ward, T. E., and Perrey, S. (2013). Nirs-measured prefrontal cortex activity in neuroergonomics: strengths and weaknesses. *Frontiers in human neuroscience*, 7.

Devoto, P., Flore, G., Saba, P., Fa, M., and Gessa, G. L. (2005). Stimulation of the locus coeruleus elicits noradrenaline and dopamine release in the medial prefrontal and parietal cortex. *Journal of neurochemistry*, 92(2):368–374.

Diamond, S. G., Huppert, T. J., Kolehmainen, V., Franceschini, M. A., Kaipio, J. P., Arridge, S. R., and Boas, D. A. (2005). Physiological system identification with the kalman filter in diffuse optical tomography. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 649–656. Springer.

Dietrich, A. (2003). Functional neuroanatomy of altered states of consciousness: the transient hypofrontality hypothesis. *Consciousness and cognition*, 12(2):231–256.

Dohrmann, K., Weisz, N., Schlee, W., Hartmann, T., and Elbert, T. (2007). Neurofeedback for treating tinnitus. *Progress in brain research*, 166:473–554.

Donoso, M., Collins, A. G., and Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191):1481–1486.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice*, volume 373. Springer Science & Business Media.

Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology*, 12(1):75–96.

Durantin, G., Dehais, F., and Delorme, A. (2015). Characterization of mind wandering using fnirs. *Frontiers in Systems Neuroscience*, 9.

Durantin, G., Gagnon, J.-F., Tremblay, S., and Dehais, F. (2014a). Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural brain research*, 259:16–23.

Durantin, G., Gagnon, J.-F., Tremblay, S., and Dehais, F. (2014b). Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural brain research*, 259:16–23.

Durantin, G., Scannella, S., Gateau, T., Delorme, A., and Dehais, F. (2014c). Moving average convergence divergence filter preprocessing for real-time event-related peak activity onset detection : Application to fnirs signals. In *Engineering in Medicine and Biology Society, 2014. EMBS 2014. 36th Annual International Conference of the IEEE*. IEEE.

Egner, T. and Gruzelier, J. H. (2003). Ecological validity of neurofeedback: modulation of slow wave eeg enhances musical performance. *Neuroreport*, 14(9):1221–1224.

Eoh, H. J., Chung, M. K., and Kim, S.-H. (2005). Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. *International Journal of Industrial Ergonomics*, 35(4):307–320.

Fazli, S., Mehnert, J., Steinbrink, J., Curio, G., Villringer, A., Müller, K.-R., and Blankertz, B. (2012a). Enhanced performance by a hybrid nirs–eeg brain computer interface. *Neuroimage*, 59(1):519–529.

Fazli, S., Mehnert, J., Steinbrink, J., Curio, G., Villringer, A., Müller, K.-R., and Blankertz, B. (2012b). Enhanced performance by a hybrid nirs–eeg brain computer interface. *Neuroimage*, 59(1):519–529.

Ferrez, P. W. and Millán, J. d. R. (2008). Simultaneous real-time detection of motor imagery and error-related potentials for improved bci accuracy. In *Proceedings of the 4th international brain-computer interface workshop and training course*, number LIDIAP-CONF-2008-019.

Fox, P. T., Raichle, M. E., Mintun, M. A., and Dence, C. (1988). Nonoxidative glucose consumption during focal physiologic neural activity. *Science*, 241(4864):462–464.

Franklin, G. F., Powell, J. D., and Emami-Naeini, A. (2006). Feedback control of dynamics systems. *Pretince Hall Inc.*

Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., and Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*, 66(12):2289–2294.

Friedland, R. P. and Iadecola, C. (1991). Roy and sherrington (1890) a centennial reexamination of "on the regulation of the blood-supply of the brain. *Neurology*, 41(1):10–10.

Friedman, J. H. (1997). On bias, variance, 0/1—loss, and the curse-of-dimensionality. *Data mining and knowledge discovery*, 1(1):55–77.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., and Turner, R. (1998). Event-related fmri: characterizing differential responses. *Neuroimage*, 7(1):30–40.

Friston, K. J., Frith, C. D., Turner, R., and Frackowiak, R. S. (1995). Characterizing evoked hemodynamics with fmri. *Neuroimage*, 2(2PA):157–165.

Fuchs, T., Birbaumer, N., Lutzenberger, W., Gruzelier, J. H., and Kaiser, J. (2003). Neurofeedback treatment for attention-deficit/hyperactivity disorder in children: a comparison with methylphenidate. *Applied psychophysiology and biofeedback*, 28(1):1–12.

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., and Owen, G. (1999). Flight deck automation issues. *The International Journal of Aviation Psychology*, 9(2):109–123.

Gagnon, J.-F., Durantin, G., Vachon, F., Causse, M., Tremblay, S., and Dehais, F. (2012). Anticipating human error before it happens: Towards a psychophysiological model for online prediction of mental workload. In *Proceedings of the human factors and ergonomics society Europe Chapter annual meeting*.

Galéra, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribéreau-Gayon, R., Masson, F., Bakiri, S., Gabaude, C., Fort, A., et al. (2012). Mind wandering and driving: responsibility case-control study. *BMJ: British Medical Journal*, 345.

Gan, J. Q. (2006). *Self-adapting BCI based on unsupervised learning*. na.

Gateau, T., Durantin, G., Lancelot, F., Scannella, S., and Dehais, F. (2015). Real-time state estimation in a flight simulator using fnirs. *PloS one*, 10(3).

George, L. and Lécuyer, A. (2010). An overview of research on" passive" brain-computer interfaces for implicit human-computer interaction. In *International Conference on Applied Bionics and Biomechanics ICABB 2010-Workshop W1" Brain-Computer Interfacing and Virtual Reality"*.

Georgiadis, S. D., Ranta-aho, P. O., Tarvainen, M. P., and Karjalainen, P. A. (2005a). Single-trial dynamical estimation of event-related potentials: a kalman filter-based approach. *Biomedical Engineering, IEEE Transactions on*, 52(8):1397–1406.

Georgiadis, S. D., Ranta-aho, P. O., Tarvainen, M. P., and Karjalainen, P. A. (2005b). Single-trial dynamical estimation of event-related potentials: a kalman filter-based approach. *Biomedical Engineering, IEEE Transactions on*, 52(8):1397–1406.

Gevins, A. and Smith, M. (2005). Neurocognitive function eeg measurement method and system. US Patent 6,947,790.

Gianaros, P. J., Van der Veen, F. M., and Jennings, J. R. (2004). Regional cerebral blood flow correlates with heart period and high-frequency heart period variability during working-memory tasks: Implications for the cortical and subcortical regulation of cardiac autonomic activity. *Psychophysiology*, 41(4):521–530.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., and Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10(2):252–269.

Giraudet, L., St-Louis, M.-E., Scannella, S., and Causse, M. (2015). P300 event-related potential as an indicator of inattentional deafness? *PLoS one*, 10:e0118556–e0118556.

Girouard, A., Solovey, E. T., and Jacob, R. J. (2013). Designing a passive brain computer interface using real time classification of functional near-infrared spectroscopy. *International Journal of Autonomous and Adaptive Communications Systems*, 6(1):26–44.

Goldberg, J. H. and Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24(6):631–645.

Goldberg, T. E., Berman, K. F., Fleming, K., Ostrem, J., Van Horn, J. D., Esposito, G., Mattay, V. S., Gold, J. M., and Weinberger, D. R. (1998). Uncoupling cognitive workload and prefrontal cortical physiology: a pet rcbf study. *Neuroimage*, 7(4):296–303.

Grandchamp, R., Braboszcz, C., and Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in psychology*, 5.

Granholm, E., Asarnow, R. F., Sarkin, A. J., and Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33(4):457–461.

Groenegress, C., Holzner, C., Guger, C., and Slater, M. (2010). Effects of p300-based bci use on reported presence in a virtual environment. *Presence: Teleoperators and virtual environments*, 19(1):1–11.

Grozea, C., Voinescu, C. D., and Fazli, S. (2011). Bristle-sensors—low-cost flexible passive dry eeg electrodes for neurofeedback and bci applications. *Journal of neural engineering*, 8(2):025008.

Gruzelier, J., Egner, T., and Vernon, D. (2006). Validating the efficacy of neurofeedback for optimising performance. *Progress in brain research*, 159:421–431.

Haapalainen, E., Kim, S., Forlizzi, J. F., and Dey, A. K. (2010). Psycho-physiological measures for assessing cognitive load. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 301–310. ACM.

Haddal, C. C. and Gertler, J. (2010). Homeland security: Unmanned aerial vehicles and border surveillance. DTIC Document.

Halas, P. and Pelaez, C. (2015). Le génie sans cerveau. *Le journal de Mickey*, 3267.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, 65(2):413.

Hancock, P. A. and Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *Systems, Man and Cybernetics, IEEE Transactions on*, 18(4):647–658.

Harrison, J., Izzetoglu, K., Ayaz, H., Willems, B., Hah, S., Woo, H., Shewokis, P. A., Bunce, S. C., and Onaral, B. (2013). Human performance assessment study in aviation using functional near infrared spectroscopy. In *Foundations of Augmented Cognition*, pages 433–442. Springer.

Harrivel, A. R., Weissman, D. H., Noll, D. C., and Peltier, S. J. (2013). Monitoring attentional state with fnirs. *Frontiers in human neuroscience*, 7.

Hart, C. S. (2010). Assessing the impact of low workload in supervisory control of networked unmanned vehicles. Technical report, DTIC Document.

Hart, S. G. and Sheridan, T. B. (1984). Pilot workload, performance, and aircraft control automation. Technical report, DTIC Document.

Hart, S. G. and Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183.

Hartmann, M. and Fischer, M. (2014). Pupillometry: The eyes shed fresh light on the mind. *Current Biology*, 24(7):R281 – R282.

Hasenkamp, W., Wilson-Mendenhall, C. D., Duncan, E., and Barsalou, L. W. (2012). Mind wandering and attention during focused meditation: a fine-grained temporal analysis of fluctuating cognitive states. *Neuroimage*, 59(1):750–760.

He, J., Becic, E., Lee, Y.-C., and McCarley, J. S. (2011). Mind wandering behind the wheel performance and oculomotor correlates. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(1):13–21.

Helton, W. S., Kern, R. P., and Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and cognition*, 18(3):600–607.

Hilburn, B. and Jorna, P. (2001). Workload and air traffic control. *Stress, Workload and Fatigue: Theory, Research and Practice. Hillsdale, New Jersey, USA: Erlbaum*.

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

Hirshfield, L. M., Gulotta, R., Hirshfield, S., Hincks, S., Russell, M., Ward, R., Williams, T., and Jacob, R. (2011). This is your brain on interfaces: enhancing usability testing with functional near-infrared spectroscopy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 373–382. ACM.

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., and Søgaard, 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.

Hockey, G. R. J. (2005). *Operator functional state: The prediction of breakdown in human performance*. Oxford, UK: OUP.

Hockey, R., Gaillard, A. W., and Coles, M. (1986). *Energetics and human information processing*, volume 31. Springer Science & Business Media.

Hoshi, Y., Tsou, B. H., Billock, V. A., Tanosaki, M., Iguchi, Y., Shimada, M., Shinba, T., Yamada, Y., and Oda, I. (2003). Spatiotemporal characteristics of hemodynamic changes in the human lateral prefrontal cortex during working memory tasks. *Neuroimage*, 20(3):1493–1504.

Hu, X.-S., Hong, K.-S., Ge, S. S., and Jeong, M.-Y. (2010). Kalman estimator-and general linear model-based on-line brain activation mapping by near-infrared spectroscopy. *Biomed. Eng. Online*, 9(1):82.

Hull, C. (1943). *Principles of behavior*. New York Appleton-century-crofts.

Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). Homer: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied optics*, 48(10):D280–D298.

ICAO (2014). Icao safety report. `http://asndata.aviation-safety.net/industry-reports/ICAO-Safety-Report-2014.pdf`. Accessed: 2015-07-15.

Iriarte, J., Urrestarazu, E., Valencia, M., Alegre, M., Malanda, A., Viteri, C., and Artieda, J. (2003). Independent component analysis as a tool to eliminate artifacts in eeg: a quantitative study. *Journal of clinical neurophysiology*, 20(4):249–257.

Izzetoglu, M., Bunce, S. C., Izzetoglu, K., Onaral, B., and Pourrezaei, K. (2007). Functional brain imaging using near-infrared technology. *IEEE Engineering in Medicine and Biology Magazine*, 26(4):38.

Izzetoglu, M., Chitrapu, P., Bunce, S., and Onaral, B. (2010). Motion artifact cancellation in nir spectroscopy using discrete kalman filtering. *Biomed Eng Online*, 9(1):16.

Jackson, M. M. and Mappus, R. (2010). Applications for brain-computer interfaces. In *Brain-Computer Interfaces*, pages 89–103. Springer.

Jain, A. K. and Chandrasekaran, B. (1982). 39 dimensionality and sample size considerations in pattern recognition practice. *Handbook of statistics*, 2:835–855.

Jang, K. E., Jeong, Y., Ye, J. C., Tak, S., Jung, J., and Jang, J. (2009). Wavelet minimum description length detrending for near-infrared spectroscopy. *Journal of biomedical optics*, 14(3):034004–034004.

Jasdzewski, G., Strangman, G., Wagner, J., Kwong, K., Poldrack, R., and Boas, D. (2003). Differences in the hemodynamic response to event-related motor and visual paradigms as measured by near-infrared spectroscopy. *Neuroimage*, 20(1):479–488.

Jentsch, F. and Bowers, C. A. (1998). Evidence for the validity of pc-based simulations in studying aircrew coordination. *The International Journal of Aviation Psychology*, 8(3):243–260.

Jin, J., Allison, B. Z., Kaufmann, T., Kübler, A., Zhang, Y., Wang, X., and Cichocki, A. (2012). The changing face of p300 bcis: A comparison of stimulus changes in a p300 bci involving faces, emotion, and movement. *PloS one*, 7(11):e49688.

Johnson, J. A. and Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage*, 31(4):1673–1681.

Kaber, D. B., Perry, C. M., Segall, N., and Sheik-Nainar, M. A. (2007). Workload state classification with automation during simulated air traffic control. *The International Journal of Aviation Psychology*, 17(4):371–390.

Kahneman, D. (1973). *Attention and effort*. Citeseer.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45.

Kamath, M. V. and Fallen, E. L. (1992). Power spectral analysis of heart rate variability: a noninvasive signature of cardiac autonomic function. *Critical reviews in biomedical engineering*, 21(3):245–311.

Kamran, M. A., Hong, K.-S., and Mannan, M. M. N. (2012). Identification of fnirs based brain activity using adaptive algorithm. In *Robotics and Artificial Intelligence (ICRAI), 2012 International Conference on*, pages 200–203. IEEE.

Kantowitz, B. H. and Knight, J. L. (1976). Testing tapping timesharing, ii: Auditory secondary task. *Acta Psychologica*, 40(5):343–362.

Kearney, P. R. and Pivec, M. (2007). Immersed and how? that is the question. In *Games in Action Conference. Gothenburg, Sweden*.

Kennedy, A. D. and Gehan, E. A. (1971). Computerized simple regression methods for survival time studies. *Computer Programs in Biomedicine*, 1(4):235–244.

Killingsworth, M. A. and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006):932–932.

Kim, B. H., Kim, M., and Jo, S. (2014). Quadcopter flight control using a low-cost hybrid interface with eeg-based classification and eye tracking. *Computers in biology and medicine*, 51:82–92.

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1655):20130474.

Koechlin, E., Corrado, G., Pietrini, P., and Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences*, 97(13):7651–7656.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648):1181–1185.

Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K.-R., Curio, G., Hagemann, K., Bruns, A., Schrauf, M., and Kincses, W. (2007). Improving human performance in a real operating environment through real-time mental workload detection. *Toward Brain-Computer Interfacing*, pages 409–422.

Kothe, C. A. and Makeig, S. (2013). Bcilab: a platform for brain–computer interface development. *Journal of neural engineering*, 10(5):056014.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. *Multiple-task performance*, pages 279–328.

Kramer, A. F., Trejo, L. J., and Humphrey, D. (1995). Assessment of mental workload with task-irrelevant auditory probes. *Biological Psychology*, 40(1):83–100.

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., and Birbaumer, N. (2004). Predictability of brain-computer communication. *Journal of Psychophysiology*, 18(2/3):121–129.

Kuhlman, W. N. (1978). Eeg feedback training of epileptic patients: clinical and electroencephalographic analysis. *Electroencephalography and clinical neurophysiology*, 45(6):699–710.

LaBerge, D. (1997). Attention, awareness, and the triangular circuit. *Consciousness and cognition*, 6(2):149–181.

LaBerge, D., Carter, M., and Brown, V. (1992). A network simulation of thalamic circuit operations in selective attention. *Neural Computation*, 4(3):318–331.

LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., and He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface. *Journal of neural engineering*, 10(4):046003.

Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., and Mcdarby, G. (2005). Steady-state vep-based brain-computer interface control in an immersive 3d gaming environment. *EURASIP journal on applied signal processing*, 2005:3156–3164.

Lavie, N., Beck, D. M., and Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1641):20130205.

Lavie, N. and Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, 56(2):183–197.

Lee, J. D. (2014). Dynamics of driver distraction: The process of engaging and disengaging. *Annals of advances in automotive medicine*, 58:24.

Leon-Carrion, J. and Leon-Dominguez, U. (2012). Functional near-infrared spectroscopy (fnirs): principles and neuroscientific applications. *Neuroimaging e Methods. ISBN*, pages 978–953.

Leontiev, A. N. (2014). Activity and consciousness. *Revista Dialectus*, (4).

Lin, K., Chen, X., Huang, X., Ding, Q., and Gao, X. (2015). A hybrid bci speller based on the combination of emg envelopes and ssvep. In *Applied Informatics*, volume 2, pages 1–12. Springer.

Liu, Y., Ayaz, H., Curtin, A., Onaral, B., and Shewokis, P. A. (2013). Towards a hybrid p300-based bci using simultaneous fnir and eeg. In *Foundations of Augmented Cognition*, pages 335–344. Springer.

Loft, S., Sanderson, P., Neal, A., and Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3):376–399.

Lotte, F., Congedo, M., Lécuyer, A., and Lamarche, F. (2007). A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4.

Lotte, F. and Guan, C. (2010). Learning from other subjects helps reducing brain-computer interface calibration time. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 614–617. IEEE.

Mack, A. and Rock, I. (1998). *Inattentional blindness.* The MIT Press.

Mackworth, N. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1(1):6–21.

Mandrick, K., Derosiere, G., Dray, G., Coulon, D., Micallef, J.-P., and Perrey, S. (2013). Prefrontal cortex activity during motor tasks with additional mental load requiring attentional demand: a near-infrared spectroscopy study. *Neuroscience research*, 76(3):156–162.

Manly, T., Robertson, I. H., Galloway, M., and Hawkins, K. (1999). The absent mind:: further investigations of sustained attention to response. *Neuropsychologia*, 37(6):661–670.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, 164(1):177–190.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007a). Wandering minds: the default network and stimulus-independent thought. *Science*, 315(5810):393–395.

Mason, S., Bashashati, A., Fatourechi, M., Navarro, K., and Birch, G. (2007b). A comprehensive survey of brain interface technology designs. *Annals of biomedical engineering*, 35(2):137–169.

McFarland, D. J., Anderson, C. W., Muller, K., Schlogl, A., and Krusienski, D. J. (2006). Bci meeting 2005-workshop on bci signal processing: feature extraction and translation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):135.

McVay, J. and Kane, M. (2010). Does mind wandering reflect executive function or executive failure? comment on smallwood and schooler (2006) and watkins (2008). *Psychological bulletin*, 136(2):188–97.

Menda, J., Hing, J. T., Ayaz, H., Shewokis, P. A., Izzetoglu, K., Onaral, B., and Oh, P. (2011). Optical brain imaging to enhance uav operator training, evaluation, and interface development. *Journal of intelligent & robotic systems*, 61(1-4):423–443.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24(1):167–202.

Miyake, S. (2001). Multivariate workload evaluation combining physiological and subjective measures. *International journal of psychophysiology*, 40(3):233–238.

Mooneyham, B. W. and Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(1):11.

Moore, M. M. (2003). Real-world applications for brain-computer interface technology. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):162–165.

Moray, N. (1967). Where is capacity limited? a survey and a model. *Acta psychologica*, 27:84–92.

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Cauwenberghs, G., and Jung, T.-P. (2013). Real-time modeling and 3d visualization of source dynamics and connectivity using wearable eeg. In *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, volume 2013, page 2184. NIH Public Access.

Müller, K.-R., Krauledat, M., Dornhege, G., Curio, G., and Blankertz, B. (2004). Machine learning techniques for brain-computer interfaces. *Biomed. Tech*, 49(1):11–22.

Murai, K., Hayashi, Y., Okazaki, T., Stone, L. C., and Mitomo, N. (2008). Evaluation of ship navigator's mental workload using nasal temperature and heart rate variability. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, pages 1528–1533. IEEE.

Näätänen, R. and Picton, T. (1987). The n1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4):375–425.

Nagaoka, T., Sakatani, K., Awano, T., Yokose, N., Hoshino, T., Murata, Y., Katayama, Y., Ishikawa, A., and Eda, H. (2010). Development of a new rehabilitation system based on a brain-computer interface using near-infrared spectroscopy. In *Oxygen transport to tissue XXXI*, pages 497–503. Springer.

Naito, M., Michioka, Y., Ozawa, K., KIGUCHI, M., and KANAZAWA, T. (2007). A communication means for totally locked-in als patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE transactions on information and systems*, 90(7):1028–1037.

Napadow, V., Dhond, R., Conti, G., Makris, N., Brown, E. N., and Barbieri, R. (2008). Brain correlates of autonomic modulation: combining heart rate variability with fmri. *Neuroimage*, 42(1):169–177.

Naseer, N. and Hong, K.-S. (2015a). Decoding answers to four-choice questions using functional near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 23(1):23–31.

Naseer, N. and Hong, K.-S. (2015b). fnirs-based brain-computer interfaces: a review. *Frontiers in human neuroscience*, 9.

Navon, D. and Gopher, D. (1979). On the economy of the human-processing system. *Psychological review*, 86(3):214.

Nishimura, E. M., Rapoport, E. D., Wubbels, P. M., Downs, T. H., and Downs III, J. H. (2010). Functional near-infrared sensing (fnir) and environmental control applications. In *Brain-Computer Interfaces*, pages 121–132. Springer.

Norman, D. A. and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive psychology*, 7(1):44–64.

Norman, D. A. and Shallice, T. (1986). *Attention to action*. Springer.

O'Donnell, R. D. and Eggemeier, F. T. (1986). Workload assessment methodology. In *Handbook of perception and human performance*, volume 2, pages 1–49. John Wiley & Sons.

Ogawa, S., Menon, R., Tank, D., Kim, S., Merkle, H., Ellermann, J., and Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. a comparison of signal characteristics with a biophysical model. *Biophysical journal*, 64(3):803.

Oron-Gilad, T., Chen, J., and Hancock, P. (2006). Remotely operated vehicles (rovs) from the top–down and the bottom–up. *Advances in Human Performance and Cognitive Engineering Research*, 7:37–47.

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

Pan, S. J. and Yang, Q. (2010). A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on*, 22(10):1345–1359.

Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theoretical Issues in Ergonomics Science*, 4(1-2):5–20.

Parasuraman, R., Mouloua, M., and Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. *Automation technology and human performance: Current research and trends*, pages 119–123.

Parasuraman, R. and Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39(2):230–253.

Parasuraman, R. and Rizzo, M. (2007). *Neuroergonomics: The brain at work*. Oxford University Press, USA.

Parasuraman, R. and Wilson, G. F. (2008). Putting the brain to work: Neuroergonomics past, present, and future. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):468–474.

Parini, S., Maggi, L., Turconi, A. C., and Andreoni, G. (2009). A robust and self-paced bci system based on a four class ssvep paradigm: algorithms and protocols for a high-transfer-rate direct brain communication. *Computational Intelligence and Neuroscience*, 2009.

Pasley, B. N. and Freeman, R. D. (2008). Neurovascular coupling. *Scholarpedia*, 3(3):5340. revision 91581.

Pellerin, L. and Magistretti, P. J. (1994). Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proceedings of the National Academy of Sciences*, 91(22):10625–10629.

Perrey, S. (2008). Non-invasive nir spectroscopy of human brain function during exercise. *Methods*, 45(4):289–299.

Peysakhovich, V., Causse, M., Scannella, S., and Dehais, F. (2015). Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort. *International Journal of Psychophysiology*, 97(1):30–37.

Pfurtscheller, G. (2001). Functional brain imaging based on erd/ers. *Vision research*, 41(10):1257–1260.

Pfurtscheller, G., Allison, B. Z., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T. O., Mueller-Putz, G., Neuper, C., and Birbaumer, N. (2010a). The hybrid bci. *Frontiers in neuroscience*, 4:3.

Pfurtscheller, G., Allison, B. Z., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T. O., Mueller-Putz, G., Neuper, C., and Birbaumer, N. (2010b). The hybrid bci. *Frontiers in neuroscience*, 4:3.

Pfurtscheller, G. and Da Silva, F. L. (1999). Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857.

Pinet, J. (2011). *Traitement de situations inattendues d'extrême urgence en vol: test d'un modèle cognitif auprès de pilotes experts*. PhD thesis, Université Toulouse le Mirail-Toulouse II.

PlaneCrashInfo.com (2015). Accident statistics. `http://www.planecrashinfo.com/cause.htm`. Accessed: 2015-07-15.

Polich, J. (2007). Updating p300: an integrative theory of p3a and p3b. *Clinical neurophysiology*, 118(10):2128–2148.

Pomeranz, B., Macaulay, R., Caudill, M. A., Kutz, I., Adam, D., Gordon, D., Kilborn, K. M., Barger, A. C., Shannon, D. C., Cohen, R. J., et al. (1985). Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology*, 248(1):H151–H153.

Poole, A. and Ball, L. J. (2006). Eye tracking in hci and usability research. *Encyclopedia of human computer interaction*, 1:211–219.

Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1):187–195.

Popescu, F., Blankertz, B., and Müller, K.-R. (2008). Computational challenges for noninvasive brain computer interfaces. *IEEE Intelligent Systems*, 23(3):78–79.

Popescu, F., Fazli, S., Badower, Y., Blankertz, B., and Müller, K.-R. (2007). Single trial classification of motor imagination using 6 dry eeg electrodes. *PloS one*, 2(7):e637.

Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1):3–25.

Power, S. D., Kushki, A., and Chau, T. (2012). Intersession consistency of single-trial classification of the prefrontal response to mental arithmetic and the no-control state by nirs. *PloS one*, 7(7):e37791.

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., and Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *The International journal of aviation psychology*, 10(4):393–410.

Qin, S., Hermans, E. J., van Marle, H. J., Luo, J., and Fernández, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological psychiatry*, 66(1):25–32.

Quiroga, R. Q. and Garcia, H. (2003). Single-trial event-related potentials with wavelet denoising. *Clinical Neurophysiology*, 114(2):376–390.

Raichlen, D. A., Foster, A. D., Gerdeman, G. L., Seillier, A., and Giuffrida, A. (2012). Wired to run: exercise-induced endocannabinoid signaling in humans and cursorial mammals with implications for the 'runner's high'. *The Journal of experimental biology*, 215(8):1331–1336.

Rakotomamonjy, A., Guigue, V., Mallet, G., and Alvarado, V. (2005). Ensemble of svms for improving brain computer interface p300 speller performances. In *Artificial Neural Networks: Biological Inspirations–ICANN 2005*, pages 45–50. Springer.

Rani, P., Sarkar, N., and Liu, C. (2005). Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th International Conference on Human Computer Interaction*, volume 58.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *Systems, Man and Cybernetics, IEEE Transactions on*, 13(3):257–266.

Raya, M. A. D. and Sison, L. G. (2002). Adaptive noise cancelling of motion artifact in stress ecg signals using accelerometer. In *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, volume 2, pages 1756–1757. IEEE.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 327(1241):475–484.

Rees, G., Frith, C. D., and Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, 278(5343):1616–1619.

Regis, N., Dehais, F., Rachelson, E., Thooris, C., Pizziol, S., Causse, M., and Tessier, C. (2014). Formal detection of attentional tunneling in human operator. *Human-Machine Systems, IEEE Transactions on*, 44(3):326–336.

Reichle, E. D., Reineberg, A. E., and Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, 21(9):1300–1310.

Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments*, 19(1):35–53.

Ridderinkhof, K. R., van den Wildenberg, W. P., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and cognition*, 56(2):129–140.

Rohmert, W. (1986). Ergonomics: Concept of work, stress and strain. *Applied Psychology*, 35(2):159–180.

Rossi, A. F., Bichot, N. P., Desimone, R., and Ungerleider, L. G. (2007). Top–down attentional deficits in macaques with lesions of lateral prefrontal cortex. *The Journal of Neuroscience*, 27(42):11306–11314.

Rowe, J. B. and Passingham, R. E. (2001). Working memory for location and time: activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage*, 14(1):77–86.

Roy, C. S. and Sherrington, C. S. (1890). On the regulation of the blood-supply of the brain. *The Journal of physiology*, 11(1-2):85–158.

Rushworth, M., Passingham, R., and Nobre, A. (2002). Components of switching intentional set. *Cognitive Neuroscience, Journal of*, 14(8):1139–1150.

Sacks, O. (1998). *The man who mistook his wife for a hat: And other clinical tales*. Simon and Schuster.

Sämann, P. G., Wehrle, R., Hoehn, D., Spoormaker, V. I., Peters, H., Tully, C., Holsboer, F., and Czisch, M. (2011). Development of the brain's default mode network from wakefulness to slow wave sleep. *Cerebral cortex*, page bhq295.

Sannelli, C., Braun, M., Tangermann, M., and Müller, K.-R. (2008). *Estimating noise and dimensionality in BCI data sets: Towards illiteracy comprehension*. Citeseer.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., and Wolpaw, J. R. (2004). Bci2000: a general-purpose brain-computer interface (bci) system. *Biomedical Engineering, IEEE Transactions on*, 51(6):1034–1043.

Schreppel, T., Egetemeir, J., Schecklmann, M., Plichta, M. M., Pauli, P., Ellgring, H., Fallgatter, A. J., and Herrmann, M. J. (2007). Activation of the prefrontal cortex in working memory and interference resolution processes assessed with near-infrared spectroscopy. *Neuropsychobiology*, 57(4):188–193.

Schudlo, L. C. and Chau, T. (2014). Dynamic topographical pattern classification of multichannel prefrontal nirs signals: Ii. online differentiation of mental arithmetic and rest. *Journal of neural engineering*, 11(1):016003.

Schulz, C., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Wasmaier, M., Kochs, E., and Schneider, G. (2011). Eye tracking for assessment of workload: a pilot study in an anaesthesia simulator environment. *British journal of anaesthesia*, 106(1):44–50.

Segal, L. D. and Wickens, C. D. (1990). Taskillan ii: Pilot strategies for workload management. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 34, pages 66–70. SAGE Publications.

Shappel, S. A. and Wiegmann, D. A. (2000). The human factors analysis and classification system–hfacs. Technical report, US Federal Aviation Administration, Office of Aviation Medicine.

Shappell, S. A. and Wiegmann, D. A. (2012). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Ashgate Publishing, Ltd.

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., and Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both "go" and "no-go" response-based vigilance tasks. *Brain and cognition*, 82(3):265–273.

Sheridan, T. B. (2011). Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: Distinctions and modes of adaptation. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41(4):662–667.

Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., Shimizu, K., and Birbaumer, N. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface. *NeuroImage*, 34(4):1416–1427.

Smallwood, J., Beach, E., Schooler, J. W., and Handy, T. C. (2008). Going awol in the brain: Mind wandering reduces cortical analysis of external events. *Journal of cognitive neuroscience*, 20(3):458–469.

Smallwood, J., Brown, K., Baird, B., and Schooler, J. W. (2012a). Cooperation between the default mode network and the frontal–parietal network in the production of an internal train of thought. *Brain research*, 1428:60–70.

Smallwood, J., Brown, K. S., Baird, B., Mrazek, M. D., Franklin, M. S., and Schooler, J. W. (2012b). Insulation for daydreams: a role for tonic norepinephrine in the facilitation of internally guided thought. *PloS one*, 7(4):e33706–e33706.

Smallwood, J. and Schooler, J. W. (2006). The restless mind. *Psychological bulletin*, 132(6):946.

Smilek, D., Carriere, J. S., and Cheyne, J. A. (2010). Out of mind, out of sight eye blinking as indicator and embodiment of mind wandering. *Psychological Science*, 21(6):786–789.

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

Sokoloff, L. (1993). Sites and mechanisms of function-related changes in energy metabolism in the nervous system. *Developmental neuroscience*, 15(3-5):194–206.

Solovey, E., Schermerhorn, P., Scheutz, M., Sassaroli, A., Fantini, S., and Jacob, R. (2012). Brainput: enhancing interactive systems with streaming fnirs brain input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2193–2202. ACM.

Solovey, E. T., Lalooses, F., Chauncey, K., Weaver, D., Parasi, M., Scheutz, M., Sassaroli, A., Fantini, S., Schermerhorn, P., Girouard, A., et al. (2011). Sensing cognitive multitasking for a brain-based adaptive user interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 383–392. ACM.

Spence, K. W. (1958). A theory of emotionally based drive (d) and its relation to performance in simple learning situations. *American Psychologist*, 13(4):131.

Sperandio, J. (1971). Variation of operator's strategies and regulating effects on workload. *Ergonomics*, 14(5):571–577.

Spreng, R. N. (2012). The fallacy of a "task-negative" network. *Frontiers in psychology*, 3.

St. John, M., Kobus, D. A., Morrison, J. G., and Schmorrow, D. (2004). Overview of the darpa augmented cognition technical integration experiment. *International Journal of Human-Computer Interaction*, 17(2):131–149.

Stassen, H. G., Johannsen, G., and Moray, N. (1990). Internal representation, internal model, human performance model and mental workload. *Automatica*, 26(4):811–820.

Stawarczyk, D., Majerus, S., Maquet, P., and D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS One*, 6(2):e16997–e16997.

Steriade, M., Gloor, P., Llinas, R., Da Silva, F. L., and Mesulam, M.-M. (1990). Basic mechanisms of cerebral rhythmic activities. *Electroencephalography and clinical neurophysiology*, 76(6):481–508.

Stevens, R., Galloway, T., and Berka, C. (2007). Allocation of time, eeg-engagement and eeg-workload resources as scientific problem solving skills are acquired in the classroom. *Proceedings of 3rd Augmented Cognition International, held in conjunction with HCI International*, pages 22–27.

Stevenson, H., Russell, P. N., and Helton, W. S. (2011). Search asymmetry, sustained attention, and response inhibition. *Brain and cognition*, 77(2):215–222.

Stokes, A., Wickens, C., and Kite, K. (1990). Display technology-human factors concepts. *NASA STI/Recon Technical Report A*, 91:27333.

Sugiyama, M. (2010). Superfast-trainable multi-class probabilistic classifier by least-squares posterior fitting. *IEICE Transactions on Information and Systems*, 93(10):2690–2701.

Sullivan, S. E. and Bhagat, R. S. (1992). Organizational stress, job satisfaction and job performance: where do we go from here? *Journal of Management*, 18(2):353–374.

Tai, K. and Chau, T. (2009). Single-trial classification of nirs signals during emotional induction tasks: towards a corporeal machine interface. *Journal of neuroengineering and rehabilitation*, 6:39.

Takeuchi, Y. (2000). Change in blood volume in the brain during a simulated aircraft landing task. *Journal of occupational health*, 42(2):60–65.

Taylor, G., Reinerman-Jones, L., Cosenzo, K., and Nicholson, D. (2010). Comparison of multiple physiological sensors to classify operator state in adaptive automation systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 54, pages 195–199. Sage Publications.

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., Rosen, A. C., and Yesavage, J. A. (2005). Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychology and aging*, 20(1):117–133.

Teixeira, A., Tome, A., Lang, E., and Martins da Silva, A. (2008). Subspace techniques to remove artifacts from eeg: A quantitative analysis. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4395–4398. IEEE.

Thatcher, R., McAlaster, R., Lester, M., Horst, R., and Cantor, D. (1983). Hemispheric eeg asymmetries related to cognitive functioning in children. *Cognitive processing in the right hemisphere*, pages 125–146.

Thayer, J. F. and Lane, R. D. (2009). Claude bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews*, 33(2):81–88.

Thomas, C. (1977). *Firefox*. Holt, Rinehart and Winston.

Thurlings, M. E., van Erp, J. B., Brouwer, A.-M., and Werkhoven, P. J. (2010). Eeg-based navigation from a human factors perspective. In *Brain-Computer Interfaces*, pages 71–86. Springer.

Tomita, Y. and Mitsukura, Y. (2013). Hemodynamic characteristics for improvement of eeg-bci performance. In *Human System Interaction (HSI), 2013 The 6th International Conference on*, pages 495–500. IEEE.

Tomita, Y., Vialatte, F.-B., Dreyfus, G., Mitsukura, Y., Bakardjian, H., and Cichocki, A. (2014). Bimodal bci using simultaneously nirs and eeg. *Transactions on Biomedical Engineering*.

Travis, T., Kondo, C., and Knott, J. (1975). Alpha enhancement research: a review. *Biological psychiatry*, 10(1):69–89.

Trejo, L. J., Kochavi, R., Kubitz, K., Montgomery, L. D., Rosipal, R., and Matthews, B. (2005). Measures and models for predicting cognitive fatigue. In *Defense and Security*, pages 105–115. International Society for Optics and Photonics.

Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., and Jung, T.-P. (2007). Task performance and eye activity: predicting behavior relating to cognitive workload. *Aviation, space, and environmental medicine*, 78(Supplement 1):B176–B185.

Utsugi, K., Obata, A., Sato, H., Katsura, T., Sagara, K., Maki, A., and Koizumi, H. (2007a). Development of an optical brain-machine interface. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5338–5341. IEEE.

Utsugi, K., Obata, A., Sato, H., Katsura, T., Sagara, K., Maki, A., and Koizumi, H. (2007b). Development of an optical brain-machine interface. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5338–5341. IEEE.

Uzzaman, S. and Joordens, S. (2011). The eyes know what you are thinking: eye movements as an objective measure of mind wandering. *Consciousness and cognition*, 20(4):1882–1886.

Van der Zee, P., Arridge, S., Cope, M., and Delpy, D. (1990). The effect of optode positioning on optical pathlength in near infrared spectroscopy of brain. In *Oxygen Transport to Tissue XII*, pages 79–84. Springer.

van Erp, J. B. F., Lotte, F., and Tangermann, M. (2012). Brain-computer interfaces: Beyond medical applications. *Computer*, 45(4):26–34.

Van Snellenberg, J. X., Slifstein, M., Read, C., Weber, J., Thompson, J. L., Wager, T. D., Shohamy, D., Abi-Dargham, A., and Smith, E. E. (2015). Dynamic shifts in brain network activation during supracapacity working memory task performance. *Human brain mapping*, 36(4):1245–1264.

Vaughan Jr, H. G. and Arezzo, J. C. (1988). The neural basis of event-related potentials. *Human event-related potentials*, 3:45–96.

Venditti, R. and Weldele, B. (2012). *Surrogates Volume 1*, volume 1. Top Shelf Productions.

Vidal, J.-J. (1973). Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*, 2(1):157–180.

Vidaurre, C. and Blankertz, B. (2010). Towards a cure for bci illiteracy. *Brain topography*, 23(2):194–198.

Villringer, A. and Obrig, H. (2002). Near infrared spectroscopy and imaging. *Brain mapping: The methods*, pages 141–158.

von Bünau, P., Meinecke, F. C., Király, F. C., and Müller, K.-R. (2009). Finding stationary subspaces in multivariate time series. *Physical Review Letters*, 103(21):214101.

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3):1195–1268.

Wassenaar, E. and Van den Brand, J. (2005). Reliability of near-infrared spectroscopy in people with dark skin pigmentation. *Journal of clinical monitoring and computing*, 19(3):195–199.

Wei, B., Wu, X., and Zhang, D. (2009). Design and implementation of an improved bci system based on alpha rhythm. In *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*, pages 1–4. IEEE.

Weibel, R. E. and Hansman, R. J. (2006). Safety considerations for operation of unmanned aerial vehicles in the national airspace system. Technical report, NASA.

Wickens, C., Kramer, A., Vanasse, L., and Donchin, E. (1983a). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science*, 221(4615):1080–1082.

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance*, 2(1):1.

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII*, 8.

Wickens, C. D. (1981). Processing resources in attention, dual task performance, and workload assessment. Technical report, DTIC Document.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):449–455.

Wickens, C. D., Huey, B. M., et al. (1993). *Workload Transition:: Implications for Individual and Team Performance*. National Academies Press.

Wickens, C. D., Sandry, D. L., and Vidulich, M. (1983b). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 25(2):227–248.

Wierwille, W. W. and Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(2):263–281.

Wilson, G. F. and Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human factors: the journal of the human factors and ergonomics society*, 49(6):1005–1018.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791.

Wolter, C. A. and Gore, B. F. (2015). A validated task analysis of the single pilot operations concept. *Moffett Field, CA: Ames Research Center*.

Woods, D. D. (2000). Behind human error: Human factors research to improve patient safety. *Washington, DC: American Psychological Association*.

Woods, D. D., Johannesen, L. J., Cook, R. I., and Sarter, N. B. (1994). Behind human error: Cognitive systems, computers and hindsight. Technical report, DTIC Document.

Xiaomeng, C. (2011). A new real-time ecg r-wave detection algorithm. In *Strategic Technology (IFOST), 2011 6th International Forum on*, volume 2, pages 1252–1255. IEEE.

Yerkes, R. M. and 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.

YouGov.com (2015). Most americans now worried about flying. `https://today.yougov.com/news/2015/04/01/most-americans-now-worried-about-flying/`. Accessed: 2015-07-15.

Young, M. S. and Stanton, N. A. (2002). Attention and automation: new perspectives on mental underload and performance. *Theoretical Issues in Ergonomics Science*, 3(2):178–194.

Yu, H., Schummers, J., and Sur, M. (2010). The influence of astrocyte activation on hemodynamic signals for functional brain imaging. In *Imaging the Brain with Optical Methods*, pages 45–64. Springer.

Zander, T. O. and Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, 8(2):025005.

Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces*, pages 181–199. Springer.

Zhang, Q., Strangman, G. E., and Ganis, G. (2009). Adaptive filtering to reduce global interference in non-invasive nirs measures of brain activation: how well and when does it work? *Neuroimage*, 45(3):788–794.